UNIVERSITAT POLITÈCNICA
DE CATALUNYA

Ph.D. Thesis

Channel State Information and Joint Transmitter-Receiver Design in Multi-Antenna Systems

Author: Antonio Pascual Iserte

Advisors: Dr. Miguel Ángel Lagunas Hernández
Dr. Ana I. Pérez Neira

Array and Multichannel Processing Group
Department of Signal Theory and Communications
Universitat Politècnica de Catalunya

Barcelona, December 2004

A mis padres,

Abstract

In this Ph.D. dissertation, the design of multi-antenna systems is addressed, where the most general case corresponds to transmitters and receivers with more than one antenna, i.e., to multiple-input-multiple-output (MIMO) channels. The main advantage is that they can provide a much better performance than single-antenna systems, both in terms of quality in the signal transmission and system capacity, i.e., number of users that can be served simultaneously.

The final objective is to carry out a joint design of the transmitter and the receiver, which depends heavily on the quantity and the quality of the channel state information (CSI) available at both sides of the system. In this dissertation, the impact of the CSI on the design has been analyzed.

First, a single-user MIMO communication system has been designed assuming the use of the orthogonal frequency division multiplexing (OFDM) modulation and according to a perfect CSI at both sides. The proposed architecture is based on a joint beamforming approach per carrier. The optimum beamvectors have been calculated and several power allocation strategies among the subcarriers have been derived. These power allocation solutions have been shown to be asymptotically related to other classical designs but with a much lower computational load.

The extension of the previous joint beamforming design to multi-user communications has also been studied assuming a perfect CSI in a scenario with several multi-antenna terminals. The problem formulation has been based on the minimization of the total transmit power subject to several quality of service constraints in terms of a maximum bit error rate (BER) for each link. The mathematical optimization problem has shown to be non-convex and, therefore, classical solutions based on gradient search or alternate & maximize schemes may find a local suboptimum design. As a possible solution, the application of the simulated annealing technique has been proposed, a powerful stochastic optimization tool able to find the global optimum design even when the problem is non-convex.

The errors in the available CSI may decrease importantly the system performance if these errors are not taken into account explicitly during the design. This degradation has been studied for the single-user MIMO-OFDM system. An analytical expression of an upper-bound on the maximum relative signal to noise plus interference ratio degradation has been obtained.

The system performance can be improved when exploiting an imperfect CSI by using adequate robustness strategies. Two approaches have been proposed to obtain robust designs: the Bayesian and the maximin solutions. The Bayesian approach is a full statistical solution that optimizes the mean value of the performance function averaged over the statistics of the actual channel conditioned to the channel estimate, i.e., including the error statistics. On the other hand, the maximin approach provides a design that optimizes the worst system performance for

any error in a predefined uncertainty region that models the imperfect CSI.

Two simple examples of Bayesian robust designs have been provided. First, a power allocation strategy has been derived for an OFDM system with one transmit and one receive antenna minimizing the mean value of an upper-bound on the mean BER. Afterwards, a design strategy for a multi-antenna transmitter with a bank of finite impulse response filters and a single-antenna receiver has been proposed, whose objective is either the maximization of the mean signal to noise ratio (SNR) or the minimization of the mean square error.

Finally, and concerning the maximin approach, a design has been proposed for a single-user MIMO system, in which the transmitter architecture is based on the combination of an orthogonal space time block code (OSTBC), a power allocation stage, and a set of beamformers coupling the transmission through the estimated channel eigenmodes. The power allocation has been found according to a channel estimate and an uncertainty region for the error in this estimate, so that the worst SNR for any error in the uncertainty region is maximized. This design has been then extended and applied to adaptive modulation schemes and multicarrier modulations, showing that the performance is much better than that achieved by a pure OSTBC solution or a non-robust beamforming scheme.

Resumen

Esta tesis aborda el problema del diseño de sistemas multiantena, siendo el caso más general el correspondiente a un canal *multiple-input-multiple-output* (MIMO) donde tanto el transmisor como el receptor cuentan con más de una antena. La ventaja de estos sistemas es que ofrecen un rendimiento mucho mejor que los de una única antena, tanto en términos de calidad en la transmisión como en capacidad entendida como número de usuarios a los que se puede prestar servicio simultáneamente.

El objetivo es diseñar conjuntamente el transmisor y el receptor, lo que depende directamente de la calidad y la cantidad de información del canal de la que se dispone. En esta tesis se analiza el impacto de dicha información en el diseño.

Primero se ha estudiado un sistema MIMO de un único usuario usando la modulación *orthogonal frequency division multiplexing* (OFDM) y asumiendo un conocimiento perfecto del canal en ambos extremos. La arquitectura propuesta se basa en conformación conjunta por portadora, calculándose los conformadores óptimos y proponiéndose diversas estrategias de distribución de potencia entre las portadoras con una baja complejidad. Se han analizado también las relaciones asintóticas de estas distribuciones de potencia con otras soluciones clásicas con mayor coste.

El diseño anterior se ha extendido a sistemas MIMO multiusuario, donde todos los terminales en el escenario tienen más de una antena y la información del canal es perfecta. El objetivo es la minimización de la potencia total transmitida sujeto a restricciones de tasa de error máxima para cada enlace. El problema matemático obtenido es no convexo, por lo que estrategias clásicas basadas en algoritmos de gradiente o de optimización sucesiva pueden llevar a soluciones subóptimas. Como posible alternativa se ha propuesto la aplicación de *simulated annealing*, una potente herramienta heurística y estocástica que permite hallar el diseño óptimo global incluso cuando el problema es no convexo.

Los errores en la información de canal disponible pueden empeorar el rendimiento del sistema si éstos no se tienen en cuenta explícitamente durante el diseño. Se ha estudiado la degradación del sistema MIMO-OFDM de un único usuario en esta situación, obteniendo una expresión analítica de una cota superior de la máxima degradación relativa de la relación señal a ruido más interferencia.

El rendimiento se puede mejorar usando técnicas robustas que tengan en cuenta la presencia de dichos errores. Existen dos aproximaciones clásicas: las Bayesianas y las maximin. En las soluciones Bayesianas el problema se formula estadísticamente, donde el objetivo es optimizar el valor medio de una función de rendimiento promediada sobre la estadística del canal real condicionado a su estimación. Por otro lado, los diseños maximin se caracterizan por optimizar el peor rendimiento para cualquier posible error en la información del canal dentro de una cierta

región de incertidumbre que modela el conocimiento imperfecto del mismo.

Se han mostrado dos ejemplos de diseños Bayesianos. Primero, una distribución de potencia en un sistema OFDM de una única antena que minimiza el valor medio de una cota superior de la tasa de error, y después un diseño de un transmisor multiantena con un banco de filtros que maximiza la relación señal a ruido media (SNR) o minimiza el error cuadrático medio.

Finalmente, se ha obtenido el diseño robusto maximin de un sistema MIMO de un único usuario donde en el transmisor se combinan un código bloque ortogonal espacio-tiempo, una distribución de potencia y un banco de conformadores correspondientes a los modos espaciales del canal estimado. La distribución de potencia se ha diseñado acorde a una región de incertidumbre para el error en la estimación de canal de manera que se maximiza la peor SNR en dicha región. Posteriormente, este diseño se ha extendido al caso de modulaciones adaptativas y multiportadora, mostrando que el rendimiento es mejor que para los códigos bloque otrogonales y la conformación no robusta.

Acknowledgements

Al final todo llega, también la tesis, y eso me hace en este momento muy feliz. Por ello no quiero desperdiciar esta oportunidad única (ya que no se escribe una tesis todos los días) para dar las gracias a todos aquéllos que me han ayudado y han estado a mi lado durante todo este tiempo.

En primer lugar, quiero agradecer a mis directores de tesis Miguel Ángel Lagunas y Ana Pérez su apoyo constante e incondicional y su confianza, además claro, de su ayuda técnica durante toda la elaboración de la tesis. Ellos siempre me han echado una mano y me han comprendido, tanto cuando todo iba bien, como cuando los ánimos no eran los mejores.

En segundo lugar no puedo dejar de nombrar a mi familia, sobretodo a mis padres, que siempre me han ayudado y que seguramente aún están más contentos que yo de que llegue este momento. A ellos les agradezco especialmente que hayan hecho siempre todo lo posible para que haya podido estudiar, sin dudarlo ni un momento. También a mi hermana y mi cuñado, y a mi sobrino Arnau, que sin ni siquiera ser consciente de ello, también me ha hecho muy feliz.

Son también varios los que me han echado una mano en las últimas revisiones de la tesis, así que tengo que agradecerles que hayan perdido un poco el tiempo en leerse unas cuantas páginas de la misma. Gracias pues a Jose L. V., Miquel P. y Toni M.

Y sobretodo, quiero también dar las gracias a todos aquellos compañeros y amigos de la Universidad que me han acompañado durante esta etapa. En primer lugar, a mis compañeros del D5-117 actuales y antiguos: Joel S., Luis G., Camilo C., Carles F., Alejandro R., Ali N., Jose L., Mónica C., Verónica V., Dani P., Javi R., Toni C., Roger G., Joan B. y René J. También a aquéllos con los que, aún sin estar en mi despacho, he tenido más contacto, como Maribel M., Helenca D., Jaume P., Xavi V., Francesc R., Xavi M., Hugo D. y Ramon M. En especial tengo que nombrar a Diego B. y Dani P., que han sido compañeros míos de penas y alegrías desde mis inicios en el doctorado. Con muchos de ellos he pasado muy buenos ratos en algún que otro viaje a congresos, como el de Orlando, el de Alaska o el de Canadá. A ver si hay suerte y en algún otro congreso podemos repetir la experiencia, y sino pues da igual, nos montamos nosotros un viaje y andando...

En Castelldefels también ha habido gente que me ayudado en mis inicios como profesor, así que un agradecimiento también a ellos: Olga M., Pere G., Eduard B. y Gabriel M. También he de agradecer a Xell L. y Alba P. su ayuda para preparar las asignaturas que he impartido.

Y ya finalmente quiero también dar gracias a todos los compañeros del CTTC que he tenido la oportunidad de conocer y con los que he tenido más contacto, en especial a mis compañeros de despacho Jose L. V., Miquel P., Ricardo M., Toni M. y Diego B., y también a Patricia G., Marta V., Jordi C., Xavi M., Jordi M. y Francisco.

Seguramente, debido a mi limitada memoria y a mis ansias por poner punto y final a este trabajo, he dejado a gente sin nombrar en estos agradecimientos, aún cuando deberían estar. Espero me disculpéis por ello y aceptéis también un agradecimiento sincero.

Ahora, a otra cosa...

Toni Pascual

Barcelona, diciembre 2004.

Contents

Notation

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

$\mathbb{R}, \mathbb{C}$	The set of real and complex numbers, respectively.
$\mathbb{R}^{n\times m}, \mathbb{C}^{n\times m}$	The set of $n \times m$ matrices with real- and complex-valued entries, respectively.
$\mathbf{X}^*$	Complex conjugate of the matrix $\mathbf{X}$.
$\mathbf{X}^T$	Transpose of the matrix $\mathbf{X}$.
$\mathbf{X}^H$	Complex conjugate and transpose (Hermitian) of the matrix $\mathbf{X}$.
$[\mathbf{X}]_{i,j}$	(i,j)th component of the matrix $\mathbf{X}$.
$\mathrm{Tr}(\mathbf{X})$	Trace of the matrix $\mathbf{X}$.
$\lvert\mathbf{X}\rvert$ or $\det(\mathbf{X})$	Determinant of the matrix $\mathbf{X}$.
$\mathrm{vec}(\mathbf{X})$	Vec-operator: if $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_n]$, then $\mathrm{vec}(\mathbf{X})$ is the column vector $\left[\mathbf{x}_1^T \cdots \mathbf{x}_n^T\right]^T$.
$\mathrm{diag}(\{x_i\})$	Diagonal matrix, whose elements are $\{x_i\}$.
$\mathbf{I}_n$	Identity matrix of dimensions $n \times n$ (the dimension is not explicitly indicated if it is clear from the context).
$\mathbf{X} \succeq \mathbf{0}$	Positive semidefinite matrix.
$\mathbf{X} \succ \mathbf{0}$	Positive definite matrix.
$\mathbf{X} \preceq \mathbf{0}$	Negative semidefinite matrix.
$\mathbf{X} \prec \mathbf{0}$	Negative definite matrix.
$\mathbf{A} \succeq \mathbf{B}$	$\mathbf{A} - \mathbf{B}$ is a positive semidefinite matrix.
$\mathbf{A} \succ \mathbf{B}$	$\mathbf{A} - \mathbf{B}$ is a positive definite matrix.
arg	Argument.
max, min	Maximum and minimum.
sup, inf	Supremum and infimum.

$(x)^+$	Positive part of the real scalar x, i.e., $(x)^+ = \max\{0, x\}$.
$\|x\|$	Modulus of the complex scalar x.
$\|\mathbf{x}\|$	Euclidean norm of the vector $\mathbf{x}$: $\|\mathbf{x}\| = \sqrt{\mathbf{x}^H\mathbf{x}}$.
$\|\mathbf{x}\|_p$	p-norm of the vector $\mathbf{x} \in \mathcal{C}^{n\times 1}$: $\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$.
$\lambda_{\max}(\mathbf{X})$	Maximum eigenvalue of the positive semidefinite matrix $\mathbf{X}$.
$\lambda(\mathbf{X})$	Set of eigenvalues of the matrix $\mathbf{X}$.
$\mathbf{u}_{\max}(\mathbf{X})$	Normalized eigenvector (norm equal to 1) associated to the maximum eigenvalue of the positive semidefinite matrix $\mathbf{X}$.
$\|\mathbf{X}\|_F$	Frobenius norm of the matrix $\mathbf{X}$: $\|\mathbf{X}\|_F = \sqrt{\mathrm{Tr}\left(\mathbf{X}^H\mathbf{X}\right)}$.
$\|\mathbf{X}\|_2$	2-norm of the matrix $\mathbf{X}$: $\|\mathbf{X}\|_2 = \sqrt{\lambda_{\max}\left(\mathbf{X}^H\mathbf{X}\right)}$.
$\|\mathbf{X}\|_p$	p-norm of the matrix $\mathbf{X}$: $\|\mathbf{X}\|_p = \sup_{\mathbf{y}\neq\mathbf{0}} \frac{\|\mathbf{X}\mathbf{y}\|_p}{\|\mathbf{y}\|_p}$.
$k_p(\mathbf{X})$	Condition number of the nonsingular matrix $\mathbf{X}$ associated to the p-norm: $k_p(\mathbf{X}) = \|\mathbf{X}\|_p\|\mathbf{X}^{-1}\|_p$.
$\mathbf{X}^{-1}$	Inverse of the matrix $\mathbf{X}$.
$\mathbf{X}^{1/2}$	Hermitian square root of the positive semidefinite matrix $\mathbf{X}$, i.e., $\mathbf{X}^{1/2}\mathbf{X}^{1/2} = \mathbf{X}$.
$\sim$	Distributed according to.
$\Pr(\cdot)$	Probability.
$\mathbb{E}[\cdot]$	Mathematical expectation.
$\mathcal{N}(\mathbf{m}, \mathbf{C})$	Real Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$.
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	Complex circularly symmetric Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$.
$\|\mathcal{A}\|$	Cardinality of the set $\mathcal{A}$, i.e., number of elements in $\mathcal{A}$.
$*$	Linear convolution.
$\otimes$	Kronecker product.
$\mathrm{Re}\{\cdot\}$	Real part.
$\mathrm{Im}\{\cdot\}$	Imaginary part.
$\mathrm{dec}\{\cdot\}$	Decision operator corresponding to QAM demodulation.
$\propto$	Equal up to a scaling factor (proportional).

$\triangleq$	Defined as.
$\simeq$	Approximately equal.
$\mathrm{dom}\, f$	Domain of the function f.
$\sup_{\mathbf{x}} f$, $\inf_{\mathbf{x}} f$	Supremum and infimum of the function f with respect to $\mathbf{x}$.
$\nabla_{\mathbf{x}} f$	Gradient of the function f with respect to $\mathbf{x}$ (the variable $\mathbf{x}$ is not explicitly indicated if it is clear from the context).
$\nabla^2_{\mathbf{x}} f$	Hessian matrix of the function f with respect to $\mathbf{x}$ (the variable $\mathbf{x}$ is not explicitly indicated if it is clear from the context).
$\mathcal{G}\{\cdot\}$	Geometric mean.
$\mathcal{H}\{\cdot\}$	Harmonic mean.
lim	Limit.
$\exp(\cdot)$	Exponential.
$\log(\cdot)$	Natural logarithm.
$\log_a(\cdot)$	Base-a logarithm.
$J_0(\cdot)$	Zeroth-order Bessel function of the first kind.

Acronyms

AM	Alternate & Maximize.
AM	Adaptive Modulation.
AP	Access Point.
As	Angular spread.
AWGN	Additive White Gaussian Noise.
BER	Bit Error Rate.
BLAST	Bell-Labs Layered Space-Time.
BPSK	Binary Phase Shift Keying.
BS	Base Station.
CCI	Co-Channel Interference.
cdf	cumulative density function.
CDMA	Code Division Multiple Access.
CP	Cyclic Prefix.
CSI	Channel State Information.
cte	constant.
DFE	Decision Feedback Equalizer.
DMT	Discrete Multi-Tone.
DPSK	Differential Phase Shift Keying.
Ds	Delay spread.
DSL	Digital Subscriber Line.
DSP	Digital Signal Processor (or Digital Signal Processing).
DVB-T	Terrestrial Digital Video Broadcasting.
etc.	etcetera.
ETSI	European Telecommunications Standards Institute.
FDD	Frequency Division Duplexing.
FDMA	Frequency Division Multiple Access.
FFT	Fast Fourier Transform.
FIR	Finite Impulse Response.

GS	Gradient Search.
GSM	Global System for Mobile communications.
ICA	Independent Component Analysis.
IEEE	Institute of Electrical and Electronics Engineers.
IFFT	Inverse Fast Fourier Transform.
i.i.d.	independent and identically distributed.
IIR	Infinite Impulse Response.
ISI	Inter-Symbol-Interference.
KKT	Karush-Kuhn-Tucker.
KL	Karhunen Loeve.
LS	Least Squares.
MAC	Medium Access Control.
MAI	Multiple Access Interference.
MC-CDMA	Multi-Carrier Code Division Multiple Access.
MIMO	Multiple-Input-Multiple-Output.
MISO	Multiple-Input-Single-Output.
ML	Maximum Likelihood.
MLSE	Maximum Likelihood Sequence Estimator.
MMSE	Minimum Mean Square Error.
MRC	Maximal Ratio Combining.
MSE	Mean Square Error.
MT	Mobile Terminal.
OFDM	Orthogonal Frequency Division Multiplexing.
OSI	Open System Interconnection.
OSTBC	Orthogonal Space-Time Block Code.
pdf	probability density function.
PHY	Physical layer.
P/S	Parallel to Serial.
PSK	Phase Shift Keying.
QAM	Quadrature Amplitude Modulation.
QoS	Quality of Service.
QPSK	Quadrature Phase Shift Keying.
RF	Radio Frequency.
RLC	Radio Link Control.
rms	root mean square.
RX	Receiver.

SA	Simulated Annealing.
SDMA	Space Division Multiple Access.
SIMO	Single-Input-Multiple-Output.
SIR	Signal to Interference Ratio.
SISO	Single-Input-Single-Output.
SM	Spatial Multiplexing.
SNIR	Signal to Noise plus Interference Ratio.
SNR	Signal to Noise Ratio.
S/P	Serial to Parallel.
STBC	Space-Time Block Code.
STTC	Space-Time Trellis Code.
SVD	Singular Value Decomposition.
T-DAB	Terrestrial Digital Audio Broadcasting.
TDD	Time Division Duplexing.
TDMA	Time Division Multiple Access.
TX	Transmitter.
UMTS	Universal Mobile Telecommunications System.
vs.	versus.
W	Watt.
WLAN	Wireless Local Area Network.
w.l.o.g.	without loss of generality.
ZF	Zero Forcing.

Chapter 1

Introduction

1.1 Motivation

Wireless communication systems have advanced significantly in the past years and played an extremely important role in our society. The demand for communication among people is increasing exponentially, requiring more connectivity, more services, and higher quality.

Many new and stringent technical requirements have to be faced when considering the design of future wireless systems. Among the different aspects that have to be taken into account, the following ones can be mentioned:

- Increase of the number of people demanding wireless communication services.
- Increase of the demand for broader coverage areas.
- Increase of the demand for the number of mobile services.
- Increase of the demand for higher quality services and, therefore, higher bit rates.
- Increase of the demand for mobile terminals with lighter batteries and longer lifetimes.

Trying to improve all the previous aspects simultaneously is extremely difficult. One possible solution is *diversity*, which consists in sending several copies of the same signal, thus, introducing reliability in the transmission. *Space diversity* is one of the possibilities of including diversity based on the availability of multiple antennas at the transmitter and/or the receiver. The use of multiple antennas allows to exploit the "spatial dimension" of the wireless channel and to provide reliability by simultaneously transmitting the same signal through the new degrees of freedom provided by this spatial dimension.

The study of multiple antennas is a classical problem, although its application to real communications systems and its inclusion in the standardization processes have been considered

Figure 1.1: Example of a multi-antenna GSM BS.

only in the last years. In fact, some of the existing systems use very simple techniques exploiting space diversity. See for example Figure 1.1, in which a sectorial GSM BS is shown, each sector being assigned two receive and one transmit antennas.

The most general case concerning the number of antennas corresponds to a MIMO channel, in which multiple antennas are available simultaneously at the transmitter and the receiver. See Figure 1.2 for a generic representation of a wireless MIMO channel with n_T transmit and n_R receive antennas.

The design of a system with multiple antennas is rather more complicated than the design of a system with a single transmit and receive antenna, i.e., a SISO channel, due to the new degrees of freedom introduced by the spatial dimension that have to be managed. Besides, the design strongly depends on the quantity and the quality of the CSI available during the design.

Most of the existing works related to the design of multi-antenna systems assume that the CSI is either perfect at both sides of the system, or that the channel is completely unknown at the transmitter. These opposite cases have been deeply analyzed and studied, obtaining a broad variety of theoretical results and techniques. Note, however, that in a real implementation of a communication system, some information of the channel is usually expected to be available at the transmitter side, although it will not be perfect or complete. In this situation, when designing a MIMO communication system the best way to exploit this limited information should be studied in order to optimize the system performance.

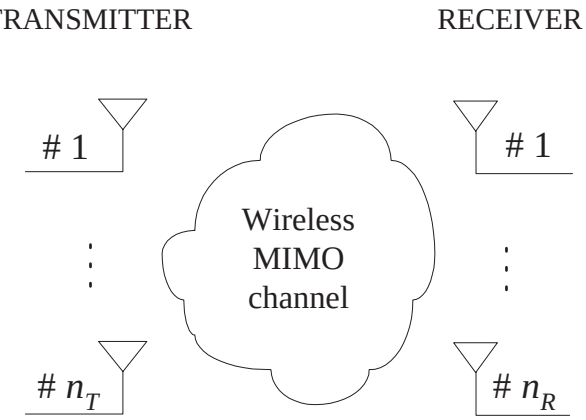


Figure 1.2: Example of a wireless MIMO channel with n_T transmit and n_R receive antennas.

This Ph.D. dissertation provides a contribution to the study of the impact of the imperfect CSI on the system performance and the design, trying to fill in the current lack of literature on these topics. The concrete issues and aspects that have been addressed in this dissertation concerning this general objective are detailed in the following subsection.

1.2 Outline of Dissertation

The focus of this dissertation is on the joint design of the transmitter and the receiver in a communication system with multiple antennas at both sides of the link, i.e., a MIMO channel. More specifically, the impact of the quality and the quantity of the CSI on the system performance and the design is analyzed, proposing also robust solutions with a lower sensitivity to the imperfections in the CSI than other classical non-robust approaches.

Chapter 1 describes the motivation of the work, the outline of the Ph.D. dissertation, and the research contributions in terms of the author's publications.

Chapter 2 provides an overview of the state of the art regarding the joint transmitter-receiver design in a multi-antenna system, with special emphasis on the impact of the CSI. Additionally, a mathematical section devoted to convex optimization is included, since this theory is used throughout the dissertation in order to solve several optimization problems.

Chapter 3 presents a single-user communication system assuming a MIMO channel. The modulation is OFDM and both the transmitter and the receiver architectures are based on a beamforming strategy per carrier. The beamvectors are first designed assuming that the CSI is perfect and, afterwards, several power allocation strategies among the subcarriers are deduced and compared with other classical solutions.

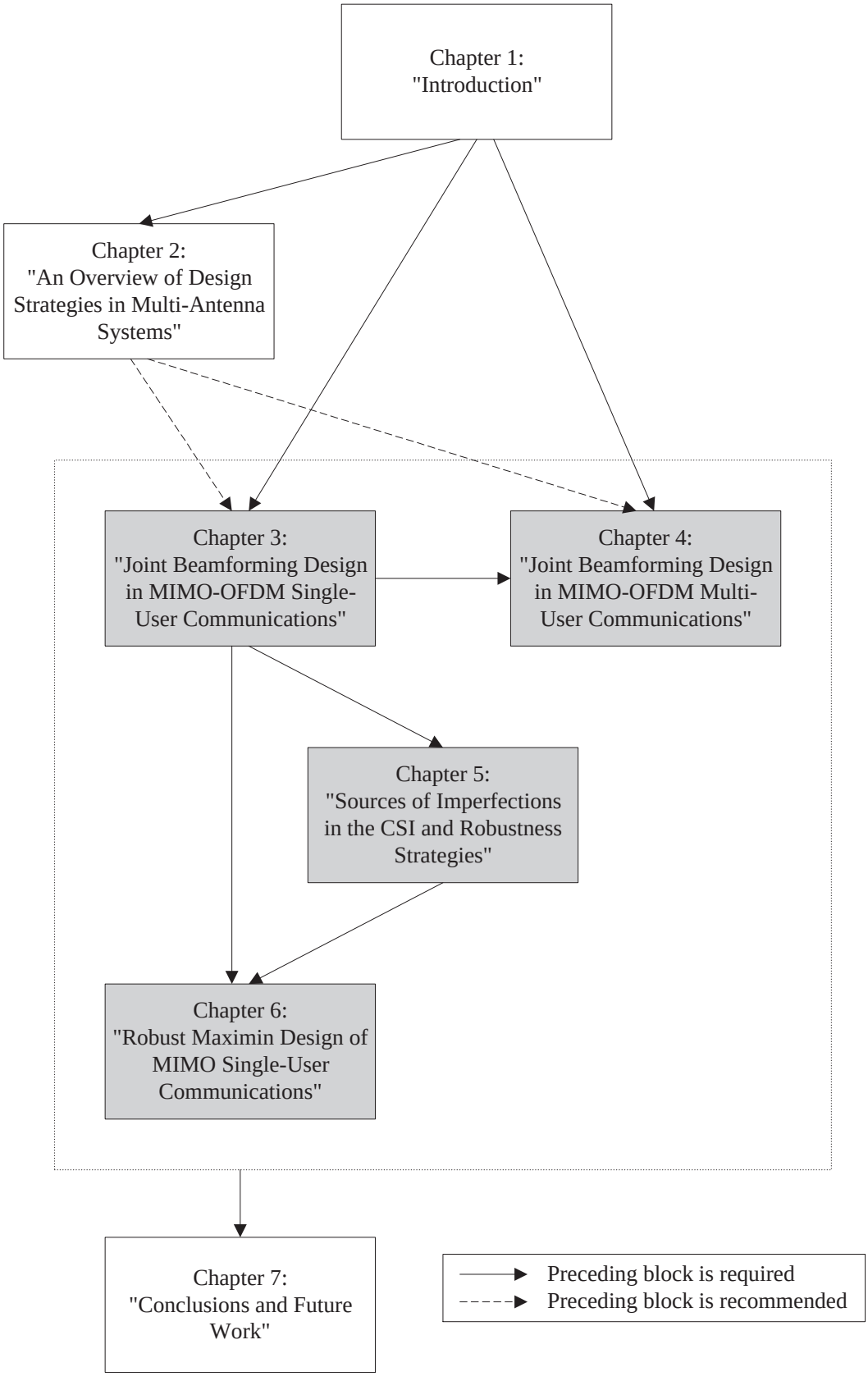


Figure 1.3: Dependence among chapters.

Chapter 4 extends the results shown in the previous chapter to a scenario with several users with multiple antennas and several links among them, leading to multiple parallel MIMO channels. As in the previous chapter, the modulation is OFDM and the transmitter and the receivers are designed according to a beamforming per carrier strategy and assuming a perfect CSI. The mathematical problem, which is formulated as the minimization of the total transmit power subject to several QoS constraints, is non-convex and is solved using the SA algorithm, an iterative and heuristic algorithm which is able to find the global optimum design.

In Chapter 5, the problem of having an imperfect CSI is addressed. First, several sources of imperfections or errors in the channel estimate available during the design stage are identified and briefly analyzed. Afterwards, the impact of the errors in the CSI on the performance of the system described in Chapter 3 is studied, i.e., the impact on a system designed according to a channel estimate which is assumed to be perfect, despite not being true. Two different robustness strategies, the Bayesian and the maximin approaches, are then presented as tools to obtain robust designs less sensitive to the errors in the CSI. Finally, two illustrative examples of Bayesian designs are given.

Chapter 6 considers a single-user MIMO flat fading channel, where the objective is to obtain a robust design of the system according to an imperfect channel estimate and under the maximin approach. The solution to the mathematical problem is simplified using convex optimization theory. The design is then extended to the case of frequency selective channels and the OFDM modulation, similarly to Chapter 3, and, finally, is applied to maximize the throughput combining the robust design with AM techniques.

Finally, Chapter 7 concludes the dissertation and proposes possible future lines of research.

Figure 1.3 illustrates the dependence among the chapters in this dissertation, differentiating recommended sequences of reading the chapters from required sequences.

1.3 Research Contributions

The main contribution of this Ph.D. dissertation is the analysis of the impact of the quality of the CSI on the design of multi-antenna communication systems. The details of the research contributions in each chapter are as follow.

Chapter 3

The main results in this chapter concerning the design of MIMO-OFDM single-user systems with perfect CSI at the transmitter and the receiver have been published in one journal paper and four conference papers:

- [PI04f] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "On Power

Allocation Strategies for Maximum Signal to Noise and Interference Ratio in an OFDM-MIMO System," *IEEE Trans. on Wireless Communications*, vol. 3, no. 3, pp. 808-820, May 2004.

- [PI01a] A. Pascual Iserte, M. A. Lagunas Hernández, and A. I. Pérez Neira, "Space-Time Diversity Applied to Single-User Environments and MIMO Transmission Channels," *Proc. IEEE International Conference on Electronics, Circuits, and Systems (ICECS'01)*, vol. 3, pp. 1179-1182, September 2001.

- [PI01c] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Pre- and Post-Beamforming in MIMO Channels Applied to HIPERLAN/2 and OFDM," *Proc. IST Mobile & Wireless Communications Summit (IST'01)*, pp. 3-8, September 2001.

- [PI02b] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Joint Beamforming Strategies in OFDM-MIMO Systems," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*, vol. 3, pp. 2845-2848, May 2002.

- [PI02e] A. Pascual Iserte, A. I. Pérez Neira, D. P. Palomar, and M. A. Lagunas Hernández, "Power Allocation Techniques for Joint Beamforming in OFDM-MIMO Channels," *Proc. European Signal Processing Conference (EUSIPCO'02)*, vol. 1, pp. 383-386, September 2002.

Chapter 4

The main results in this chapter regarding the design of MIMO-OFDM multi-user systems with perfect CSI at the transmitter and the receiver have been published in one journal paper and two conference papers:

- [PI04e] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "An Approach to Optimum Joint Beamforming Design in a MIMO-OFDM Multiuser System," *accepted for publication in EURASIP Journal on Wireless Communications and Networking*, 2004.

- [PI02d] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Simulated Annealing Techniques for Joint Transmitter-Receiver Design in a Multiple User Access MIMO-OFDM Channel," *Proc. IST Mobile & Wireless Communications Summit (IST'02)*, pp. 325-329, June 2002.

- [PI02c] *Invited paper:* A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Joint Transceiver Optimization in Wireless Multiuser MIMO-OFDM Channels Based on Simulated Annealing," *Proc. European Signal Processing Conference (EUSIPCO'02)*, vol. 2, pp. 421-424, September 2002.

Chapter 5

One journal paper and four conference papers contain the main results in this chapter regarding the impact of the errors in the CSI on the system performance and two examples of Bayesian robust designs:

- [PI04f] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "On Power Allocation Strategies for Maximum Signal to Noise and Interference Ratio in an OFDM-MIMO System," *IEEE Trans. on Wireless Communications*, vol. 3, no. 3, pp. 808-820, May 2004.

- [PI03c] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Exploiting Transmission Spatial Diversity in Frequency Selective Systems with Feedback Channel," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, vol. 4, pp. 85-88, April 2003.

- [PI03b] A. Pascual Iserte, M. A. Lagunas Hernández, and A. I. Pérez Neira, "Robust Power Allocation for Minimum BER in a SISO-OFDM System," *Proc. IEEE International Conference on Communications (ICC'03)*, vol. 2, pp. 1263-1267, May 2003.

- [PI03d] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Performance Degradation of an OFDM-MIMO System with Imperfect Channel State Information at the Transmitter," *Proc. IST Mobile & Wireless Communications Summit (IST'03)*, vol. 2, pp. 396-400, June 2003.

- [PI03a] *Invited paper:* A. Pascual Iserte, M. A. Lagunas Hernández, and A. I. Pérez Neira, "Diseño Cooperativo de Sistemas de Comunicaciones con Información Parcial del Canal," *Proc. Simposium Nacional de la Unión Científica Internacional de Radio (URSI'03)*, September 2003.

Chapter 6

The main results obtained in this chapter concerning the robust design of a MIMO communication system under the maxmin philosophy have been published in one book chapter, one paper submitted to a journal, and two conference papers:

- [Pal05] D. P. Palomar, A. Pascual Iserte, J. M. Cioffi, and M. A. Lagunas Hernández, "Convex Optimization Theory Applied to Joint Transmitter-Receiver Design in MIMO Channels," A. B. Gershman and N. D. Sidiropoulos (eds.), *Space-Time Processing for MIMO Communications (to appear)*, John Wiley & Sons, 2005.

- [PI04b] A. Pascual Iserte, D. P. Palomar, A. I. Pérez Neira, and M. A. Lagunas Hernández, "A Robust Maximin Approach for MIMO Communications with Partial Channel State Information Based on Convex Optimization," *submitted to IEEE Trans. on Signal Processing*, September 2004.

- [PI04d] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "A Maximin Approach for Robust MIMO Design: Combining OSTBC and Beamforming with Minimum Transmission Power Requirements," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04)*, vol. 2, pp. 1-4, May 2004.

- [PI04c] A. Pascual Iserte, M. Payaró, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Robust Adaptive Modulation for Throughput Maximization in MIMO Systems Combining OSTBC and Beamforming," *Proc. IST Mobile & Wireless Communications Summit (IST'04)*, June 2004.

Other contributions not presented in this dissertation

During the Ph.D. period, the author has published other papers not directly related to the main topic of this dissertation.

Some results on the estimation of parameters related to spatially distributed sources have been published in the following conference paper:

- [PI01b] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Iterative Algorithm for the Estimation of Distributed Sources Localization Parameters," *Proc. IEEE Workshop on Statistical Signal Processing (SSP'01)*, pp. 528-531, August 2001.

A conference paper on the design of FIR filters in MIMO channels has also been published:

- [Pal01] D. P. Palomar, M. A. Lagunas Hernández, A. Pascual Iserte, and A. I. Pérez Neira, "Practical Implementation of Jointly Designed Transmit-Receive Space-Time IIR Filters," *Proc. IEEE International Symposium on Signal Processing and its Applications (ISSPA'01)*, pp. 521-524, August 2001.

A semiblind technique for designing FIR filters in a multi-antenna OFDM receiver has been deduced in the following conference paper:

- [Bar02] D. Bartolomé, A. I. Pérez Neira, and A. Pascual Iserte, "Blind and Semiblind Spatio-Temporal Diversity for OFDM Systems," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*, vol. 3, pp. 2769-2772, May 2002.

The application of adaptive array techniques to Bluetooth devices has been presented in

- [PI02a] A. Pascual Iserte, N. Ahmed Awad, and A. I. Pérez Neira, "Array Antennas for Packet Transmission Networks," *Proc. ETSI Workshop on Broadband Wireless Ad-Hoc Networks and Services*, September 2002.

A conference paper concerning the application of antenna arrays to multi-user systems, including spatial scheduling strategies, has also been published:

- [Bar03a] D. Bartolomé, A. Pascual Iserte, and A. I. Pérez Neira, "Spatial Scheduling Algorithms for Wireless Systems," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, vol. 4, pp. 185-188, April 2003.

Some results related to the realistic hardware implementation of adaptive antenna algorithms have been presented in two conference papers:

- [Bar03b] D. Bartolomé, A. Pascual Iserte, A. I. Pérez Neira, and P. Rosson, "From a Theoretical Framework to a Feasible Hardware Implementation of Antenna Array Algorithms for WLAN," *Proc. IST Mobile & Wireless Communications Summit (IST'03)*, vol. 1, pp. 1-5, June 2003.
- [Hen04] C. Hennebert, P. Rosson, D. Bartolomé, A. Pascual Iserte, and A. I. Pérez Neira, "Practical Implementation of Space-Diversity Receivers in OFDM Systems: Structure, Performance, and Complexity," *Proc. IST Mobile & Wireless Communications Summit (IST'04)*, June 2004.

A paper concerning the application of ICA based algorithms for the detection of OSTBC signals has been presented in a conference:

- [Liu03] J. Liu, A. Pascual Iserte, and M. A. Lagunas Hernández, "Blind Separation of OSTBC Signals Using ICA Neural Networks," *Proc. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'04)*, December 2003.

Two additional contributions regarding robustness criteria and design strategies have been presented in a paper submitted to a journal and a conference paper:

- [Mor04] A. Morell, A. Pascual Iserte, and A. I. Pérez Neira, "Fuzzy Inference Robust Beamforming," *submitted to EURASIP Signal Processing*, 2004.

- [PI04a] A. Pascual Iserte, M. A. Lagunas Hernández, and A. I. Pérez Neira, "Robustness Criteria for Transmit Spatial Diversity Systems in Frequency Selective Channels," *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM'04)*, July 2004.

Chapter 2

An Overview of Design Strategies in Multi-Antenna Systems

2.1 Space Diversity in Communication Systems

The demand for communication among people in the current society has become clearly important in the last years, and is expected to be a key and essential issue in the very near future. From a technical point of view, many challenges arise from this fact. Current wireless communication systems such as GSM [Mou92], UMTS [Hol02], or some WLAN's, including Bluetooth [Blu01], HiperLAN/2 [ETS00], and IEEE 802.11a [IEE99], have a limited capacity, not only in terms of the number of users that can be served simultaneously, but also in terms of BER, bit rate, delay, and, in general, in terms of QoS. Keeping in mind which will be the new and increased demands for service of the information society in the future, both in quantity and quality, the main technical challenges will consist in the design of new and advanced wireless communication systems able to cope with all these new and increasingly stringent requirements.

Many are the direct consequences and technical demands derived from the needs of the future wireless communication systems. The first one is, obviously, an increase of the bit rate so that the new multimedia applications can also be served "over the air". The first possible solution consists in the increase of the transmission bandwidth, although this is not always possible due to the physical constraints in the system. Note however that, even in case that the increase of the bandwidth is affordable, the negative effects of the channel over the transmitted signal may become very important, specially due to the fact that the time dispersion effect or frequency selectivity of the channel increases as the transmission bandwidth is enlarged [Pro95]. The number of users is also a key parameter to keep in mind when designing future communication systems. It is expected that this number of users will increase importantly in such a way that the current systems will not be able to accommodate all the demands for service. As an evidence of this, it is only necessary to have a look at the evolution of this kind of capacity demands

for cellular communication systems and wireless telephony in the last years. Because of that, the new systems should take into account this fact and, therefore, the development of new and efficient multiple access technologies will also be a key technological issue both from the research point of view and their applicability in a real system.

As a summary, the new technical challenges can be divided into two main groups:

- *Signal related issues:* within this group of challenges, those aspects related to the quality of the signal itself transmitted through the wireless channel are considered: signal distortion, interference rejection, etc.; i.e., problems related to the QoS of the bit stream transmitted "over the air" to a concrete user. In this point, the effects of the scattered wireless channel should be taken into account, including both the dispersive behavior in the time domain (frequency selective channel) and in the frequency domain (time selective channel).
- *System related issues:* within this group, the aspects concerning protocol layers higher than the physical one are taken into account, such as MAC and RLC. The main objective is to increase the capacity of the system in terms of number of users while maintaining the same physical resources and, therefore, one of the challenges is to provide effective multiple access strategies.

When looking for possible solutions to the aforementioned problems, several constraints should be taken into account in order to design systems attractive to the telecommunications manufacturers, in the sense that they can be implemented and commercialized with a minimum guarantee of success. Within these constraints, many groups can be identified: physical, cost, and complexity constraints, etc.; although very frequently, many of them are interconnected. As an example, the following constraints can be mentioned: maximum signal bandwidth, maximum transmit power, maximum delay, maximum BER, minimum battery lifetime, etc.

There are many possible solutions to cope with these problems while still fulfilling the a-priori defined constraints. One of the main concepts that has been used in order to design wireless communication systems is *diversity*. Mainly, diversity consists in improving the reliability of the signal transmission by exploiting several copies of the signal which are sent through "parallel" channels, where the concept of "channel" should be understood in a more general way than its physical meaning. Among the different kinds of diversity, it is possible to cite the *channel coding*, the *frequency diversity*, the *time diversity*, or the *space diversity*. While channel coding and frequency and time diversities were some of the first ones considered in communication systems, the main disadvantage is that their use implies an increase of the signal bandwidth and the transmission delay. The other possibility consists in *space diversity* schemes, i.e., schemes that make use of more than one antenna, also called *antenna arrays*, at the transmitter and/or the receiver. By means of this, the quality of the signal and the

system performance can be importantly improved without increasing the signal bandwidth and the transmit power. When considering space diversity schemes, several possibilities may arise depending on the configuration adopted by the transmitter and the receiver:

- *Single-Input-Single-Output (SISO) channels:* one antenna at the transmitter and one antenna at the receiver.
- *Single-Input-Multiple-Output (SIMO) channels:* one antenna at the transmitter and multiple antennas at the receiver.
- *Multiple-Input-Single-Output (MISO) channels:* multiple antennas at the transmitter and one antenna at the receiver.
- *Multiple-Input-Multiple-Output (MIMO) channels:* multiple antennas at the transmitter and multiple antennas at the receiver.

As has been already stated, space diversity architectures are able to improve the quality and reliability of the radio link, so that the received signal can be decoded and detected with a higher quality without increasing the bandwidth and power, which are very valuable physical resources. However, this is not the only application of antenna arrays. Thanks to such systems, it is possible to implement the SDMA. This is a multiple access technique similar to TDMA, FDMA, and CDMA, but in this case, the spatial dimension is used to separate the signals transmitted from/to different users. The key advantage is that this can be performed without increasing the bandwidth or decreasing the bit rate assigned to each user; in other words, space diversity systems offer an attractive solution to increase the capacity of the wireless communication systems with a high spectral efficiency. Indeed, this last point concerning the multiple access capability of antenna arrays shows that in the next communication systems, there will be a link between the PHY layer, directly related to the front-ends design based on space diversity, and other upper layers such as MAC and RLC, which are more related, not only to the multiple access capabilities, but also to other system and network parameters that can be efficiently improved by means of a joint design of the different layers in the OSI protocol stack in the so called *cross-layer* perspective [AH02].

The main objective of this Ph.D. dissertation is to present several design strategies of multi-antenna systems keeping in mind that the ultimate objective is to improve the quality of the signal and the capacity of the system. The optimum solution when designing such a system is to carry out the design of the transmitter and the receiver jointly. Taking this ultimate goal into account, many possible schemes and architectures arise. The optimum joint design is obtained when all the information about the channel response, the interferences, or other users' channel responses (embraced by the concept of CSI) are available at both the transmitter and

the receiver during the design stage. However, in a realistic scenario, this is not always possible and, sometimes, only some imperfect or partial information is available. Besides, the quantity and the quality (i.e., the accuracy and reliability) of this information may be different at the transmitter and the receiver. Concerning the channel response, usually only a channel estimate can be exploited to carry out the design. In TDD systems with a coherence time longer than the duplexing period, the channel estimate obtained during the transmission in the reverse link can be used for the design of the system when transmitting in the forward link due to the channel reciprocity principle. In the case of FDD, this is not a possible solution and, therefore, some kind of *feedback channel* has to be implemented from the receiver to the transmitter so that the transmitter can obtain an estimate of the channel, or any other related information, in order to design the front-end. In the opposite case, there exists the possibility of having no information at all at the transmitter side and, therefore, the only solution is to use a channel coding approach.

From a theoretical point of view, there are several possibilities of exploiting the available information during the design. For example, this information can be assumed to be complete and exact, i.e., with no error. This corresponds to the assumption of an ideal case. This kind of designs may be very sensitive to errors in this information and, therefore, are not attractive from a practical and realistic point of view. There are other design strategies that take into account explicitly the fact that this information may have some error or may be incomplete and partial. This kind of strategies are *robust* to errors and have a more practical sense when implementing real systems, since they are less sensitive to these errors than the non-robust counterparts.

As a summary, in this Ph.D. dissertation, several designs of multi-antenna systems are presented according to different degrees of quality of the CSI. First, some designs are shown assuming that the CSI is perfect, both for the case of single-user and multi-user scenarios. Then, errors in the CSI are assumed to, as a first stage, analyze the impact of these errors on the system performance when assuming that the CSI is perfect, despite not being true and, afterwards, to obtain robust solutions that take into account explicitly the presence of these errors, improving the performance.

2.2 Impact of the CSI on the Design: An Overview of the State of the Art

As explained previously, the optimum solution in a design problem consists in a joint optimization of the transmitter and the receiver. The theory of optimum terminal filters design has been a classical one since 1970 and has been addressed, for example, in [Car86] and [Pro95]. There, the frequency response of the channel is assumed to be known at both the transmitter and the receiver and the objective is the maximization of the SNR at the detection stage subject to two different design constraints. The first one is a ZF condition, which is equivalent to forcing no ISI

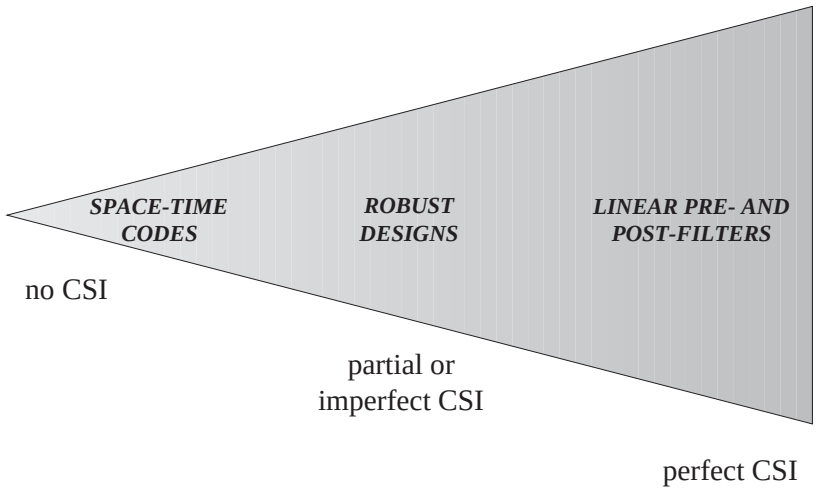


Figure 2.1: Diagram showing the different possibilities of CSI at the transmitter and the main solutions.

at the detection stage, whereas the second constraint is related to the maximum transmit power. Since that analysis, many researchers have been working actively on the topic, generalizing the results for systems using antenna arrays at the transmitter and/or the receiver.

As mentioned in the previous subsection, many possible situations arise depending on the quantity and the quality of the information available during the design stage. In this information it may be included, for example, the channel impulse response for the desired communication or link, the statistics of the interferences, the channel responses for other users in the same area, the mobile terminal speed, the signal strength, etc., among other aspects. In this dissertation, the most general definition and concept of "information" is used under the name *Channel State Information* (CSI).

As far as the practical implementation of real systems is concerned, the most common situation corresponds to the case in which most information is available at the receiver side of the communication system. This is obvious, as the channel estimation process is performed at the receiver during the training period in which the transmitter sends a pilot or training sequence. Besides, during this period the statistics of the interferences can also be estimated. The most problematic issue concerns the case in which it is desired that the transmitter has also some CSI in order to carry out the design of the front-end. Taking this into account, 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. Example: space-time coding.
- **Perfect CSI:** the transmitter has full knowledge of the channel and, possibly, of the

interferences statistics at the receiver. Example: optimum terminal filtering. There are many possible strategies and optimization criteria to carry out the design.

- **Partial or imperfect CSI:** the transmitter has only a partial or imperfect knowledge of the channel and the interference statistics at the receiver, which can be obtained through a feedback channel from the receiver to the transmitter or based on the channel reciprocity principle in TDD systems. Examples: transmission of a quantized version of the channel estimate, transmission of some digital parameters calculated at the receiver, transmission of some statistics of the channel, etc.

In the following, a brief overview of the state of the art is given for the three possibilities described above and represented in Figure 2.1. The objective is not to provide an exhaustive description of the state of the art, but only to show a general view of the current research lines regarding the joint transmitter-receiver design in multi-antenna systems. A more detailed description of the state of the art is given in each chapter when necessary.

2.2.1 Designs with No CSI at the Transmitter

The problem regarding the case in which no CSI is available at the transmitter has received much attention from the research community for many years because of its evident importance from a practical implementation point of view. One of the first works considering this case is [Ses93], in which the *delay diversity* technique is presented in a multiple transmit antenna system. This technique consists in transmitting the same information symbols through all the antennas, but applying a different delay at each antenna. This strategy has been afterwards generalized to the case of *linear precoding* with no CSI in other works such as [Wor97], in which a linear filter is applied at each transmit antenna. These filters can be designed based on Fourier or Hadamard structures, among others.

Besides the linear precoding approach, much work and research has been focused on *space-time coding* techniques. Basically, this approach results from the generalization of the classical channel coding techniques to the case of a transmit multi-antenna system (MISO or MIMO), introducing redundancy in the transmission both in the temporal and the spatial dimensions. In [Tar98], the design criteria of the optimum space-time codes for a Rayleigh flat fading MIMO channel are derived and several *space-time convolutional or trellis codes* (STTC) are proposed according to these criteria, providing both diversity gain and coding gain. In that work, it is considered that the CSI is perfect at the receiver, which is responsible for detecting the symbols under the ML criterion. In [Tar99c], the effect produced by the errors in the CSI on the BER performance is studied. Besides, a new metric for the Viterbi decoder is proposed to improve the performance taking into account the errors in the CSI. One of the problems of STTC is that the computational complexity of both the transmitter and the ML receiver increases importantly as

the number of transmit and/or receive antennas and the code rate are increased. As a possible solution, in [Tar99b], the use of STTC combined with beamforming at the receiver is considered. The idea consists in having parallel space-time codes working at different non-overlapping groups of transmit antennas. The receiver is composed of a bank of beamformers, each one responsible for detecting the symbols transmitted through a concrete antenna group while cancelling the contributions of the signals transmitted through the other groups of antennas. Obviously, this receiver structure is not the optimum one as it does not correspond to the ML detector and, therefore, the diversity gain is decreased; however, the great advantage is that the computational complexity can be reduced dramatically.

The main drawback of STTC, as stated previously, is that the optimum decoder based on the ML criterion may have a not affordable computational complexity. This fact has motivated the introduction of the *space-time block codes* (STBC) for MIMO channels. [Ala98] is the first work proposing a block code for a flat fading channel with two transmit antennas and any number of receive antennas. The optimum ML receiver is based on the assumption of having a perfect CSI and can be implemented by only linear processing with a very low computational complexity. In [Gan01], [Gan02a], and [Tar99a], Alamouti's work is generalized to the case of any number of transmit antennas, proposing possible STBC solutions based on orthogonal designs both for real and complex signal constellations. There, it is shown that full rate codes are available for any number of transmit antennas in the case of real constellations, whereas in the complex case, full rate codes are possible only for two transmit antennas. Using these designs, the optimum ML receiver can be implemented by only linear processing assuming perfect CSI thanks to the "orthogonality" of the codes; besides, it is shown that this kind of codes provides full diversity. In [Gan02a], a simple theoretical analysis of the performance degradation is carried out when the CSI is not perfect at the receiver and a possible architecture for robust transmit antenna selection is derived. In [Jaf01], the quasi-orthogonal block codes are presented. These codes have full rate even for complex constellations and for any number of transmit antennas. The main disadvantage is that the optimum ML receiver has a higher computational complexity than the corresponding one for the orthogonal codes and that there is a loss in the diversity order. [Sto02] provides a new interpretation of orthogonal STBC, also called "amicable designs", as the robust solutions maximizing the worst SNR when the flat fading MIMO channel is totally unknown at the transmitter.

In most of the previous works, the CSI has been considered to be perfect at the receiver during the detection stage. In a real system, this is not always true and, therefore, a performance degradation is expected. In [Fen00], the BER degradation is studied by means of simulations and a simple theoretical analysis is deduced for a single channel realization comparing the case of MRC and a 2-branchs diversity Alamouti's scheme. In [Ste01], the problem of imperfect CSI at the detection stage is also addressed. In that work, both the MRC technique with two receive

antennas and Alamouti's code with one receive antenna are compared in terms of SNR reduction by means of a theoretical analysis and in terms of BER degradation through simulations. The main conclusion is that MRC performs better than Alamouti's code if both techniques are evaluated under the same SNR during the channel estimation process.

From the works mentioned previously, it has been concluded that the imperfect CSI at the receiver may have a very negative impact on the system performance if the error in this CSI is not negligible when designing and applying the corresponding detector and space-time decoder. This fact has motivated many researchers to investigate the case in which there is no CSI at the transmitter and the receiver. Most of these works consider the concrete case of Rayleigh uncorrelated flat fading MIMO channels. In [Mar99], the channel is assumed to be unknown at both ends and to be constant over T symbol periods, changing independently after each of these periods. The objective is to find the pdf of the transmitted signal that maximizes the mutual information, i.e., the capacity achieving solution, as described in [Sha48]. From the information theory point of view, it is shown that it is not worthwhile having more transmit antennas than T. There, it is shown that as T and/or the SNR tend to infinity, the optimum transmitted signal tends to the solution achieving the capacity with CSI at the receiver and no CSI at the transmitter, i.e., in which the symbol sequences transmitted through different antennas are mutually orthogonal, leading to the so called unitary codes. In [Hoc00a], unitary code designs are presented according to the asymptotic optimum solution shown in the previous work. The optimum detector is based on the ML criterion, although the final implementation depends on the availability of CSI at the receiver. If CSI is available, a conventional "coherent" ML detector is derived, whereas in the case of unknown channel, the ML detector is an "uncoherent" receiver with a quadratic structure. Basically, the gain of a coherent detector in front of the uncoherent version is around 3 dB in terms of SNR. Besides, in that paper, different criteria for the design of the unitary codes are derived under the ultimate goal of minimizing the mean error probability. In [Hoc00b], a systematic design of unitary codes is provided based on Fourier and algebraic models.

In [Hoc00c], the design of unitary space-time matrices is applied to the case of *differential modulation*, which is basically the extension of DPSK to the case of multiple transmit antennas. In this case, the optimum quadratic ML detector with no CSI at the receiver working over the entire received signal sequence can be simplified to a suboptimum differential detector which works only over pairs of received signal blocks, lowering the computational complexity. Several families of codes with a good performance and with different characteristics are also provided. Concretely, most of these families are groups from an algebraic point of view, which simplifies importantly the coding process. Also in [Tar00] and [Hug00], the same problem is addressed. The first paper presents the differential extension of Alamouti's code using PSK signal constellations with very simple encoders and decoders. The second paper proposes a system using differential

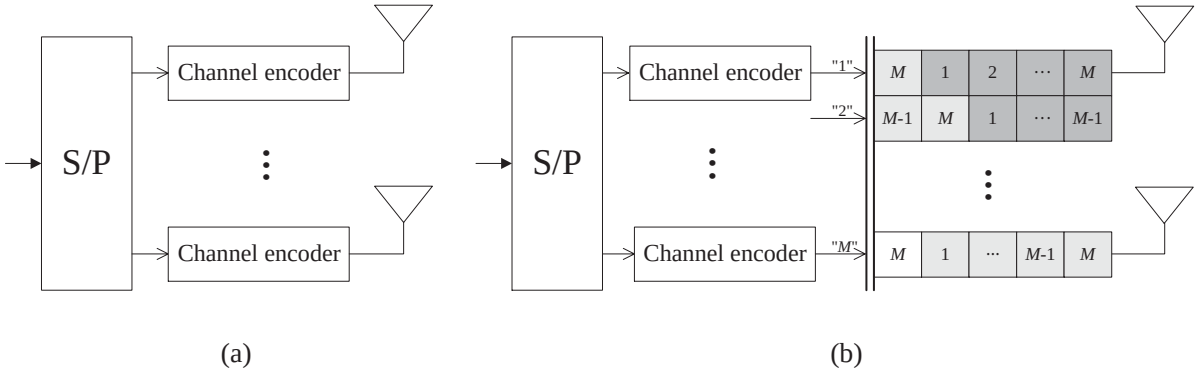


Figure 2.2: Transmitter scheme for BLAST. (a) V-BLAST. (b) D-BLAST.

unitary space-time block codes, where the optimum ML receiver with unknown channel is shown to be quadratic for the case of Rayleigh uncorrelated MIMO channels, but can be simplified to a suboptimum low complexity differential detector. In [Gan02b], a similar but more simplified analysis is carried out, focusing the attention on the differential receiver. In [Has02], a systematic approach for the design of differential codes for any number of transmit antennas, receive antennas, and code rate is provided. Those families are based on the Cayley transform, which enables to find very simple encoders based on a linear mapping and a transformation, and very low complex suboptimum receivers based on the resolution of a set of linear equations. As shown, the great advantage of the differential systems is that they do not need CSI at the receiver and, therefore, training or pilot sequences are not necessary to estimate the channel, which increases the spectral efficiency. Besides, the differential detectors at the receiver are quite simple and do not suffer from error propagation.

Finally, it is important to remark that there exist other techniques that can be used in the case of having no CSI at the transmitter to allow to increase the transmission rate by exploiting the *spatial multiplexing* (SM) capabilities provided by the multiple antennas. BLAST based strategies were first developed by Bell-Labs [Fos96] and are the most known SM techniques. The most popular schemes are the Vertical-BLAST (V-BLAST) and Diagonal-BLAST (D-BLAST). The latter is more complex to implement and that is the reason why the former is the most widely used BLAST strategy. Both techniques obtain full diversity (assuming full multiplexing is performed) and full rate. The use of multiple antennas aims to gain in channel capacity at the expenses of increasing the number of symbols transmitted in each channel access. More precisely, as many symbols as antennas at the transmitter side are sent in one channel attempt. The transmission scheme for V-BLAST [Loz02] is yet very simple. It takes the substreams going to be transmitted (which is equal to the number of antennas at the transmitter side) and performs a one to one mapping from each symbol to each antenna. D-BLAST scheme is a more complex scheme. Let us assumed that M transmit antennas are available. In that case, the

original stream is divided into M parallel substreams, so that the symbols from each of them are transmitted sequentially through all the antennas in successive channel attempts [Fos96]. See Figure 2.2 for a generic representation of a BLAST transmitter.

2.2.2 Designs with Perfect CSI at the Transmitter

The opposite approach from the previous one consists in assuming that the CSI at both the transmitter and the receiver is perfect. Obviously, a system designed using a perfect CSI is expected to perform better than any space-time coding technique, since the design can be perfectly matched to the actual channel realization. In this subsection, a brief overview of the state of the art for the case of perfect CSI is provided. The difference among the works and papers rely on the different system architectures, optimization criteria, and design constraints that are adopted. Due to the great quantity of available works and references, only some examples are given, excluding those that are not directly related to the main objectives of this dissertation.

[Pau97] is a tutorial introducing the main concepts of space-time processing and optimum design with perfect CSI assuming space diversity at either the transmitter or the receiver, i.e., SIMO or MISO channels. In [Sal85], a technique is presented for designing a MIMO system with the same number of antennas at both the transmitter and the receiver (indeed, the proposed design in that paper is intended for a wireline system with multiple inputs and outputs suffering from the cross-coupling effect). In this system, several data streams are transmitted in parallel and at the same time through the MIMO channel. The transmitter and the receiver are based on linear time-invariant filters which are designed under the MMSE criterion and assuming no excess bandwidth. In [Yan94b], the results are extended to the case of a MIMO system with any number of transmit and receive antennas and considering excess bandwidth. In both cases, a single-carrier modulation is assumed and the optimum filters are designed in the frequency domain. The result corresponds to a solution in which the channel is diagonalized at each frequency and each data stream is transmitted through a different eigenmode or "spatial subchannel". Afterwards, the optimum power allocation is calculated so that the MSE is minimized. In that work, it is considered that there may be interferences at the receiver by assuming the general case of non-white spatial and frequency noise. Other related works and/or generalizations are [Yan94a] and [Hon92]. In the first paper, a system with a DFE at the receiver is proposed. The transmit and receive filters are designed jointly to minimize the geometric mean of the MSE, providing a solution in the frequency domain that also diagonalizes the channel. The second work addresses the difficult problem of designing a MIMO system with structural constraints at the transmitter and the joint design of FIR transmit and receive filters. Usually, closed-form solutions cannot be found and iterative suboptimum algorithms have to be used.

In [Ral98], the problem concerning the design of a transmitter in a MIMO channel with

another criterion different from MMSE is proposed. In that work, the ultimate objective is to design the transmitter so as to maximize the mutual information as defined in [Sha48]. The channel is frequency selective and has any number of transmit and receive antennas. Two different techniques to achieve capacity are proposed, both of them assuming perfect CSI at the transmitter and carrying out the optimum power allocation *water-filling* [Cov91]: spatio-temporal vector-coding and discrete matrix multitone.

In [Sca02], a block transmission mode is assumed in a frequency selective MIMO channel, i.e., the transmitted symbols are divided into disjoint groups, and then each group is encoded using a matrix called *linear precoder* that operates simultaneously in the spatial and temporal dimensions. At the receiver, also a *linear decoder* matrix is applied to the received samples to estimate the transmitted data. Both matrices are designed jointly using three different criteria: MMSE, the minimization of the determinant of the MSE matrix, and the maximization of the minimum eigenvalue of the SNR-like matrix. The designs are carried out subject to a mean power constraint and a maximum eigenvalue power constraint. All the solutions diagonalize the space-time MIMO channel matrix but allocates the power according to different policies. Indeed, [Sca02] extends the results of [Sca99a] and [Sca99b], in which a frequency selective SISO system is considered. The transmitter and the receiver are also based on a block transmission and the signal model is equivalent to considering a MIMO channel. In those papers, the optimum transmitter and receiver are found under maximum mutual information, MMSE, and maximum SNR criteria subject to maximum power constraints and, in some cases, to ZF constraints. In [Sam01], also the linear precoding and decoding problem in MIMO channels is addressed, where the objective is to design the system according to a weighted MMSE criterion. In that paper, it is shown that the solution also diagonalizes the channel and that, by correctly choosing the weighting matrix, many other known criteria can be accommodated, such as maximum information rate, QoS constraints, the unweighted MMSE, and equal error design. In [Din02], it is considered the very important case in which the ultimate goal is the minimization of the BER also in a block transmission system using either zero padding or cyclic prefix. This is carried out subject to two constraints: the maximum mean transmit power and ZF. The main result is that the precoder minimizing the BER also minimizes the MSE, although, in general, does not diagonalize the MIMO channel matrix.

An important case in many current systems, such as the WLAN's HiperLAN/2 [ETS00] and IEEE 802.11a [IEE99], is the utilization of the multicarrier OFDM modulation [Bin90, Nee98, Wan00] at the physical layer. Taking that into account, some authors have considered the combination of this modulation with the use of antenna arrays. [Won01] considers a MIMO channel and the OFDM modulation in a joint single beamforming approach. That means that a single symbol is transmitted through each carrier of the OFDM modulation using the best spatial eigenmode of the channel at each frequency. In that paper, the same power is allocated

to all the carriers. [Pal03a] and [Pal04] extend and generalize the previous results by considering several data streams that are transmitted simultaneously through the MIMO channel at each subcarrier under a *joint multiple beamforming* approach and according to several optimization criteria and constraints.

As stated previously, one of the main advantages of using antenna arrays is that it is possible to implement the SDMA in a multi-user scenario. By means of this, it is possible to increase substantially the number of users that can be served simultaneously. In [RF98b], an uplink transmission in a flat fading channel is considered with one BS using an antenna array and several single-antenna MT's, whose transmitted signals are detected by the BS. The problem consists in finding the optimum beamvectors at the BS and the transmit powers at the MT's so that the total transmit power is minimized subject to QoS constraints formulated as a minimum SNIR for each link. In that paper, an iterative algorithm is proposed able to find the global optimum design, in addition to indicate that also the base assignment problem can be introduced in the optimization. In [Ben01], the inverse case is analyzed, i.e., a downlink scenario where the transmit beamvectors have to be calculated. An additional problem in this case is that all the MT's must have all the CSI regarding all the mobiles in the same area. In that work, it is shown that the optimization problem is NP-complete, quadratic, and non-convex. In order to find the optimum solution, the problem can be reformulated as an equivalent "virtual uplink problem" and be solved by means of the iterative algorithm proposed in [RF98b], which is able to find the global optimum design. Another way of solving the stated optimization consists in reformulating it as a convex problem [Boy00] and using efficient iterative algorithms such as the "interior point method". The formulation as a convex problem permits the introduction of other kind of constraints related to a feasible implementation of the design in a real system. In [Ben01], other design strategies with heuristic cost functions are proposed in which each MT needs only to know the CSI regarding the channel between itself and the BS. [RF98a] also analyzes the uplink and downlink cases.

[Cha02] extends the results in [RF98b] to the case of an uplink flat fading channel in which both the MT's and BS have multiple antennas. The optimization problem consists in designing the optimum transmit and receive beamvectors in two different situations. In the first one, the minimization of the total transmit problem is addressed, taking into account that the SNIR for each MT must be higher than a prefixed threshold. In the second approach, the objective is to maximize the minimum SNIR subject to a total transmit power constraint. Different iterative algorithms are proposed to calculate the beamvectors. The main problem is that these algorithms may converge to a local suboptimum design as the stated optimization problems are not convex [Boy00]. The reason for that is that a coupling effect appears in the problem. When calculating the optimum receive beamvector for a concrete communication, the interference covariance matrix for this communication has to be calculated. This covariance matrix depends on

the transmit beamvectors for the other communications in the system, producing the coupling effect in the design and making the optimization problem non-convex and highly nonlinear.

Previously to [Cha02], other works and papers have been published where the same problem has been addressed, such as [Jan98] and [Lok00]. In all of them, the same optimization mathematical problem as in [Cha02] is obtained, in which the main task is to solve the coupling effect between the transmitters and the receivers design. In [Jan98], a multi-user scenario is analyzed with no space diversity at the transmitter and the receiver. There, the diversity is introduced in the system by utilizing a set of available signature waveforms. In that work, an iterative algorithm based on the AM technique is proposed to minimize the MSE subject to a total transmit power constraint, although it is admitted that it may converge to a local suboptimum design. [Lok00] considers a multi-user system using MC-CDMA [Wan00]. The goal is to design the transmitters and receivers for all the users, where each user is assigned a different random signature common to all the carriers. The transmitter has to decide how to distribute each symbol of each user among the different carriers, whereas the receiver has to combine the samples at all the carriers in an optimum way so that the interferences created by other users are minimized. Also in this case, the diversity is introduced by means of considering several available frequencies or carriers, and the objective is to minimize the total transmit power taking into account that the SNIR must be higher than a prefixed threshold for each user. A criterion in order to decide whether a feasible solution exists is provided, and an iterative algorithm is proposed to find a solution, that may be suboptimum due to the non-convexity behavior of the problem. The iterative algorithm is based on the Lagrange multipliers method and the quadratic penalty function [Ber82].

2.2.3 Designs with Partial or Imperfect CSI at the Transmitter

In the previous subsections, two opposite cases have been considered: no CSI and perfect CSI at the transmitter side when carrying out the design of the communication system. In a realistic implementation, the most adequate assumption would consist in considering that the transmitter has some *imperfect or partial CSI*, which means that only a limited information about the channel or the interferences behavior can be used when designing the transmitter in order to increase the performance when compared to a system in which no CSI is available. The simplest example is represented by a transmitter that has a noisy estimate of the channel response instead of the actual channel realization. In this subsection, the state of the art considering the imperfect or partial CSI case is summarized from two different points of view:

- The transmitter has a noisy estimate of the channel and the design takes into account explicitly this fact, obtaining robust solutions less sensitive to the uncertainties in the channel estimate.
- A low-rate feedback channel from the receiver to the transmitter is implemented, so that a

much more simple information than the whole channel estimate is sent in order to improve the quality of the transmission when compared to a system designed with no CSI.

In this subsection, the attention is mainly focused on the partial CSI at the transmitter. In §2.2.1, some references have been already given studying the performance degradation of a system in which space-time codes are used and only noisy estimates of the channel are available at the receiver. In §5.3, some additional references are provided in which the same analysis is performed for systems in which a multi-antenna transmitter is designed according to a channel estimate, which is assumed to be perfect despite not being true.

Robust Designs with Imperfect CSI at the Transmitter

In [Nar98], a transmitter with multiple antennas in a Rayleigh flat fading channel is considered. The objective is to design the transmitter using only a channel estimate in order to maximize the mean SNR and the mutual information from a Bayesian perspective, i.e., a pure statistical approach is taken. Two different kinds of errors in the channel estimate are considered: Gaussian and quantization noise. The receiver is assumed to have a perfect CSI. In that work, the solution to the stated design problem is found for two different scenarios: one single-antenna receiver and multiple single-antenna receivers. For the case of maximizing the mean SNR, the optimum solution corresponds to beamforming, whereas in the case of maximizing the mutual information, the general solution may correspond to the transmission through several eigenmodes of the channel covariance matrix combined with an adequate power allocation. A similar approach is taken in [Jön02] for a flat fading MIMO channel. There, the transmitter combines an OSTBC and a matrix carrying out a linear transformation of the outputs of the block encoder before transmitting the signals through the antennas. The design of the matrix is based on the channel estimate available at the transmitter, taking as the objective the minimization of the Chernoff upper-bound on the mean error probability from a Bayesian perspective, i.e., averaging over the statistics of the actual channel conditioned to the channel estimate, and assuming that the receiver uses a ML detector based on a perfect CSI. An asymptotic analysis is carried out showing that, when the quality of the CSI is perfect, the solution tends to be a beamformer, whereas in the case of no CSI, the optimum transmitter is based only on the OSTBC.

[Rey02a], [Rey02b] and [Rey03] propose an optimum design of a single-user MIMO system using the OFDM modulation. In these works, the obtained solution assumes an imperfect CSI at both the transmitter and the receiver and is based on linear pre- and post-processing structures. The ultimate objective is to minimize the MSE or an upper-bound on the mean error probability averaged over the statistics of the actual channel conditioned to the channel estimate, i.e., a Bayesian point of view of the problem is taken.

In some of the references considered above, a Bayesian perspective [Kay93] has been taken

for the robust design of the optimum transmitter. Statistically, this corresponds to the optimum approach, but from a realistic point of view, this is not always a feasible solution since it is necessary to know the a-priori statistical distribution of the channel in a very accurate way. Other possible approaches to design robust systems could be based on the optimization of the *worst case performance*. Mainly, the idea consists in assuming an *uncertainty region* around the channel estimate and trying to improve and optimize the worst performance in this region, i.e., a *maximin* point of view of the problem is considered as described in [Kas85]. [Vor03] applies this technique to the design of a robust receive beamformer in which a mismatch is assumed between the actual spatial signature and the presumed one. In [Pal03b], the optimal solution for a vector communication through a MIMO channel in the sense of satisfying a set of QoS requirements for the simultaneously established substreams with minimum transmit power is derived. A robust design under channel estimation errors is also considered by taking also the worst-case approach. For this kind of maximin approaches, the uncertainty region plays an important role and should be directly related to the quality of the CSI if high performance designs are desired.

A more detailed and complete description of the state of the art is given in §5.4.

Improved Designs with Low-Rate Feedback from the Receiver

In [Sca02], it is considered that the transmitter has no channel estimate, but knows its statistical distribution, i.e., its pdf, which is assumed to be stationary and follow a frequency selective Nakagami-m distribution [Pap91]. That could correspond to a system in a variant channel environment in which the low-rate of the feedback is not high enough so as to transmit an updated channel estimate. In that work, the simple SISO case is assumed and the objective is to find the optimum power allocation for the OFDM modulation minimizing the mean BER averaged over the channel statistics. A similar approach is assumed in [Vis01] for a MISO flat fading channel. The rate of the feedback channel is not high enough to transmit each channel estimate. Instead of that, only a limited statistically description of the channel is available at the transmitter, either its mean or its covariance matrix, assuming a channel with complex Gaussian coefficients. The design goal is to find the transmitter maximizing the mutual information averaged over the presumed channel statistics. In general, the optimum solution implies a high rank of the transmitted signal covariance matrix. The authors also propose a relationship between the quality of the channel estimate and the statistically characterization of the channel. The extension of the previous results is presented in [Mou02] and [Sim02], analyzing both the ergodic and the outage mutual information for the two possibilities concerning the channel knowledge described in [Vis01]. Besides, they evaluate the situations in which beamforming is the optimum solution from an information theory point of view.

Other papers such as [Muk01] consider the feedback transmission of parameters related to

each channel realization. Concretely, in that work, a MISO flat fading scenario is analyzed, where the channel parameters that are allowed to be known at the transmitter are either its phases or its gains. For the case of phase only knowledge, the transmitter minimizing the mean error probability is an equal gain beamformer with phases matched to the channel realization, whereas in the gain only knowledge, the optimum solution is antenna selection choosing the branch with the highest gain.

A more general case corresponding to a MIMO flat fading channel has been addressed in [Cho02d]. In that paper, it is shown that for a concrete channel realization, the optimum transmit beamvector is equal to the right singular vector associated to the maximum singular value of the channel matrix. The receiver is assumed to have a perfect estimate of the channel and is responsible to send the optimum beamvector to the transmitter through the digital feedback channel. Instead of sending directly the optimum beamvector, the receiver sends an "error vector" correcting the current beamvector being used at the transmitter, i.e., a differential encoding of the optimum transmit beamvector is carried out. This error vector belongs to a discrete set and, therefore, only a few bits are necessary to be sent through the feedback channel. In that work, it is deduced that asymptotically, if the adaptation gain in the beamvector updating process is low enough, this transmit beamvector tends to the optimum maximum right singular vector. This simple technique works well if the updating rate is chosen accordingly to the Doppler frequency of the channel.

[Cho02a] analyzes the case of a FDD MISO frequency selective channel using a CDMA modulation and a single-antenna RAKE receiver. Taking this into account, the signal model is equivalent to a flat fading single-user MIMO channel. The optimum beamformer at the transmitter is also the maximum eigenvector of the channel matrix, which is calculated by the transmitter based on a local estimate. This local estimate is obtained based on the geometric reconstruction of the multipath propagation in the scenario. The angles of the multipath components are estimated by the transmitter during the transmission in the inverse link, whereas the gains are sent from the receiver to the transmitter through a feedback channel. At the transmitter, a linear predictor is used to update the values of the gains so as to cope well in channels with a non-zero Doppler frequency and feedback delay.

Finally, in [Lar02] and [Akh03], the objective is to increase the performance of OSTBC by means of using limited feedback in a flat fading MISO system. In the first paper, a weighting matrix is applied after the space-time encoder. This matrix belongs to a finite set, and the receiver is responsible for choosing the optimum matrix based on its channel knowledge. As the set is finite and discrete, only a few bits are necessary to be sent through the feedback channel. If all these matrices are full rank, full diversity order is guaranteed whatever choice. For the concrete case of diagonal weighting matrices and fixed rate designs, it is shown that the optimum solution corresponds to antenna selection, i.e., only the antenna with the highest

gain is chosen for transmission. In case that there are errors in the feedback channel, antenna selection is shown to produce a loss in the diversity gain. In order to solve this problem, a robust technique is proposed so that the total transmit power is distributed among the antennas and the diversity gain is not reduced. Finally, in [Akh03], an algorithm is proposed to extend OSTBC to any number of transmit antennas and complex constellation. This can be done while maintaining full diversity gain and full rate with only a very limited channel knowledge. The idea relies on using different codes on different non-overlapping groups of antennas, and applying a phase, which depends on the channel, to each group. By means of this, all the signals transmitted through the different encoders and groups are coherently combined at the receiver. In [Akh03], it is shown that only $p-1$ bits are necessary in the feedback so that a system with p antenna groups performs at full diversity gain.

2.3 Mathematical Preliminaries

In this section, some mathematical preliminaries are given concerning the theory of *convex optimization*. These preliminaries will be useful in some parts of this dissertation, in which the theory of convex optimization will be applied to solve some constrained optimization problems.

2.3.1 Convex Optimization

The theory of convex optimization [Roc71a, Boy04] provides a general framework for solving many constrained optimization problems. The main advantage is that, thanks to this theory, closed-form solutions can be found to many problems under some mild conditions and based on the application of the KKT conditions, as will be shown later in this section. In case that a closed-form solution does not exist or cannot be found, the solution to a convex optimization problem can always be calculated by applying efficient numerical methods. Currently, there exists a wide range of algorithms and public software packages that can solve any kind of convex problem in an admissible period of time.

2.3.2 Convex Sets and Convex Functions

Some necessary definitions are first given before presenting some general results on convex optimization, which will be useful in this dissertation.

A set $\mathcal{A}$ is a *convex set* if the line segment between any two points in the set lies in the set. This can be expressed mathematically as

$$\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{A}, \qquad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{A}, \quad \forall \theta \in [0,1]. \tag{2.1}$$

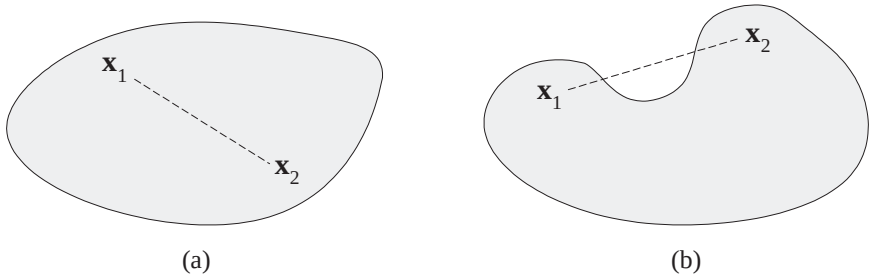


Figure 2.3: (a) Example of a convex set. (b) Example of a non-convex set.

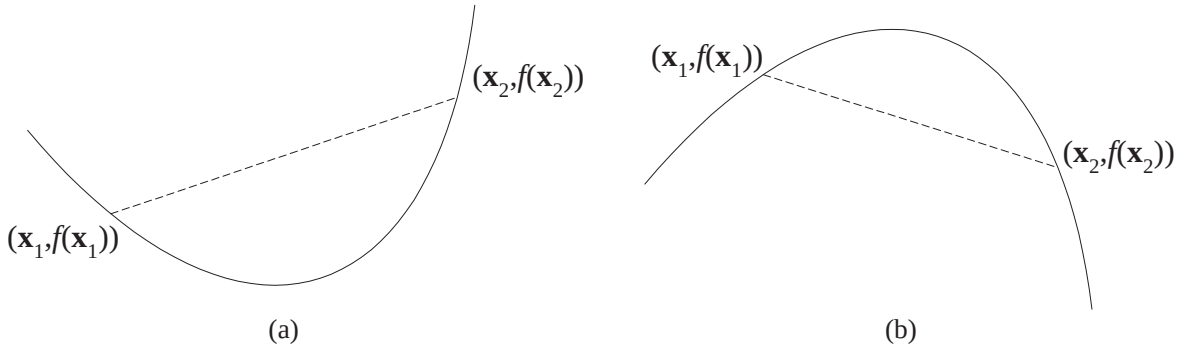


Figure 2.4: (a) Example of a convex function. (b) Example of a concave function.

See Figure 2.3 for a simple example of a convex and a non-convex set. There are many examples of convex sets, such as the vector subspaces, the hyperplanes, and the halfspaces, among others. There are also many properties regarding convex sets. One of them states that the set resulting from the intersection of an arbitrary number of convex sets is also convex. See [Boy04] for a more complete list of examples and properties of convex sets.

A real-valued function f is a *convex function* if its domain $\operatorname{dom} f$ is a convex set and the following inequality holds:

$$f\left(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2\right) \leq \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2), \qquad \forall \mathbf{x}_1, \mathbf{x}_2 \in \operatorname{dom} f, \quad \forall \theta \in [0,1], \tag{2.2}$$

which means that the line segment between $(\mathbf{x}_1, f(\mathbf{x}_1))$ and $(\mathbf{x}_2, f(\mathbf{x}_2))$ lies above the graph of f. The simplest example of convex function is the set of affine (also called linear) functions of the form $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$. In this case, the inequality in (2.2) holds with equality. See [Boy04] for a more complete list of examples and properties of convex functions.

A *concave function* is defined as a function f such that its domain is convex and the following inequality is fulfilled:

$$f\left(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2\right) \geq \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2), \qquad \forall \mathbf{x}_1, \mathbf{x}_2 \in \operatorname{dom} f, \quad \forall \theta \in [0,1], \tag{2.3}$$

which is equivalent to saying that the function $-f$ is convex. Note that, according to these definitions, the linear functions are simultaneously convex and concave. See Figure 2.4 for an example of a convex and a concave function.

One of the properties of a convex function is that the associated *sublevel sets* are convex, where the α-sublevel set $\mathcal{S}_\alpha f$ is defined as

$$\mathcal{S}_\alpha f \triangleq \{\mathbf{x} \in \operatorname{dom} f : f(\mathbf{x}) \leq \alpha\}. \tag{2.4}$$

In case that a function f is differentiable, it will be convex if and only if its domain is convex and

$$f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \nabla f(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1), \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in \operatorname{dom} f, \tag{2.5}$$

which constitutes the so called first order condition. If f is twice differentiable, it will be convex if and only if its domain is convex and

$$\nabla^2 f(\mathbf{x}) \succeq \mathbf{0}, \quad \forall \mathbf{x} \in \operatorname{dom} f, \tag{2.6}$$

i.e., the Hessian matrix of f is positive semidefinite, which constitutes the second order condition.

2.3.3 Definition of a Convex Problem

The general expression of a *constrained optimization problem* is as follows:

$$\begin{array}{lll} \underset{\mathbf{x}}{\text{minimize}} & f_0(\mathbf{x}) & \\ \text{subject to} & f_i(\mathbf{x}) \leq 0, & 1 \leq i \leq m, \\ & h_i(\mathbf{x}) = 0, & 1 \leq i \leq p, \end{array} \tag{2.7}$$

which consists in finding the infimum of the function $f_0(\mathbf{x})$, which is called the *objective function* or *cost function*, among all $\mathbf{x}$ that satisfy the conditions $f_i(\mathbf{x}) \leq 0,\ i = 1, \ldots, m$, and $h_i(\mathbf{x}) = 0,\ i = 1, \ldots, p$, simultaneously. The variable $\mathbf{x}$ is called the *(primal) optimization variable*, $f_i(\mathbf{x}) \leq 0$ and $f_i(\mathbf{x})$ are the *inequality constraints* and *inequality constraint functions*, respectively, and $h_i(\mathbf{x}) = 0$ and $h_i(\mathbf{x})$ are the *equality constraints* and *equality constraint functions*, respectively.

The set of points $\mathcal{D}$ for which the objective and all the constraint functions are defined,

$$\mathcal{D} \triangleq \bigcap_{i=0}^{m} \operatorname{dom} f_i \cap \bigcap_{i=1}^{p} \operatorname{dom} h_i, \tag{2.8}$$

is called the *domain* of the optimization problem (2.7). A point $\mathbf{x} \in \mathcal{D}$ is *feasible* if it satisfies all the constraints $f_i(\mathbf{x}) \leq 0,\ i = 1, \ldots, m$, and $h_i(\mathbf{x}) = 0,\ i = 1, \ldots, p$, simultaneously. The problem (2.7) is feasible if there exists at least one feasible point, and *infeasible* otherwise. The set of all feasible points is called the *feasible set* or the *constraint set*.

The *optimal value* $f^\star$ of the problem (2.7) is defined as

$$f^\star \triangleq \inf \{ f_0(\mathbf{x}) : f_i(\mathbf{x}) \leq 0,\ i = 1, \ldots, m,\ h_i(\mathbf{x}) = 0,\ i = 1, \ldots, p \} . \tag{2.9}$$

If the problem is infeasible, then $f^\star = +\infty$, whereas if the problem is unbounded below, $f^\star = -\infty$. A point $\mathbf{x}^\star$ is an *optimal solution* or *optimal point*, i.e., solves the problem (2.7), if $\mathbf{x}^\star$ is feasible and $f_0(\mathbf{x}^\star) = f^\star$. The set of all optimal points is the *optimal set*. If there exists an optimal point for the problem, then it is said that the optimal value is *attained* or *achieved*, otherwise the optimal value is not attained or not achieved.

A feasible point $\mathbf{x}$ is said to be *locally optimal* if there is $R > 0$ such that

$$f_0(\mathbf{x}) = \inf \{ f_0(\mathbf{z}) : f_i(\mathbf{z}) \leq 0,\ i = 1, \ldots, m,\ h_i(\mathbf{z}) = 0,\ i = 1, \ldots, p,\ \|\mathbf{z} - \mathbf{x}\| \leq R \} . \tag{2.10}$$

All optimal points are also locally optimal, but there may exist locally optimal points that are not optimal solutions.

A problem (2.7) is said to be a *convex optimization problem* if the cost function and the inequality constraint functions f_i are convex, and if the equality constraint functions h_i are affine, i.e., $h_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{x} + b_i$. An important consequence of this definition is that the domain $\mathcal{D}$ of the optimization problem is convex since it is the intersection of the set of domains of all the functions, which are convex. Additionally, the constraint set is also convex since it is the intersection of m convex sublevel sets and p hyperplanes. A fundamental property of convex optimization problems is that any locally optimal point is also (globally) optimal [Boy04].

Some examples of convex optimization problems are *linear programs*, in which all the functions $f_i,\ i = 0, \ldots, m$, and $h_i,\ i = 1, \ldots, p$, are affine; and *quadratic programs*, in which the objective function is convex quadratic, and all the equality and inequality constraint functions are affine. If the objective function as well as all the inequality constraint functions are convex quadratic, and the equality constraint functions are affine, then the optimization problem is also convex and is called *quadratically constrained quadratic problem*. See [Boy04] for a more complete list of kinds of convex problems, including second order cone programming and semidefinite programming, among others.

A maximization problem where the objective function $f_0(\mathbf{x})$ is concave, all the inequality constraint functions $f_i,\ i = 1, \ldots, m$, are convex, and the equality constraint functions h_i are affine, is a *concave optimization problem*. This problem is equivalent to a convex optimization problem in which the convex objective function $-f_0$ is to be minimized subject to the same constraints.

2.3.4 Duality and Karush-Kuhn-Tucker Conditions

Given the optimization problem (2.7), where its domain is assumed to be non-empty, the *Lagrangian function* L associated to the problem is defined as

$$L(\mathbf{x};\boldsymbol{\lambda},\boldsymbol{\nu}) \triangleq f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}), \tag{2.11}$$

where $\boldsymbol{\lambda} = \begin{bmatrix}\lambda_1 \cdots \lambda_m\end{bmatrix}^T$ and $\boldsymbol{\nu} = \begin{bmatrix}\nu_1 \cdots \nu_p\end{bmatrix}^T$ are the vectors of *Lagrange multipliers*, also called *dual variables*. Based on the Lagrangian function, the *Lagrange dual function* (or just *dual function*) is defined as the minimum value of the Lagrangian over $\mathbf{x}$:

$$g(\boldsymbol{\lambda},\boldsymbol{\nu}) \triangleq \inf_{\mathbf{x}\in\mathcal{D}} L(\mathbf{x};\boldsymbol{\lambda},\boldsymbol{\nu}), \tag{2.12}$$

where $\mathcal{D}$ is the domain of the original problem (2.7). When the Lagrangian is unbounded below, the dual function takes on the value $-\infty$. The dual function $g(\boldsymbol{\lambda},\boldsymbol{\nu})$ is concave even if the original problem is not convex, since g is the infimum of a family of affine functions of $(\boldsymbol{\lambda},\boldsymbol{\nu})$ (see [Boy04]). A point $(\boldsymbol{\lambda},\boldsymbol{\nu})$ is *dual feasible* if $\lambda_i \geq 0,\ i = 1,\ldots,m$, and $g(\boldsymbol{\lambda},\boldsymbol{\nu}) > -\infty$. The dual function evaluated at any dual feasible point is a lower bound on the optimal value $f^\star$ of the original problem (2.7), which is proved as follows. For any feasible $\mathbf{x}$ and $(\boldsymbol{\lambda},\boldsymbol{\nu})$:

$$\begin{aligned} f_0(\mathbf{x}) &\geq f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) && (2.13)\\ &\geq \inf_{\mathbf{z}\in\mathcal{D}} \left(f_0(\mathbf{z}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{z}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{z}) \right) && (2.14)\\ &= g(\boldsymbol{\lambda},\boldsymbol{\nu}), && (2.15) \end{aligned}$$

therefore,

$$g(\boldsymbol{\lambda},\boldsymbol{\nu}) = \inf_{\mathbf{z}\in\mathcal{D}} L(\mathbf{z};\boldsymbol{\lambda},\boldsymbol{\nu}) \leq L(\mathbf{x};\boldsymbol{\lambda},\boldsymbol{\nu}) \leq f_0(\mathbf{x}). \tag{2.16}$$

Since $g(\boldsymbol{\lambda},\boldsymbol{\nu}) \leq f_0(\mathbf{x})$ holds for every feasible point $\mathbf{x}$, it follows the basic result:

$$g(\boldsymbol{\lambda},\boldsymbol{\nu}) \leq f^\star, \tag{2.17}$$

thus, the dual function gives a lower bound on the optimal value of the original problem (2.7), which is also called the *primal problem*. The best lower bound can be found by solving the following optimization problem, which is concave and is called the *(Lagrange) dual problem*:

$$\begin{array}{ll} \underset{\boldsymbol{\lambda},\boldsymbol{\nu}}{\text{maximize}} & g(\boldsymbol{\lambda},\boldsymbol{\nu}) \\ \text{subject to} & \lambda_i \geq 0, \quad i = 1,\ldots,m, \end{array} \tag{2.18}$$

where the optimal value is denoted by $d^\star$ and is attained at the dual optimal variables $(\boldsymbol{\lambda}^\star,\boldsymbol{\nu}^\star)$. This optimal value corresponding to the dual problem is the best lower bound on $f^\star$. In particular,

$$d^\star \leq f^\star, \tag{2.19}$$

which is called *weak duality* and holds even if the original problem is not convex. In case that

$$d^\star = f^\star, \tag{2.20}$$

then it is said that *strong duality* holds or that *the duality gap is zero*.

Strong duality does not always hold. However, in case that the original problem (2.7) is convex, and under some mild technical conditions, strong duality can be assured. The conditions that assure strong duality are called *constraint qualifications*. An example of them is given by the *Slater's conditions*, that state that strong duality holds if there exists a feasible $\mathbf{x}$ such that $f_i(\mathbf{x}) < 0,\ i = 1, \ldots, m$ (i.e., a strictly feasible point exists).

Under the assumption that the duality gap is zero, the following expression holds for the primal optimal $\mathbf{x}^\star$ and dual $(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ variables:

$$\begin{aligned} f_0(\mathbf{x}^\star) &= g(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star) && (2.21)\\ &= \inf_{\mathbf{x}\in\mathcal{D}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^\star f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^\star h_i(\mathbf{x}) \right) && (2.22)\\ &\leq f_0(\mathbf{x}^\star) + \sum_{i=1}^{m} \lambda_i^\star f_i(\mathbf{x}^\star) + \sum_{i=1}^{p} \nu_i^\star h_i(\mathbf{x}^\star) && (2.23)\\ &\leq f_0(\mathbf{x}^\star). && (2.24) \end{aligned}$$

From the previous equations it is concluded that all the inequalities are fulfilled with equality. A result from this observation is:

$$\sum_{i=1}^{m} \lambda_i^\star f_i(\mathbf{x}^\star) = 0 \quad \Rightarrow \quad \lambda_i^\star f_i(\mathbf{x}^\star) = 0, \quad i = 1, \ldots, m, \tag{2.25}$$

since each of the terms $\lambda_i^\star f_i(\mathbf{x}^\star)$ are nonpositive and $\nu_i^\star h_i(\mathbf{x}^\star)$ are equal to 0, due to the feasibility of the variables $\mathbf{x}^\star$ and $\boldsymbol{\lambda}^\star$. Note also that $\mathbf{x}^\star$ minimizes $L(\mathbf{x}; \boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ over $\mathcal{D}$.

Using all these results, the following expressions have to be fulfilled for any primal optimal $\mathbf{x}^\star$ and dual optimal $(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ variables, in case that all the functions are differentiable and if strong duality holds:

$$\begin{aligned} h_i(\mathbf{x}^\star) &= 0, \qquad i = 1, \ldots, p, && (2.26)\\ f_i(\mathbf{x}^\star) &\leq 0, \qquad i = 1, \ldots, m, && (2.27)\\ \lambda_i^\star &\geq 0, \qquad i = 1, \ldots, m, && (2.28)\\ \lambda_i^\star f_i(\mathbf{x}^\star) &= 0, \qquad i = 1, \ldots, m, && (2.29)\\ \nabla f_0(\mathbf{x}^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla f_i(\mathbf{x}^\star) + \sum_{i=1}^{p} \nu_i^\star \nabla h_i(\mathbf{x}^\star) &= \mathbf{0}, && (2.30) \end{aligned}$$

which are the so called KKT conditions.

To summarize, for *any* optimization problem with differentiable objective and constraint functions for which strong duality holds, any pair of primal and dual points must satisfy the KKT conditions, i.e., they are necessary conditions. Additionally, for the case of convex problems, the KKT are also sufficient conditions, i.e., if the points $\mathbf{x}$ and $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ satisfy the KKT conditions, then they are optimal points for the primal and dual problems and have a zero duality gap (note that the feasibility of the points is assured by the KKT conditions).

2.3.5 Solving Convex Problems

As commented, in case that the objective and constraint functions in a convex problem are differentiable and strong duality holds, which is the most usual situation, the KKT conditions are necessary and sufficient conditions for a solution to be optimal. Hence, these equations can be used to find a closed-form solution to the optimization problem, if possible.

Note, however, that even in the case that a closed-form solution cannot be found, an optimal solution can always be calculated by applying an adequate numerical method. Currently, there exist many efficient algorithms for solving any kind of convex problem with excellent convergence properties and not suffering from the usual problems of traditional methods, such as sensitivity to the algorithm initialization. Among the efficient algorithms, the so called *interior point methods* (also called barrier methods), which were first introduced by Nesterov and Nemirovski in [Nes94], have to be highlighted. These algorithms find the optimal solution by solving a sequence of smooth (i.e., continuous second derivatives are assumed) unconstrained problems, usually using the Newton's method [Ber99, Boy04]. There also exist interior-point based methods that are able to calculate not only the optimal value of the primal variables, but also the optimal dual variables, i.e., the Lagrange multipliers (these techniques are called the *primal-dual interior point methods*).

Even when the functions involved in the optimization problem are not differentiable, the solution can also be found by applying algorithms such as the *cutting-plane and ellipsoidal methods* [Gof99, Boy04]. These techniques start with the feasible set, and iteratively divide it into two halfspaces rejecting the one that is known not to contain any optimal solution (in the case of cutting-plane), or sequentially reduce and ellipsoid known to contain an optimal solution. According to this, the optimal solution is asymptotically found.

Chapter 3

Joint Beamforming Design in MIMO-OFDM Single-User Communications

3.1 Introduction

As explained previously, the use of space diversity, consisting in the exploitation of multiple antennas at the transmitter and/or the receiver, is an effective method to improve the performance of the communication link and the capacity of the scattered wireless channel while maintaining the required transmit power and bandwidth. The most general case corresponds, as commented in the previous chapters, to a MIMO channel, in which multiple antennas are available simultaneously at both the transmitter and the receiver.

The design of the communication system depends on the quantity and the quality of the CSI available at both sides of the communication link. Obviously, the highest performance is attained when the channel estimate is perfect and known at the transmitter and the receiver. In this situation, and if the bit rate is fixed, joint beamforming schemes can be applied to maximize the SNIR, leading to solutions in which the best spatial eigenmode is chosen for transmission.

In [Lo99], this approach is taken and the problem of designing the optimum transmit beamformer is addressed, although no CCI is considered, i.e., the interferences contribution at the receiver is not taken into account. A more general case corresponding to a single-carrier signaling through a MIMO frequency selective channel is analyzed in [Yan94b] and [Yan94a], where the optimum transmit and receive filters are designed in the frequency domain taking into account the presence of interferences at the receiver and according to the MMSE criterion.

In many practical systems, such as the terrestrial broadcasting standards DVB-T and T-DAB, and the WLAN's HiperLAN/2 [ETS00] and IEEE 802.11a [IEE99], the multicarrier OFDM

[Bin90, Nee98, Wan00] modulation is used. This multicarrier signaling has also been adopted by some wireline standards, such as those corresponding to DSL, in which the DMT modulation is used. Thanks to the multicarrier modulation, the transmission bandwidth is divided into several parallel subchannels or subcarriers, so that the equivalent channel for each of them can be assumed to be frequency flat, even when the channel response is time dispersive and, therefore, frequency selective. The complexity of the OFDM modulation is extremely low, since the modulation and demodulation processes, based on the IFFT and the FFT, respectively, and the equalization, have a very low computational load.

In the literature, there are some works in which the combination of OFDM with MIMO channels has been addressed. In [Won01], a MIMO channel combined with the use of the OFDM modulation is considered, assuming a perfect CSI at both the transmitter and the receiver. In that paper, the design of the transmit and receiver beamformers is deduced, so that the SNIR per carrier is maximized. However, the total available transmit power is distributed in a uniform way among the subcarriers, leading to an inefficient power allocation.

This chapter is devoted to the design of a single-user MIMO-OFDM communication system in which the CSI is assumed to be perfect and complete at both the transmitter and the receiver. The system, whose architecture is based on a joint beamforming approach per carrier, is designed so that the beamvectors maximizing the SNIR are obtained, and appropriate power allocation policies among the subcarriers of the OFDM modulation are found. The proposed power allocation strategies have a low computational load and are based on several global performance functions.

These techniques were afterwards applied to the case of multiple beamforming and taking general optimization criteria under the theories of convex optimization [Roc71a, Boy04] and majorization [Mar79]. In [Pal03a], the design is addressed in a multicarrier system under a global transmit power constraint, whereas in [Pal04], the problem consists in minimizing the transmit power subject to QoS constraints.

The chapter is organized as follows. In Section 3.2, the main properties of the OFDM modulation are first summarized in §3.2.1. Afterwards, in §3.2.2, the system model corresponding to the combination of MIMO channels and OFDM based on joint beamforming is deduced. The expression of the transmit and receive beamvectors maximizing the SNIR is obtained in §3.2.3. The proposed power allocation strategies are analyzed in Section 3.3, whereas they are evaluated and compared by means of several simulation results in Section 3.4. Finally, the chapter summary and some conclusions are provided in Section 3.5.

The results obtained in this chapter have been published in [PI04f, PI01a, PI01c, PI02b, PI02e].

3.2 The Communication System

The description of the system considered in this chapter is now presented. A single-user communication is assumed, i.e., only one communication link is to be designed according to a perfect channel estimate available at both sides of the link. The system architecture is based on the combination of the OFDM modulation and a MIMO channel, where the number of transmit and receive antennas is n_T and n_R, respectively.

The final objective consists not only in deducing the optimum joint design of the transmit and receive beamvectors, but also in proposing power allocation strategies among the subcarriers of the OFDM modulation that are suitable for the multicarrier communication and with a lower complexity than other classical solutions.

3.2.1 The OFDM Modulation

The OFDM modulation is a multicarrier signaling scheme that has been adopted by many systems for the physical layer. The general principles of this modulation are now given before explaining how it can be exploited in a MIMO system. See [Bin90, Nee98, Wan00], among other references, for more details about this multicarrier modulation. The presentation is given assuming a SISO channel and no pre-processing at the transmitter, although the extension to MIMO channels with transmit and receive beamforming is direct, as will be shown in the following subsections.

Let us assume a N-carriers OFDM modulation, i.e., the number of subchannels in which the available transmission bandwidth is divided is N, where $k = 0, \ldots, N-1$ is the subcarrier index. The stream of QAM symbols to be transmitted is divided into N parallel substreams using a S/P converter, so that they are transmitted simultaneously through the N subcarriers. Let $s_k(m)$ be the QAM symbol to be transmitted through the kth subcarrier during the mth slot or period of time of duration T_s corresponding to one OFDM symbol (in OFDM, the duration of a symbol corresponds to the period of time during which each of the carriers is modulated by the same symbol). In the following, the information symbols will be assumed to be i.i.d. with zero-mean and to have a normalized energy ($\mathbb{E}\left[|s_k(m)|^2\right] = 1$).

Each OFDM symbol is composed of two parts: the CP, of duration T_{CP}, and the useful part, of duration T, i.e., $T_s = T_{CP} + T$. Usually, and if no oversampling is applied, the sampling frequency used by the DSP's is equal to $f_s = N/T$, so that the total number of samples in one OFDM symbol is $P = D + N$, where $D = NT_{CP}/T$ and N are the number of samples in the CP and the useful part, respectively. The CP is typically added at the beginning of each OFDM symbol, and its D samples are equal to the last D samples of the useful part of the symbol (see Figure 3.1).

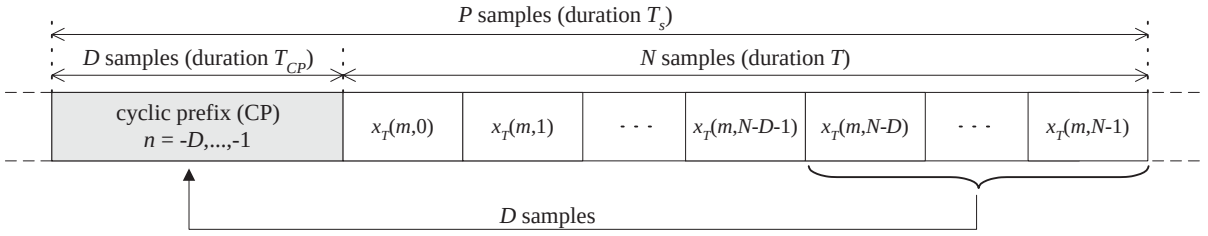


Figure 3.1: Structure of an OFDM symbol composed of the CP and the useful part.

The base-band frequency corresponding to the kth carrier is represented by f_k. In OFDM, the values of these frequencies are $f_k = k/T,\ k = 0, \ldots, N-1$ and, therefore, the carriers are separated by a frequency equal to $1/T$. According to this, the $D+N$ samples corresponding to the transmitted signal in the time domain during the mth OFDM symbol can be expressed as

$$x_T(m,n) \triangleq \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k(m) e^{j2\pi f_k \frac{n}{f_s}} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k(m) e^{j\frac{2\pi}{N}kn}, \quad -D \leq n < N, \tag{3.1}$$

where n is the temporal sample index, the first D samples ($-D \leq n < 0$) correspond to the CP, and the last N ones ($0 \leq n < N$) constitute the useful part of the symbol (see Figure 3.1). Note that the generation of the samples in (3.1) can be easily implemented through a N-points unitary IFFT (the IFFT is said to be unitary since the factor $1/\sqrt{N}$ is added in (3.1), so that the energy is preserved after the inverse Fourier transform). In Figure 3.2, the block diagram of the OFDM modulator is shown. Consequently, the final transmitted sampled signal resulting from the concatenation of all the transmitted OFDM symbols is

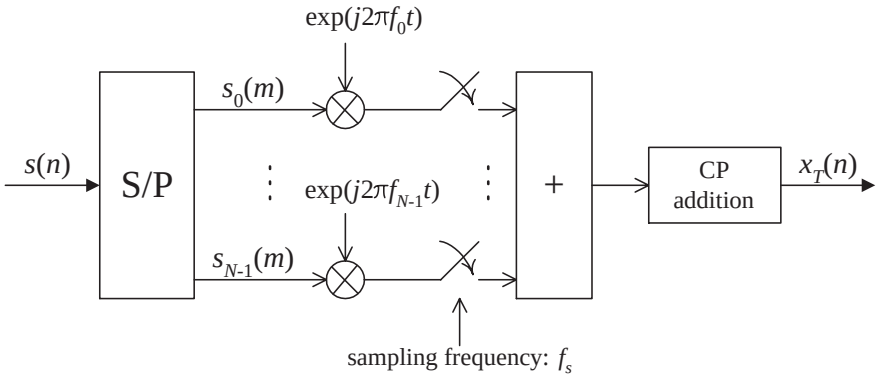
$$x_T(n) \triangleq \sum_{m=-\infty}^{\infty} x_T(m, n - mP). \tag{3.2}$$

The channel is assumed to be time invariant and frequency selective and, therefore, it has to be represented by a time response with a number of taps L generally greater than 1. Let $h(n),\ n = 0, \ldots, L-1$, represent the time response of the wireless channel sampled at the frequency f_s. Thus, the received signal $x_R(n)$ can be expressed as

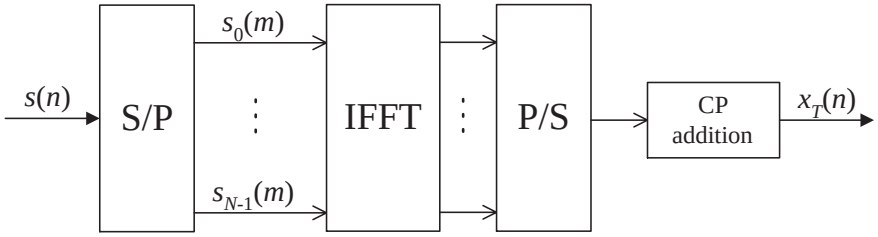
$$x_R(n) = \sum_{l=0}^{L-1} h(l) x_T(n-l) + n(n) = x_T(n) * h(n) + n(n), \tag{3.3}$$

where $n(n)$ represents the noise plus interferences contribution at the receiver. If the number of samples of the CP D is equal to or higher than the channel order $L-1$, then the equalization process and the detection of the QAM symbols is extremely simplified. As shown in Figure 3.3, the detection is based on the output of the unitary FFT applied to the received samples. Let us define the variable $y_k(m)$ as

$$y_k(m) \triangleq \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_R(mP+n) e^{-j\frac{2\pi}{N}kn}, \tag{3.4}$$



(a)



(b)

Figure 3.2: Block diagram of an OFDM modulator. (a) OFDM modulator as a bank of complex subcarriers modulated by the QAM symbols at the output of the S/P converter. (b) Practical implementation of the OFDM modulator based on the application of a unitary IFFT.

resulting from the removal of the CP corresponding to the mth OFDM symbol and the application of the unitary FFT to the samples of the useful part of the symbol. Under the presumed assumptions regarding the channel length ($D \geq L-1$), the variable $y_k(m)$ can be shown to be

$$y_k(m) = H(k)s_k(m) + n_k(m), \tag{3.5}$$

where $H(k)$ and $n_k(m)$ represent the Fourier transform of the channel and the noise plus interferences contribution:

$$H(k) \triangleq \sum_{n=0}^{L-1} h(n)e^{-j\frac{2\pi}{N}kn}, \qquad n_k(m) \triangleq \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} n(mP+n)e^{-j\frac{2\pi}{N}kn}. \tag{3.6}$$

Consequently, according to (3.5), the transmission of the symbol $s_k(m)$ through the kth subcarrier is equivalent to the transmission through a flat fading channel with a gain equal to $H(k)$ and a noise plus interferences contribution equal to $n_k(m)$.

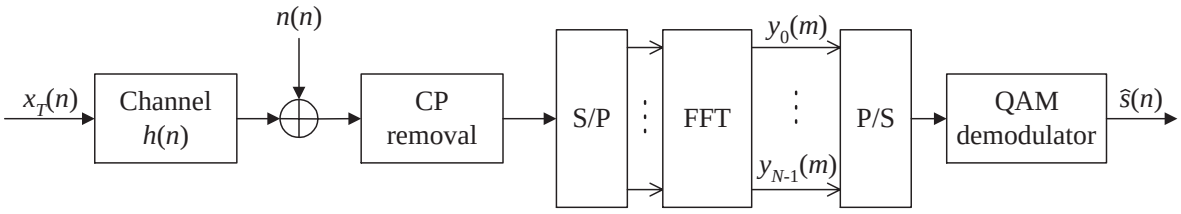


Figure 3.3: Block diagram of the OFDM demodulator based on the application of the unitary FFT after the CP removal for each OFDM symbol.

3.2.2 The MIMO Channel

In this chapter, the objective is to design the transmitter and the receiver in a communication system where both sides of the link have multiple antennas and the signaling scheme is based on the OFDM modulation. In this case, a perfect CSI is assumed to be available at both the transmitter and the receiver.

The proposed front-end architecture, as shown in Figure 3.4, is based on joint beamforming. The N QAM symbols to be transmitted in the same OFDM symbol are collected in the vector $\mathbf{s}(m) \triangleq \left[s_0(m) \cdots s_{N-1}(m)\right]^T \in \mathcal{C}^{N\times 1}$, where $\mathbb{E}\left[\mathbf{s}(m)\mathbf{s}^H(m)\right] = \mathbf{I}$, since the symbols are i.i.d. with a normalized energy and zero-mean. The independence of the symbols is true if no channel coding is applied before modulating them. Indeed, all the commercial communication systems use always some coding to improve the performance of the system. In this dissertation, however, the attention is focused on the uncoded performance of the system and, consequently, no coding will be considered and the symbols will be assumed to be independent.

At the transmitter, the symbol corresponding to the kth subcarrier is multiplied by the transmit beamvector $\mathbf{b}_k \triangleq \left[b_1(k) \cdots b_{n_T}(k)\right]^T \in \mathcal{C}^{n_T\times 1}$ before being modulated by the IFFT. That means that, at the pth transmit antenna, the amplitude and the phase of the kth subcarrier is modulated by $s_k(m)b_p(k)$ instead of $s_k(m)$, obtaining the transmitted signal $x_T^{(p)}(n)$ as a result (see Figure 3.5(a)).

The channel is assumed to be time invariant. Let $h_{q,p}(n)$ be the L-taps long time response of the MIMO channel between the pth transmit and qth receive antennas, respectively, with associated frequency response

$$H_{q,p}(k) \triangleq \sum_{n=0}^{L-1} h_{q,p}(n) e^{-j\frac{2\pi}{N}kn}. \tag{3.7}$$

As commented in the previous subsection, the length of the CP will be assumed to be higher than or equal to $L-1$. The received signal samples $x_R^{(q)}(n)$ at the qth antenna result from the linear combination of the transmitted signals from all the transmit antennas taking into account

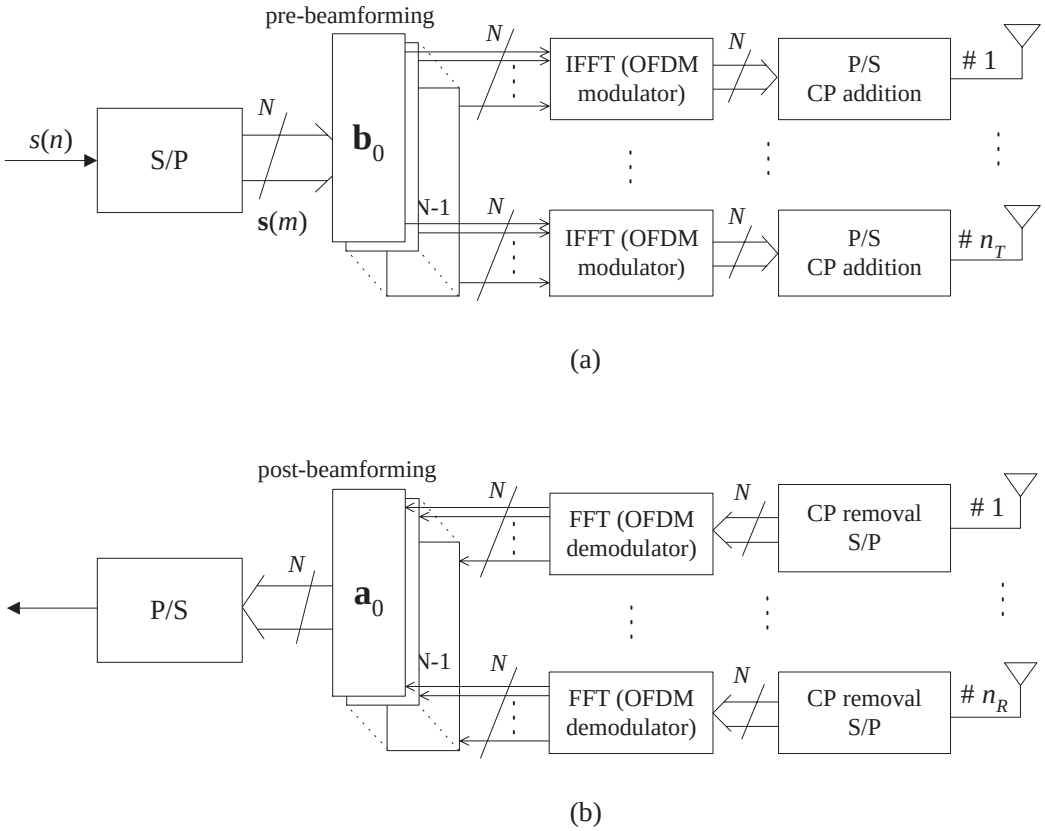


Figure 3.4: Beamforming architecture for the single-user MIMO-OFDM system. (a) Transmitter architecture based on a pre-beamforming per subcarrier. (b) Receiver architecture based on a post-beamforming per subcarrier.

the channel effects and the noise plus interferences contribution:

$$x_R^{(q)}(n) = \sum_{p=1}^{n_T} x_T^{(p)}(n) * h_{q,p}(n) + n^{(q)}(n). \tag{3.8}$$

Taking into account the result (3.5), the sample corresponding to the kth subcarrier after the FFT demodulation at the qth receive antenna can be expressed as

$$y_k^{(q)}(m) = \sum_{p=1}^{n_T} s_k(m) H_{q,p}(k) b_p(k) + n_k^{(q)}(m), \tag{3.9}$$

where $n_k^{(q)}(m) \triangleq \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} n^{(q)}(mP+n) e^{-j\frac{2\pi}{N}kn}$.

All the samples corresponding to all the receive antennas and the kth subcarrier can be collected in a single snapshot vector $\mathbf{y}_k(m) \triangleq \left[y_k^{(1)}(m) \cdots y_k^{(n_R)}(m)\right]^T \in \mathbb{C}^{n_R \times 1}$. Using (3.9), the snapshot vector model can be rewritten in a compact way as

$$\mathbf{y}_k(m) = \mathbf{H}_k \mathbf{b}_k s_k(m) + \mathbf{n}_k(m) \in \mathbb{C}^{n_R \times 1}, \tag{3.10}$$

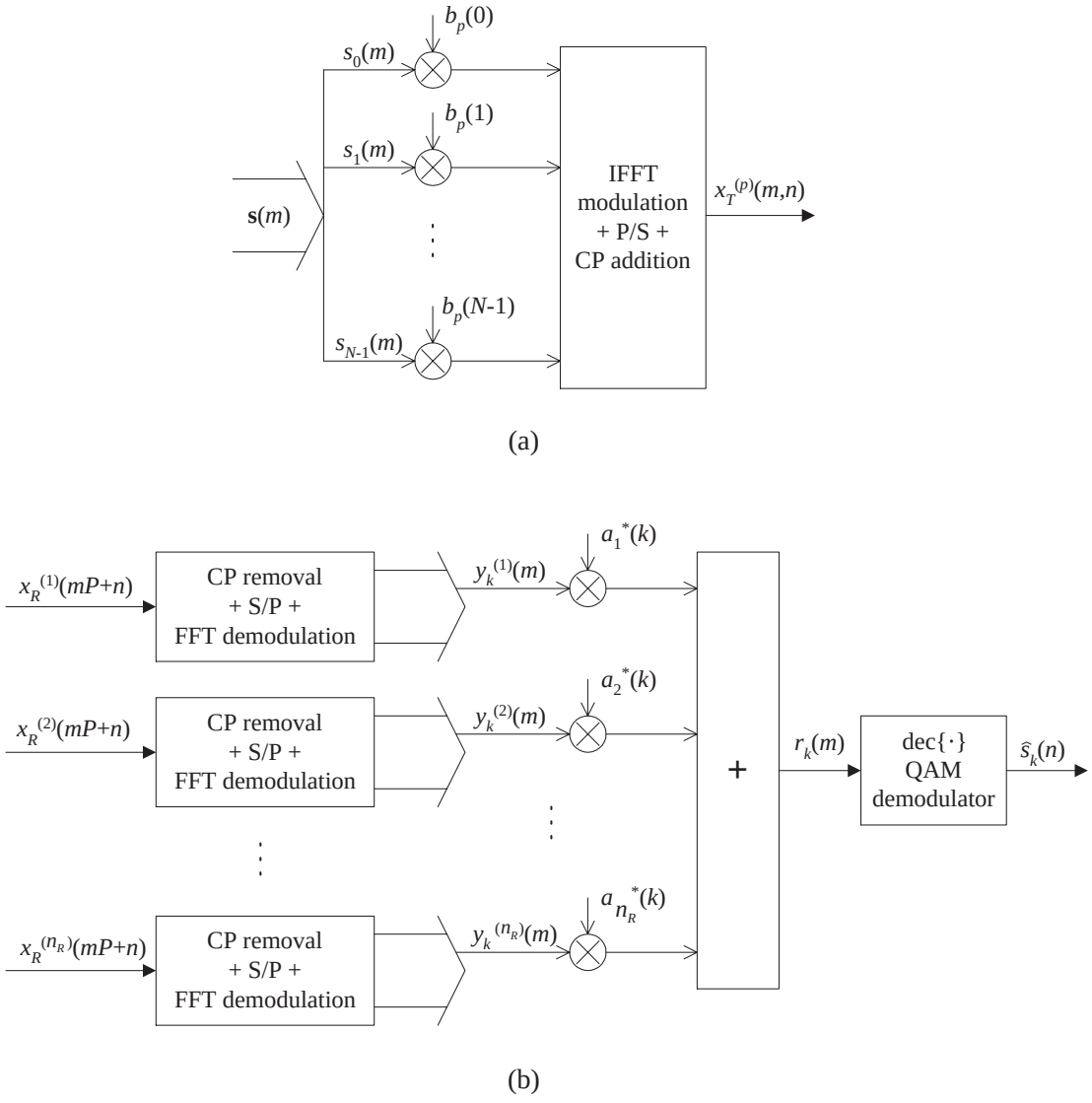


Figure 3.5: Detailed structure of the transmitter and the receiver based on beamforming per carrier. (a) OFDM modulation scheme corresponding to the pth transmit antenna based on beamforming. (b) Receiver structure based on beamforming corresponding to the kth subcarrier.

where

$$\mathbf{H}_k \triangleq \begin{bmatrix} H_{1,1}(k) & \cdots & H_{1,n_T}(k) \\ \vdots & \ddots & \vdots \\ H_{n_R,1}(k) & \cdots & H_{n_R,n_T}(k) \end{bmatrix} \in \mathbb{C}^{n_R \times n_T} \tag{3.11}$$

is the MIMO channel response at the kth subcarrier and $\mathbf{n}_k(m) \triangleq \left[n_k^{(1)}(m) \cdots n_k^{(n_R)}(m)\right]^T \in \mathbb{C}^{n_R \times 1}$ represents the noise plus interferences contribution, i.e., the noise plus interferences snapshot vector, at the same frequency.

The receiver is also based on beamforming, which means that the received samples after

the FFT are combined before detecting the transmitted symbols. The coefficients that are used for the linear combination of the samples corresponding to the kth subcarrier are collected in the receive beamvector $\mathbf{a}_k \triangleq \left[a_1(k) \cdots a_{n_R}(k)\right]^T \in \mathbb{C}^{n_R \times 1}$. Using this, the following operation is performed (see Figures 3.4(b) and 3.5(b)):

$$r_k(m) \triangleq \sum_{q=1}^{n_R} a_q^*(k) y_k^{(q)}(m) = \mathbf{a}_k^H \mathbf{y}_k(m) = \mathbf{a}_k^H \mathbf{H}_k \mathbf{b}_k s_k(m) + \mathbf{a}_k^H \mathbf{n}_k(m), \tag{3.12}$$

where $r_k(m)$ is the soft-estimate of the transmitted symbol $s_k(m)$. The final decision or hard-estimate of the transmitted symbol is obtained by demodulating $r_k(m)$, i.e.,

$$\widehat{s}_k(m) \triangleq \mathrm{dec}\left\{r_k(m)\right\}. \tag{3.13}$$

3.2.3 The Optimum Transmit and Receive Beamvectors

The quality of the detection and the symbol estimation depends directly on the SNIR. For the concrete system and signal models presented previously, the SNIR at the kth subcarrier is expressed as

$$\mathrm{SNIR}_k = \frac{\mathbb{E}\left[|\mathbf{a}_k^H \mathbf{H}_k \mathbf{b}_k s_k(m)|^2\right]}{\mathbb{E}\left[|\mathbf{a}_k^H \mathbf{n}_k(m)|^2\right]} = \frac{|\mathbf{a}_k^H \mathbf{H}_k \mathbf{b}_k|^2}{\mathbf{a}_k^H \mathbf{R}_n(k) \mathbf{a}_k}, \tag{3.14}$$

where $\mathbf{R}_n(k)$ is the correlation matrix of the noise plus interferences contribution, i.e.,

$$\mathbf{R}_n(k) \triangleq \mathbb{E}\left[\mathbf{n}_k(m) \mathbf{n}_k^H(m)\right]. \tag{3.15}$$

An upper-bound on the SNIR can be found by applying the Cauchy-Schwarz inequality[1] to the numerator of (3.14) using the vectors $\mathbf{R}_n^{1/2}(k)\mathbf{a}_k$ and $\mathbf{R}_n^{-1/2}(k)\mathbf{H}_k\mathbf{b}_k$.[2] The following result is then obtained:

$$\mathrm{SNIR}_k = \frac{|\mathbf{a}_k^H \mathbf{H}_k \mathbf{b}_k|^2}{\mathbf{a}_k^H \mathbf{R}_n(k) \mathbf{a}_k} \leq \mathbf{b}_k^H \mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k \mathbf{b}_k. \tag{3.16}$$

From (3.16) it is concluded that the upper bound depends only on the transmit beamvector $\mathbf{b}_k$. The receive beamvector $\mathbf{a}_k$ that maximizes the SNIR and attains the upper-bound is given by the equality condition of the Cauchy-Schwarz inequality, leading to

$$\mathbf{a}_k = \alpha_k \mathbf{R}_n^{-1}(k) \mathbf{H}_k \mathbf{b}_k, \tag{3.17}$$

which corresponds to the whitened matched filter or Wiener filter [Pro95, Won01]. The scalar α_k does not affect the SNIR and can be chosen, for example, to obtain a normalized value of the

[1] The Cauchy-Schwarz inequality is given by $|\mathbf{x}^H\mathbf{y}|^2 \leq \|\mathbf{x}\|^2\|\mathbf{y}\|^2$. The equality holds if and only if $\mathbf{x}$ and $\mathbf{y}$ are colinear, i.e., $\mathbf{x} \propto \mathbf{y}$.

[2] The inverse and the Hermitian square root of the matrix $\mathbf{R}_n(k)$ exist always since $\mathbf{R}_n(k)$ is a Hermitian positive definite matrix, i.e., it diagonalizes and all its eigenvalues are strictly positive due to the presence of AWGN [Gol96].

equivalent channel at the detection stage, i.e., $\mathbf{a}_k^H \mathbf{H}_k \mathbf{b}_k = 1$ (in (3.12), it is seen that $\mathbf{a}_k^H \mathbf{H}_k \mathbf{b}_k$ is the factor that multiplies the transmitted symbol $s_k(m)$).

When considering the design of the transmit beamvectors $\{\mathbf{b}_k\}$, the power constraints at the transmitter have to be taken into account. The transmit power through all the antennas for the kth subcarrier is proportional to $\|\mathbf{b}_k\|^2 = \mathbf{b}_k^H \mathbf{b}_k$. The objective is to maximize the SNIR per carrier assuming that the matched filter (3.17) is used at the receiver, whose expression is

$$\mathrm{SNIR}_k = \mathbf{b}_k^H \mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k \mathbf{b}_k, \tag{3.18}$$

subject to a carrier power constraint formulated as $\|\mathbf{b}_k\|^2 = P_k$.

In order to deduce the expression of the optimum transmit beamvector, which has also been found in [Won01], let us use the eigendecomposition of the Hermitian matrix $\mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k$, which is given by [Gol96]

$$\mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^H, \tag{3.19}$$

where $\mathbf{\Lambda}_k = \mathrm{diag}(\{\lambda_1(k), \ldots, \lambda_{n_T}(k)\})$ is the diagonal matrix containing the eigenvalues sorted in decreasing order, i.e., $\lambda_1(k) = \lambda_{\max}\big(\mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k\big) = \lambda_{\max}(k)$, and $\mathbf{U} = \big[\mathbf{u}_1(k) \cdots \mathbf{u}_{n_T}(k)\big]$ is the unitary matrix containing the normalized eigenvectors associated to the eigenvalues, i.e., $\mathbf{u}_1(k) = \mathbf{u}_{\max}\big(\mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k\big)$. Let us also represent the transmit beamvector $\mathbf{b}_k$ to be designed in terms of its coordinates $\boldsymbol{\beta}_k = \big[\beta_1(k) \cdots \beta_{n_T}(k)\big]^T$ when expressed in the basis given by the eigenvectors, i.e., $\mathbf{b}_k = \mathbf{U}_k \boldsymbol{\beta}_k$. Using this representation, the transmit power constraint in terms of the coordinates $\{\beta_i(k)\}$ is given by

$$\mathbf{b}_k^H \mathbf{b}_k = \boldsymbol{\beta}_k^H \mathbf{U}_k^H \mathbf{U}_k \boldsymbol{\beta}_k = \boldsymbol{\beta}_k^H \boldsymbol{\beta}_k = \sum_{i=1}^{n_T} |\beta_i(k)|^2 = P_k, \tag{3.20}$$

whereas the SNIR can be rewritten as

$$\mathrm{SNIR}_k = \mathbf{b}_k^H \mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k \mathbf{b}_k = \boldsymbol{\beta}_k^H \mathbf{U}_k^H \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^H \mathbf{U}_k \boldsymbol{\beta}_k = \boldsymbol{\beta}_k^H \mathbf{\Lambda}_k \boldsymbol{\beta}_k = \sum_{i=1}^{n_T} \lambda_i(k) |\beta_i(k)|^2. \tag{3.21}$$

The maximization of (3.21) subject to the constraint (3.20) corresponds to a linear program with linear constraints, taking $\{|\beta_i(k)|^2\}$ as the optimization variables, whose solution is given by $|\beta_1(k)|^2 = P_k$, $|\beta_i(k)|^2 = 0$, $i > 1$. That means that all the power of the kth subcarrier P_k is allocated to the maximum eigenvector and, therefore, the closed-form solution to the problem of the transmit beamvector design is

$$\mathbf{b}_k = \sqrt{P_k}\, \mathbf{u}_{\max}\big(\mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k\big), \tag{3.22}$$

i.e., the optimum transmit beamvector is a scaled version of the maximum eigenvector of $\mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k$ and, thus, the resulting SNIR is

$$\mathrm{SNIR}_k = \lambda_{\max}(k) P_k, \tag{3.23}$$

which means that the final SNIR depends on the MIMO channel through the gain $\lambda_{\max}(k)$ and on the power distribution among the subcarriers through P_k. In this joint beamforming design, the main role of the MIMO channel when compared to a SISO transmission is to provide diversity through $\lambda_{\max}(k)$.

3.3 Power Allocation Strategies

In a real system, the power constraint refers to a global transmit power constraint, i.e., to the total power P_0 that has to be distributed among the subcarriers according to the CSI, i.e., according to the channel $\{\mathbf{H}_k\}$ and the noise plus interferences correlation $\{\mathbf{R}_n(k)\}$ matrices. This global power constraint, formulated in the frequency domain, is expressed as

$$\sum_{k=0}^{N-1} P_k = P_0. \tag{3.24}$$

The power distribution among the subcarriers depends on the adopted optimization criterion, which is closely related to the definition of a figure of merit that measures a global system performance as a function of the SNIR at all the subcarriers.

In the following, some optimization criteria are identified in order to deduce the corresponding power allocation strategies. These criteria are compared asymptotically to other classical solutions, showing that the proposed techniques are asymptotically equivalent while having a quite lower computational load.

3.3.1 GEOM: Maximization of the Geometric Mean of the SNIR

The first proposed performance function is the geometric mean of the SNIR averaged over the subcarriers $\mathcal{G}\left(\{\mathrm{SNIR}_k\}\right)$, which is defined as

$$\mathcal{G}\left(\{\mathrm{SNIR}_k\}\right) \triangleq \prod_{k=0}^{N-1} \mathrm{SNIR}_k^{1/N} = \prod_{k=0}^{N-1} \left(\lambda_{\max}(k) P_k\right)^{1/N}, \tag{3.25}$$

assuming that all the maximum eigenvalues are greater than 0 because, otherwise, the geometric mean would be equal to 0 for any power distribution.

The power distribution that maximizes the geometric mean subject to the global transmit power constraint (3.24) can be calculated in closed-form, since the geometric mean is a concave function of the power allocation variables $\{P_k\}$ and the KKT conditions can be applied and solved (see §2.3 on convex optimization and Appendix 3.A, where the proof of the concavity of the geometric mean is given and the solution to the KKT conditions is found).

The solution to this problem can be shown to correspond to a uniform power allocation, i.e., the same power is allocated to all the subcarriers:

$$P_k^{\mathrm{GEOM}} = \frac{P_0}{N}, \tag{3.26}$$

and, therefore, the SNIR is given by

$$\mathrm{SNIR}_k^{\mathrm{GEOM}} = \frac{P_0}{N}\lambda_{\max}(k). \tag{3.27}$$

The main problem of the uniform power allocation is that it does not cope well with frequency selective channels, i.e., with scenarios in which the channel or the noise plus interferences correlation matrices depend on the frequency. In the following subsections, other power allocation strategies more suitable for frequency selective channels are proposed, while still having a low computational complexity.

The uniform power allocation is related to the solution that maximizes the mutual information, i.e., the power distribution that achieves capacity [Sha48]. If the capacity is to be achieved, more than one eigenmode of $\mathbf{H}_k^H \mathbf{R}_n^{-1}(k)\mathbf{H}_k$ may be required to be used for transmission [Ral98] at each carrier. In case that the system architecture is constrained so that only one eigenmode is used per carrier, then the maximum one is chosen and the power among the subcarriers is allocated according to the water-filling solution given by [Ral98]

$$P_k^{\mathrm{CAP}} = \max\left\{0, \mu - \frac{1}{\lambda_{\max}(k)}\right\}, \tag{3.28}$$

where μ is a constant calculated so that the global power constraint (3.24) is fulfilled. Note that in this solution, the subcarriers with the highest $\lambda_{\max}(k)$ are allocated more power, and that some of the subcarriers, the ones with the lowest MIMO channel gain $\lambda_{\max}(k)$, may be cancelled, i.e., may be allocated no power ($P_k^{\mathrm{CAP}} = 0$). Consequently, calculating the value of the parameter μ requires some computational effort, since an iterative algorithm has to be applied in order to identify which are the active subcarriers. It is important to remark that the capacity is achieved, not only if the power is distributed in this way, but also if the statistical distribution of the transmitted symbols is Gaussian [Sha48] (assuming that at the receiver, only AWGN is present). Obviously, in a realistic deployment, the modulation format is fixed and the signal constellation is finite and, therefore, the capacity will not be achieved, even if the power allocation (3.28) is applied.

If the global transmit power P_0 is high enough and all the maximum eigenvalues are different from 0, then all the subcarriers are allocated a power different from 0, i.e., all the subcarriers are suitable for transmission. In this case, the parameter μ can be calculated as

$$\mu = \frac{P_0}{N} + \frac{1}{N}\sum_{l=0}^{N-1}\frac{1}{\lambda_{\max}(l)}. \tag{3.29}$$

As a consequence of this result, the following asymptotic result holds, which relates the GEOM and the CAP power allocation techniques:

$$\lim_{P_0 \to \infty} \frac{P_k^{\mathrm{CAP}}}{P_k^{\mathrm{GEOM}}} = 1, \tag{3.30}$$

i.e., both techniques are asymptotically equal when the transmit power is high.

3.3.2 HARM: Maximization of the Harmonic Mean of the SNIR

A possible alternative different from the previous ones consists in distributing the total transmit power P_0 so that the harmonic mean of the SNIR averaged over the subcarriers $\mathcal{H}\left(\{\mathrm{SNIR}_k\}\right)$ is maximized. The harmonic mean is defined as

$$\mathcal{H}\left(\{\mathrm{SNIR}_k\}\right) \triangleq \frac{N}{\sum_{k=0}^{N-1} \frac{1}{\mathrm{SNIR}_k}} = \frac{N}{\sum_{k=0}^{N-1} \frac{1}{\lambda_{\max}(k) P_k}}. \tag{3.31}$$

As shown in Appendix 3.B, the harmonic mean is a concave function of the power allocation variables $\{P_k\}$. Taking into account the theory on convex optimization presented in §2.3, the KKT conditions can be applied to the optimization problem, as also shown in Appendix 3.B, obtaining the following closed-form expression of the optimum power allocation and assuming that all the maximun eigenvalues are greater than 0:

$$P_k^{\mathrm{HARM}} = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\sqrt{\lambda_{\max}(l)}}} \frac{1}{\sqrt{\lambda_{\max}(k)}}, \tag{3.32}$$

and, therefore, the resulting SNIR is

$$\mathrm{SNIR}_k^{\mathrm{HARM}} = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\sqrt{\lambda_{\max}(l)}}} \sqrt{\lambda_{\max}(k)}. \tag{3.33}$$

From (3.32) it is deduced that the power allocated to each subcarrier scales linearly with P_0 and that more power is allocated to those carriers with a lower $\lambda_{\max}(k)$, i.e., to those frequencies in which the channel has a fading and/or the level of the noise plus interferences is high. It is possible that, for a concrete channel or noise plus interferences realization, a deep fading occurs at some subcarrier. In that case, most of the power would be wasted in that frequency and, therefore, the performance of the other subcarriers would be lower. However, this seldom happens in a MIMO channel with a minimum angular and delay spread since, as the number of antennas is higher, the maximum eigenvalues also increase. In Figure 3.6, the cdf's of the maximum eigenvalues are shown for different number of transmit and receive antennas (the notation $n_T + n_R$ is used to indicate that n_T transmit and n_R receive antennas are available). The angular spread is 30^{o} at the transmitter and 15^{o} at the receiver, whereas the power delay profile is that corresponding to model A [ETS98] standardized for the HiperLAN/2 system

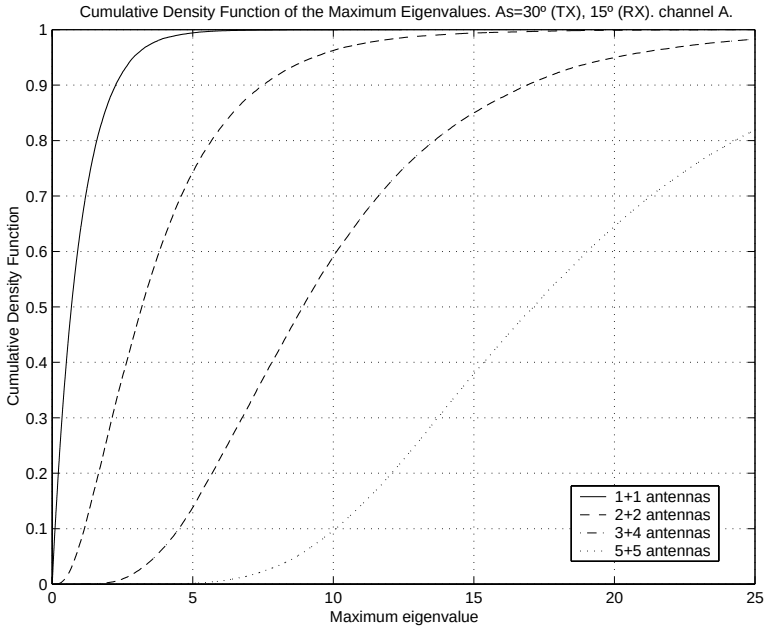


Figure 3.6: Cdf's of the maximum eigenvalues for a MIMO channel with an angular spread of 30° at the transmitter and 15° at the receiver, and a power delay profile corresponding to model A, for different number of transmit and receive antennas.

[ETS00]. The antenna arrays at both the transmitter and the receiver are linear, with a half wavelength ($\lambda/2$) interelement separation. From the figure it is concluded that the probability of deep fades is quite low as the number of antennas increases.

As shown in [PI01c], this criterion is equivalent to the ZF solution, in which the MSE averaged over the subcarriers, which is defined as

$$\xi \triangleq \sum_{k=0}^{N-1} \mathbb{E}\left[|r_k(m) - s_k(m)|^2\right] \tag{3.34}$$

$$= \sum_{k=0}^{N-1} \mathbb{E}\left[|\mathbf{a}_k^H \mathbf{y}_k(m) - s_k(m)|^2\right] = \sum_{k=0}^{N-1} \left(|\mathbf{a}_k^H \mathbf{H}_k \mathbf{b}_k - 1|^2 + \mathbf{a}_k^H \mathbf{R}_n(k) \mathbf{a}_k\right), \tag{3.35}$$

is minimized subject to constraints related to the equivalent channel gains, which must be equal to 1 for all the carriers, i.e., $\mathbf{a}_k^H \mathbf{H}_k \mathbf{b}_k = 1$.

The HARM technique is asymptotically related to the power allocation resulting from the MMSE criterion with no constraint over the equivalent channel gains. The MMSE is a classical design criterion [Pro95] and can be extended to the case of a MIMO channel, as presented in [Yan94b]. Here, it is applied to the case of a multicarrier modulation. The optimization must be carried out by finding the optimum beamvectors at both sides of the communication system,

obtaining the following solution:

$$\mathbf{b}_k = \sqrt{P_k^{\mathrm{MMSE}}}\mathbf{u}_k, \quad \lambda_{\max}(k)\mathbf{u}_k = \mathbf{H}_k^H \mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{u}_k, \quad \|\mathbf{u}_k\| = 1, \tag{3.36}$$

$$P_k^{\mathrm{MMSE}} = \max\left\{0, \frac{\mu}{\sqrt{\lambda_{\max}(k)}} - \frac{1}{\lambda_{\max}(k)}\right\}, \tag{3.37}$$

$$\mathbf{a}_k = \frac{\mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{b}_k}{1 + \mathbf{b}_k^H\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{b}_k}, \tag{3.38}$$

where μ is a constant calculated to satisfy the global transmit power constraint (3.24). From (3.37) it is deduced that the power allocated to each subcarrier does not scale linearly with P_0. Besides, it is possible that, under certain transmit power conditions, the most degraded subcarriers are not suitable for transmission, i.e., $P_k^{\mathrm{MMSE}} = 0$. One disadvantage of this strategy is that the parameter μ has to be calculated by means of iterative methods, which increases the computational load when compared to the GEOM and the HARM strategies.

When the transmit power P_0 is high enough and $\lambda_{\max}(k)$ is different from 0 for all the subcarriers, then all the frequencies are suitable for transmission and the constant μ is calculated as

$$\mu = \frac{P_0 + \sum_{l=0}^{N-1}\frac{1}{\lambda_{\max}(l)}}{\sum_{l=0}^{N-1}\frac{1}{\sqrt{\lambda_{\max}(l)}}}. \tag{3.39}$$

As a consequence of this result, the following asymptotic relationship between the HARM and the MMSE power allocation strategies holds:

$$\lim_{P_0\to\infty}\frac{P_k^{\mathrm{MMSE}}}{P_k^{\mathrm{HARM}}} = 1, \tag{3.40}$$

i.e., for high transmit powers, both solutions tend to distribute the available power in the same way.

The behaviors of the MMSE strategy, in the case of high transmit power, and the HARM solution are opposed to the water-filling (CAP distribution), as the first ones allocate more power to the most degraded subcarriers, while the CAP design injects more power in the best frequencies. In the case of a low transmit power, the MMSE strategy behaves in a similar way than CAP, allocating more power to the most degraded frequencies.

3.3.3 MAXMIN: Maximization of the Minimum SNIR

In an OFDM system, the worst subcarrier is the one that makes the error probability rise and, therefore, it is convenient to force the minimum SNIR over the subcarriers to be maximum.

Let us assume that an arbitrary power allocation P_k is applied. Given this power distribution, let $\overline{\mathcal{M}}$ be the set of subcarriers attaining the maximum SNIR, and let $\underline{\mathcal{M}}$ bet the set of carriers

with the minimum SNIR. If $\overline{\mathcal{M}} \neq \underline{\mathcal{M}}$, the minimum SNIR, i.e., the SNIR for the subcarriers in $\underline{\mathcal{M}}$, can be increased by decreasing the power allocated to the carriers in $\overline{\mathcal{M}}$ and increasing the power for the subcarriers in $\underline{\mathcal{M}}$, while still fulfilling the global transmit power constraint (3.24). This procedure can be applied until $\overline{\mathcal{M}} = \underline{\mathcal{M}}$, which corresponds to the situation in which all the carriers of the OFDM modulation have the same SNIR, i.e.:

$$\mathrm{SNIR}_k = \mathrm{SNIR}_{k'} = cte, \quad \forall k, k' \quad \Rightarrow \quad \lambda_{\max}(k) P_k = \lambda_{\max}(k') P_{k'} = cte, \quad \forall k, k'. \tag{3.41}$$

According to this result, the MAXMIN power allocation technique is expressed as follows:

$$P_k^{\mathrm{MAXMIN}} = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\lambda_{\max}(l)}} \frac{1}{\lambda_{\max}(k)}, \tag{3.42}$$

and, consequently, the obtained SNIR is

$$\mathrm{SNIR}_k^{\mathrm{MAXMIN}} = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\lambda_{\max}(l)}}, \tag{3.43}$$

which, as commented previously, is equal for all the subcarriers.

From (3.42) it is deduced that the powers allocated to the subcarriers are proportional to the global transmit power P_0. As in the case of the HARM technique, the MAXMIN solution allocates more power to the subcarriers in which the global frequency response of the channel has a fading and/or the level of the noise plus interferences contribution is high.

The computational load associated to this power allocation strategy is the lowest one compared to all the other stategies, except GEOM, i.e., it has the lowest complexity among the techniques suitable for frequency selective channels. Additionally, in the case of a scenario in which the noise plus interferences contribution are Gaussian distributed, this solution minimizes asymptotically the Chernoff upper-bound on the mean error probability, as deduced in the following.

Let us define the effective error probability $P_{\mathrm{e,eff}}$ as the error probability averaged over all the subcarriers, assuming that no channel coding is applied:[3]

$$P_{\mathrm{e,eff}} \triangleq \frac{1}{N} \sum_{k=0}^{N-1} P_e(k), \tag{3.44}$$

where $P_e(k)$ is the error probability associated to the kth subcarrier

In the case of a scenario in which the interferences and the noise are Gaussian distributed, the error probability corresponding to the kth subcarrier can be expressed approximately as

$$P_e(k) \simeq \alpha_m \mathcal{Q}\left(\sqrt{k_m \mathrm{SNIR}_k}\right), \tag{3.45}$$

[3]The expression of the effective error probability is correct when the detection of the symbol $s_k(m)$ is based only on $r_k(m)$, i.e., when the information received in other frequencies is not used, which is optimum when the interferences and the noise contributions at different frequencies are independent.

where $\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} \mathrm{d}t$, and α_m and k_m are constants depending on the signal constellation applied to the subcarriers [Pro95] (for BPSK, $\alpha_m = 1$ and $k_m = 2$).

The solution to the problem of minimizing the effective error probability (3.44) subject to the global transmit power constraint (3.24) is difficult and cannot be calculated efficiently and, therefore, numerical methods have to used in order to apply this design criterion. An approximated solution to the problem can be found by minimizing the Chernoff upper-bound [Boy04, Ong03] on the error probability deduced from $\mathcal{Q}(x) \leq e^{-x^2/2}$:

$$P_{\mathrm{e,eff}} \leq \frac{\alpha_m}{N} \sum_{k=0}^{N-1} e^{-\frac{1}{2} k_m \lambda_{\max}(k) P_k}. \tag{3.46}$$

The minimization of the expression above subject to the global power constraint (3.24) has been performed in works such as [Ong03], leading to the following solution under the name MEP (minimum effective error probability):

$$P_k^{\mathrm{MEP}} = \frac{2}{k_m} \frac{\max\{0, \log(\lambda_{\max}(k)) + \mu\}}{\lambda_{\max}(k)}, \tag{3.47}$$

where μ is a constant calculated to satisfy the global transmit power constraint (3.24). As in the CAP and the MMSE solutions, in the case of the minimization of the Chernoff upper-bound on the effective error probability, some of the subcarriers may not be suitable for transmission, i.e., $P_k^{\mathrm{MEP}} = 0$ and, therefore, the parameter μ has to be calculated by means of iterative algorithms, increasing the computational load and making the HARM and the MAXMIN approaches more attractive in this sense.

If the transmit power P_0 is high enough and $\lambda_{\max}(k)$ is different from 0 for all the subcarriers, all the carriers are suitable for transmission and the parameter μ can be calculated as

$$\mu = \frac{\frac{k_m P_0}{2} - \sum_{l=0}^{N-1} \frac{\log(\lambda_{\max}(l))}{\lambda_{\max}(l)}}{\sum_{l=0}^{N-1} \frac{1}{\lambda_{\max}(l)}}. \tag{3.48}$$

Using the previous expression, it can be verified that, asymptotically, the following result holds:

$$\lim_{P_0 \to \infty} \frac{P_k^{\mathrm{MEP}}}{P_k^{\mathrm{MAXMIN}}} = 1, \tag{3.49}$$

concluding that, for high transmit powers P_0, both strategies are equivalent.

In case that the interferences in the scenario are not Gaussian distributed, the MEP power allocation does not minimize the upper-bound on the error probability. The optimum technique should then take into account the interferences' statistical distribution, although this is difficult in a realistic and practical system.

Table 3.1: Summary of the different power allocation techniques and asymptotic analysis.

GEOM power alloc.	**CAP power alloc.**	**Asymptotic analysis**
$P_k = \frac{P_0}{N}$	$P_k = \max\left\{0, \mu - \frac{1}{\lambda_{\max}(k)}\right\}$	$\lim_{P_0\to\infty} \frac{P_k^{\mathrm{CAP}}}{P_k^{\mathrm{GEOM}}} = 1$
HARM power alloc.	**MMSE power alloc.**	**Asymptotic analysis**
$P_k = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\sqrt{\lambda_{\max}(l)}}} \frac{1}{\sqrt{\lambda_{\max}(k)}}$	$P_k = \max\left\{0, \frac{\mu}{\sqrt{\lambda_{\max}(k)}} - \frac{1}{\lambda_{\max}(k)}\right\}$	$\lim_{P_0\to\infty} \frac{P_k^{\mathrm{MMSE}}}{P_k^{\mathrm{HARM}}} = 1$
MAXMIN power alloc.	**MEP power alloc.**	**Asymptotic analysis**
$P_k = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\lambda_{\max}(l)}} \frac{1}{\lambda_{\max}(k)}$	$P_k = \frac{2}{k_m} \frac{\max\{0, \log(\lambda_{\max}(k)) + \mu\}}{\lambda_{\max}(k)}$	$\lim_{P_0\to\infty} \frac{P_k^{\mathrm{MEP}}}{P_k^{\mathrm{MAXMIN}}} = 1$

3.3.4 Summary and Asymptotic Relationships

In Table 3.1, all the techniques deduced in this chapter are presented as a summary, showing both the power allocation techniques jointly with the asymptotic relationships among them.

3.4 Simulation Results

Some simulation results are now presented based on the HiperLAN/2 system, an European standard for WLAN [ETS00]. The modulation specified for the physical layer of this system is a 64-carriers OFDM, although only 52 subcarriers are active, where 4 of them are pilot carriers used for synchronization, channel estimation, etc. The number of samples in the CP is $D = 16$, with a sampling frequency equal to $f_s = 20$ MHz, leading to $T = 3.2$ μs and $T_s = 4$ μs. In the simulations, BPSK subcarrier modulation is assumed and no channel coding is used; thus, the results refer to the uncoded or raw BER as defined in (3.44). The simulated MIMO channels are normalized as $\mathbb{E}\big[\sum_{n=0}^{L-1} |h_{q,p}(n)|^2\big] = 1, \ \forall p, q$. The channel power delay profiles are those standardized by the ETSI in [ETS98].

In the first simulations, some examples are given concerning the distribution of the available transmit power among the subcarriers and the resulting SNIR. In Figure 3.7, the maximum eigenvalues $\lambda_{\max}(k)$ are shown for a single realization of a 3+4 MIMO channel with the power delay profile specified for the model A in [ETS98], corresponding to a rms delay spread of 50 ns. The angular spread at the transmitter and the receiver is 30^{o} and 15^{o}, respectively. The arrays are linear with a half wavelength ($\lambda/2$) interelement separation. At the receiver, there are no interferences, but only AWGN. Figure 3.7 shows that the global MIMO channel transfer function has two fading bands around the 8th and 55th subcarriers.

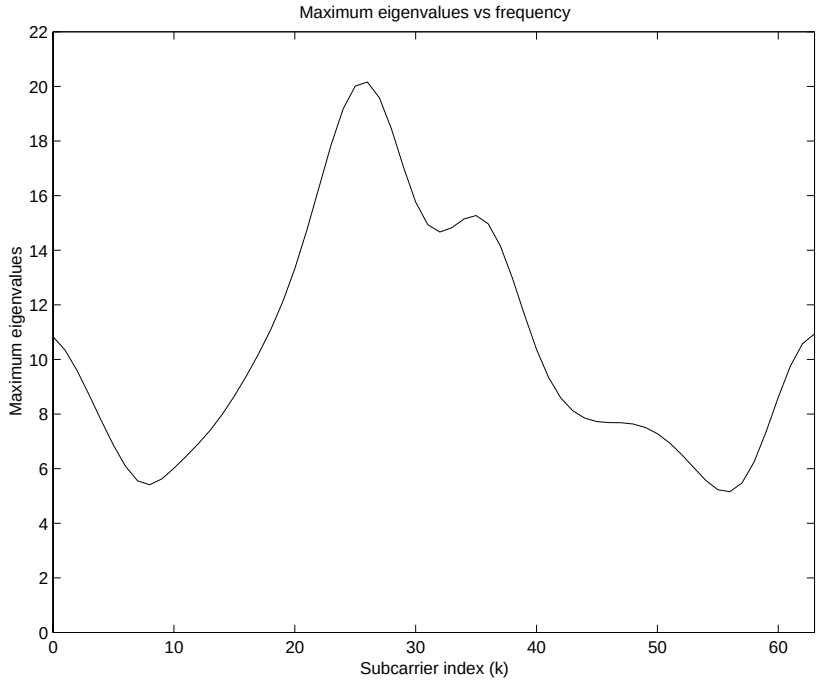
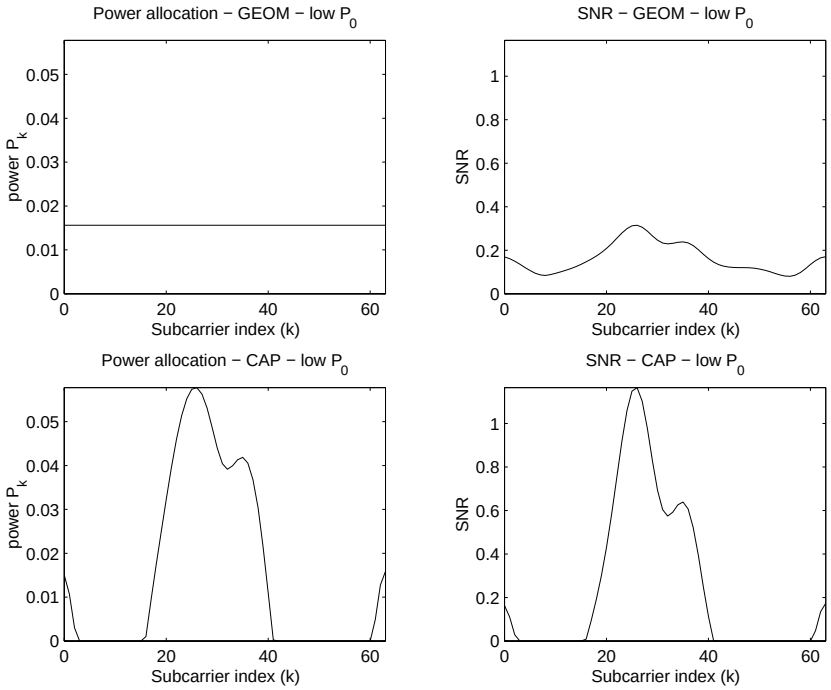


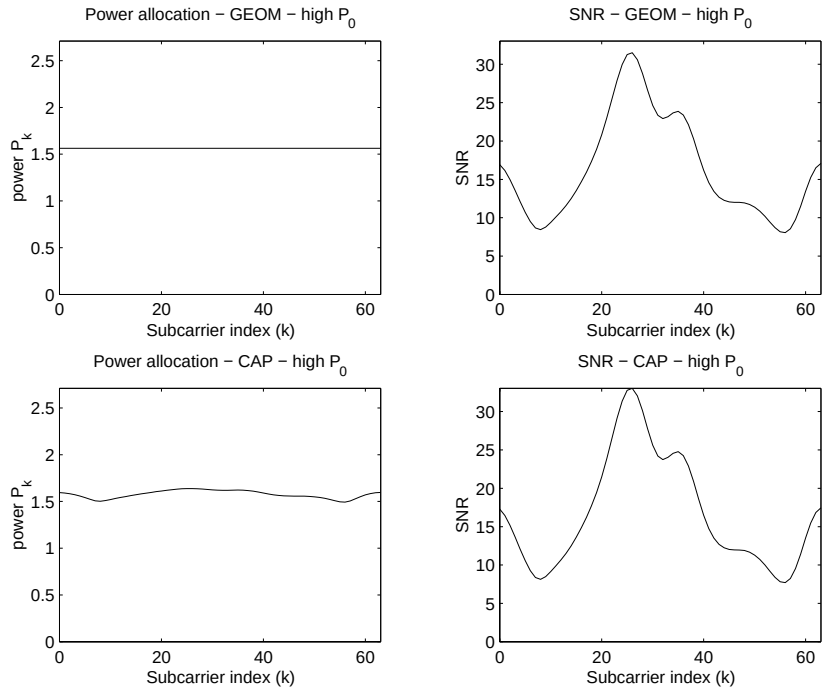
Figure 3.7: Maximum eigenvalue of $\mathbf{H}_k^H \mathbf{R}_n^{-1}(k)\mathbf{H}_k$ vs. the subcarrier index k for a single realization of a 3+4 MIMO channel A.

Figures 3.8, 3.9, and 3.10 show the power allocated to each subcarrier and the resulting SNR for the techniques GEOM-CAP, HARM-MMSE, and MAXMIN-MEP, respectively. These results are obtained considering two situations: a low transmit power ($P_0 = 1$ W) and a high transmit power ($P_0 = 100$ W). In the case of a low transmit power, it is shown that the CAP, MMSE, and MEP criteria (see Figures 3.8(a), 3.9(a), and 3.10(a)) decide not to transmit in those bands in which the channel presents fading, while allocates more power to the bands with a higher frequency response. As expected, the GEOM strategy distributes the power uniformly among the subcarriers and, therefore, the final SNR function is proportional to the maximum eigenvalues $\lambda_{\max}(k)$ (Figure 3.8). Finally, the HARM and MAXMIN techniques provide more power to those frequencies in a deeper fading (Figures 3.9 and 3.10), differing from CAP, MMSE, and MEP. The MAXMIN presents an equalized behavior for the SNR under the criterion of maximizing the minimum SNR over the subcarriers.

In the case of a high transmit power ($P_0 = 100$ W), the behavior of the GEOM technique is the same as in the case of a low transmit power. The most important difference when increasing P_0 refers to the behavior of the MMSE and MEP algorithms. In this case, more power is allocated to the subcarriers in which the MIMO channel frequency response has low values, which is the opposed behavior to the case of a low transmit power. Concerning the CAP strategy, it tends to distribute the power in a uniform way, as the GEOM technique (see Figure 3.8(b)), and as

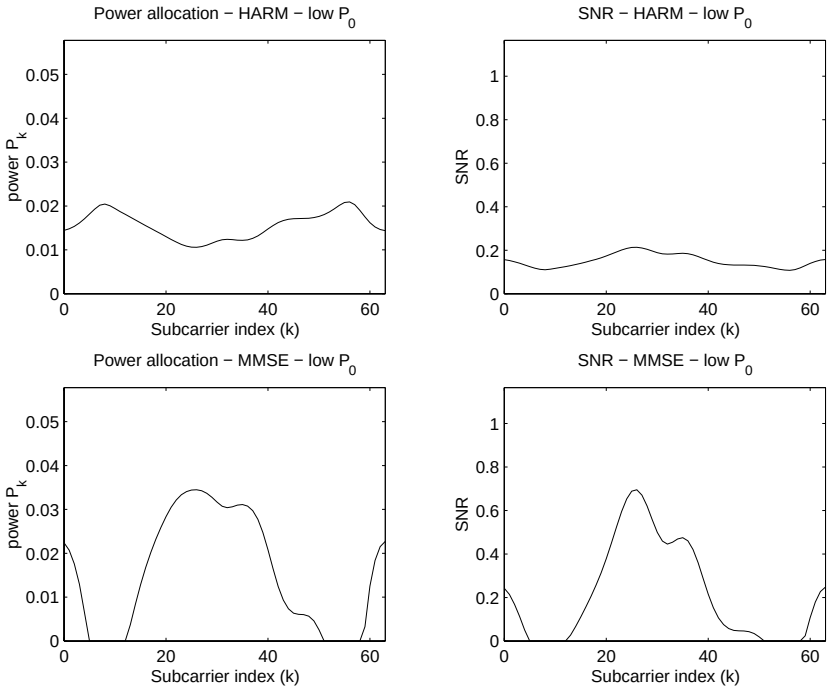


(a)

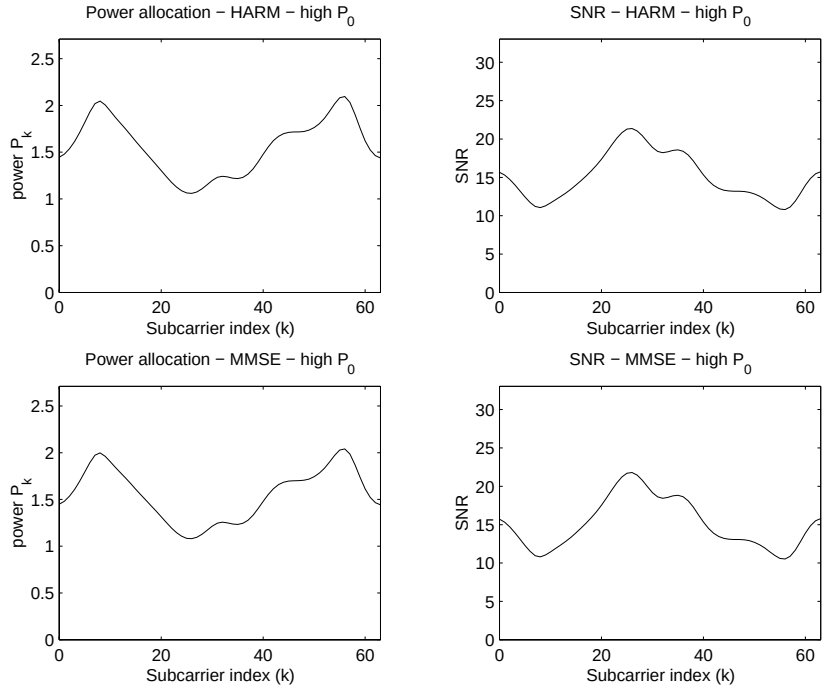


(b)

Figure 3.8: Power distribution among the subcarriers and resulting SNR for the GEOM and CAP techniques. (a) Low transmit power ($P_0 = 1$ W). (b) High transmit power ($P_0 = 100$ W).

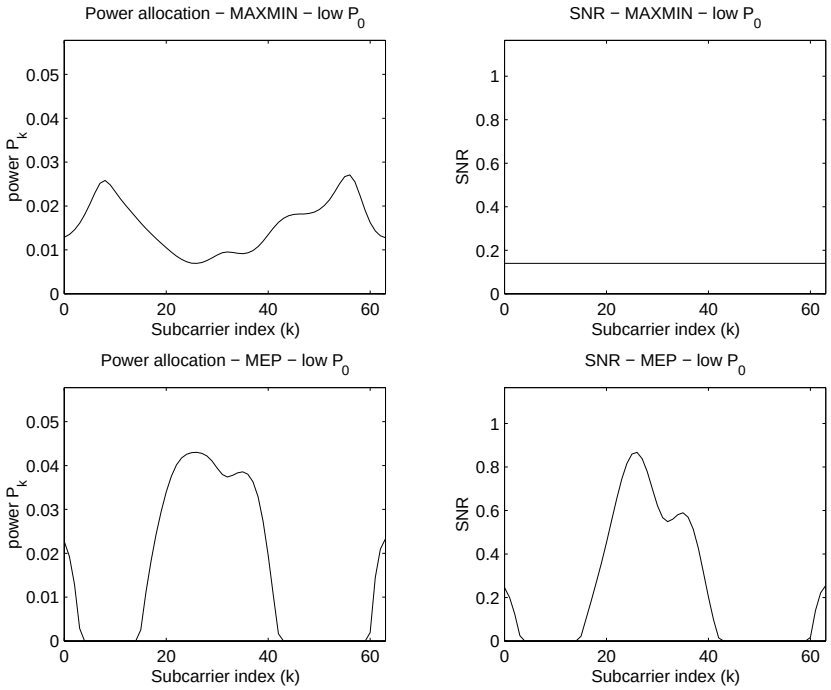


(a)

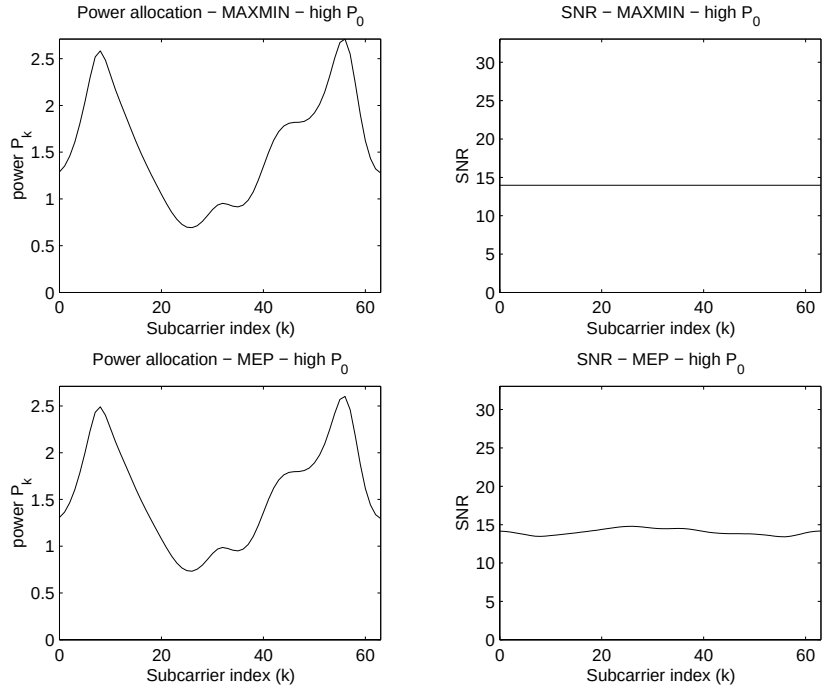


(b)

Figure 3.9: Power distribution among the subcarriers and resulting SNR for the HARM and MMSE techniques. (a) Low transmit power ($P_0 = 1$ W). (b) High transmit power ($P_0 = 100$ W).



(a)



(b)

Figure 3.10: Power distribution among the subcarriers and resulting SNR for the MAXMIN and MEP techniques. (a) Low transmit power ($P_0 = 1$ W). (b) High transmit power ($P_0 = 100$ W).

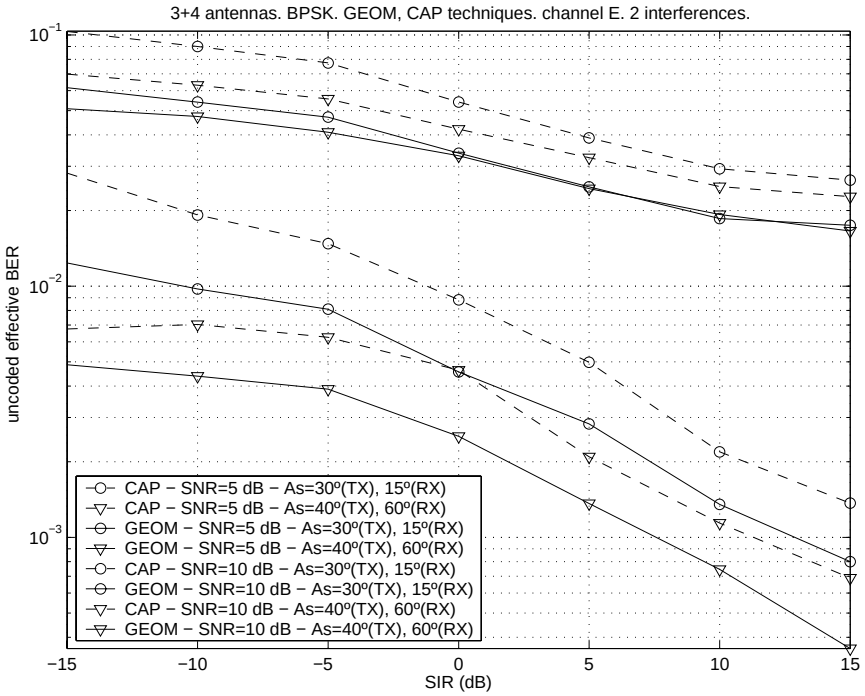


Figure 3.11: Comparison of the performance of the GEOM and CAP power allocation techniques in terms of the uncoded effective BER vs. the SIR for two SNR conditions: 5 and 10 dB. In the simulated scenario, 3 transmit and 4 receive antennas are available. The channel power delay profile is that corresponding to model E. Two different angular spread conditions have been considered: 30º at the transmitter and 15º at the receiver, and 40º at the transmitter and 60º at the receiver. There are 2 equal-level interferences from other OFDM transmitters.

proved in §3.3.1. It is also seen that the pairs of techniques HARM-MMSE and MAXMIN-MEP distribute the power in a very similar way when P_0 is high (see Figures 3.9(b) and 3.10(b)), as deduced theoretically in §3.3.2 and §3.3.3, respectively.

In the following, some figures are given showing the performance of the techniques in terms of the uncoded BER. In these figures, the SNR and the SIR are defined as follows: SNR $= P_0 n_T n_R/\sigma_n^2$ and SIR $= P_0 n_T/P_I$, where σ_n^2 and P_I are the mean power of the AWGN and the interferences at each receive antenna, respectively. The interferences correspond to other OFDM transmissions with the same modulation format as the desired signal.

Figure 3.11 shows the comparison between the GEOM and CAP power allocation strategies in terms of the uncoded BER in a 3+4 antennas scenario assuming the power delay profile for channel E (rms delay spread equal to 250 ns). The performance has been represented vs. the SIR assuming 2 equal-level interferences and two SNR conditions: 5 and 10 dB. Two different angular spreads have been simulated, 30º at the transmitter and 15º at the receiver, and 40º at the transmitter and 60º at the receiver. From the figure it is concluded that GEOM performs better than CAP, even taking into account that GEOM distributes the power in a uniform

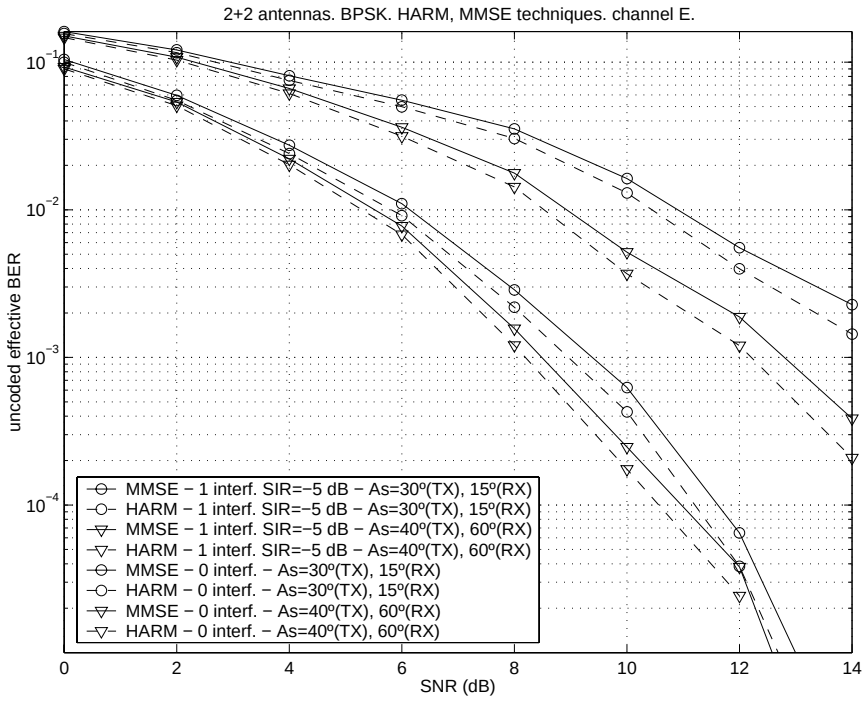


Figure 3.12: Comparison of the performance of the HARM and MMSE power allocation techniques in terms of the uncoded effective BER vs. the SNR for two interference conditions: 1 interference from another OFDM transmitter with a SIR $= -5$ dB and no interference. In the simulated scenario, 2 transmit and 2 receive antennas are available. The channel power delay profile is that corresponding to model E. Two different angular spread conditions have been considered: 30^o at the transmitter and 15^o at the receiver, and 40^o at the transmitter and 60^o at the receiver.

way. An additional conclusion is that, in this case, an increase of the angular spread implies an improvement of the BER, due to the increase of the space diversity.

The performance of the HARM and MMSE techniques is shown in Figure 3.12, where a 2+2 antennas scenario is considered. The BER is represented vs. the SNR assuming two interferences conditions: one interference with a SIR equal to -5 dB, and no interference. The simulated angular spreads and the power delay profile are the same as in the previous figure. In the figure, it is seen that the HARM technique performs better than the MMSE strategy in all the simulated SNR range. Besides, and as commented in §3.3.2, the computational load of HARM is lower than in the case of MMSE, concluding that HARM is more suitable for OFDM than MMSE. Regarding the angular spread, the same conclusion as in the previous case is obtained, i.e., and increase of the angular spread implies an improvement of the BER.

The same scenario and channel conditions have been assumed in the simulations corresponding to Figure 3.13, where the MAXMIN and the MEP power allocation strategies are compared in terms of the uncoded BER. As expected, the performance of the MEP strategy, which minimizes the Chernoff upper-bound on the BER, is better than the performance corresponding to

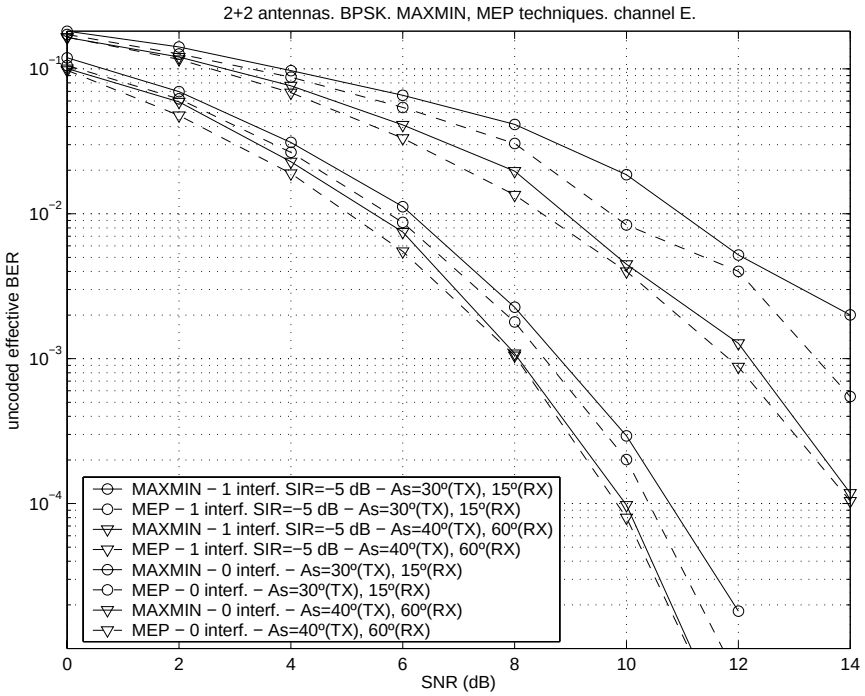


Figure 3.13: Comparison of the performance of the MAXMIN and MEP power allocation techniques in terms of the uncoded effective BER vs. the SNR for two interference conditions: 1 interference from another OFDM transmitter with a SIR $= -5$ dB and no interference. In the simulated scenario, 2 transmit and 2 receive antennas are available. The channel power delay profile is that corresponding to model E. Two different angular spread conditions have been considered: 30^o at the transmitter and 15^o at the receiver, and 40^o at the transmitter and 60^o at the receiver.

the MAXMIN strategy. Note, however, that the maximum difference between the performance of both techniques is around 1 dB in terms of SNR, i.e., the performance degradation of the MAXMIN strategy when compared to MEP is not high. Furthermore, the MEP technique has the highest computational complexity and, therefore, in some cases it may be more suitable to use the MAXMIN strategy. As in the previous simulations, the BER improves if the angular spread increases.

In Figures 3.14 and 3.15, the GEOM, MMSE, HARM, and MAXMIN approaches are compared. In the considered scenario, 2+2 antennas are available and the angular spread is equal to 30^o at the transmitter and 15^o at the receiver. Two interference conditions have been simulated: one interference with a SIR equal to -5 dB, and no interference. In Figure 3.14, the power delay profile corresponding to channel A with a delay spread equal to 50 ns has been assumed, whereas in Figure 3.15, the channel model is C with a delay spread equal to 150 ns. From both figures similar conclusions can be obtained. In an acceptable range, the worst technique is GEOM, i.e., the technique performing a uniform power allocation, whereas the MAXMIN and HARM techniques have the best performance. Note also that, as the transmit power or SNR increases,

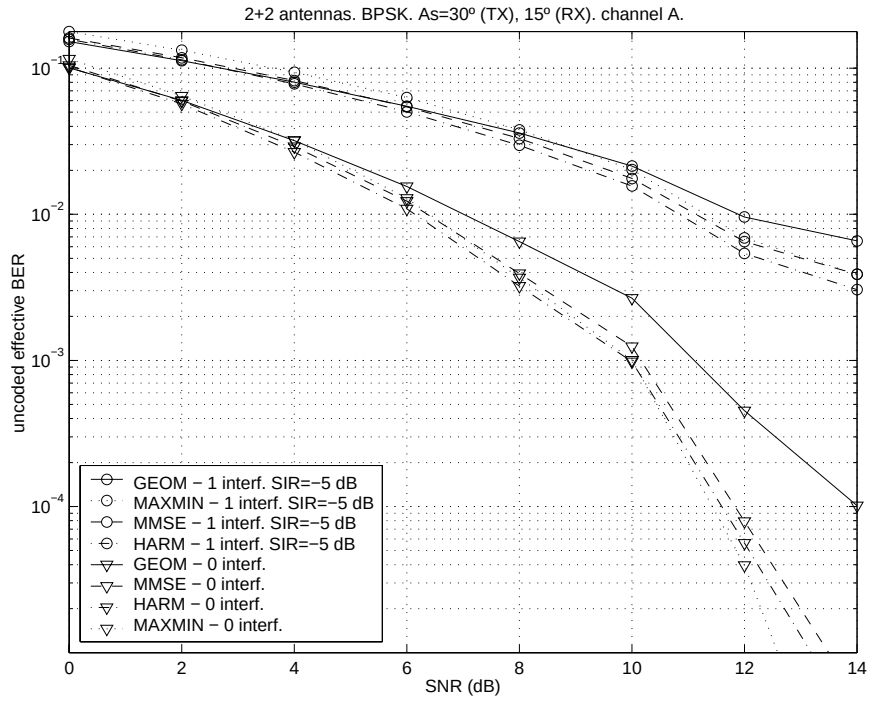


Figure 3.14: Comparison of the performance of the GEOM, MMSE, HARM, and MAXMIN power allocation techniques in terms of the uncoded effective BER vs. the SNR for two interference conditions: 1 interference from another OFDM transmitter with a SIR $= -5$ dB and no interference. In the simulated scenario, 2 transmit and 2 receive antennas are available. The channel power delay profile is that corresponding to model A. The angular spread is 30^o at the transmitter and 15^o at the receiver.

MAXMIN tends to perform better than HARM. In case that there is no interference and the noise is Gaussian, this result is justified by the fact that the MAXMIN strategy minimizes asymptotically the Chernoff upper-bound on the BER (see §3.3.3). Note, however, that even in the case that there are non-Gaussian interferences, MAXMIN performs better than HARM for a high enough SNR and SIR. Note also that for low SNR, i.e., for low values of the transmit power P_0, the MAXMIN strategy may have a worse BER than GEOM and MMSE. This is due to the fact that MAXMIN allocates more power to the most degraded subcarriers and, therefore, in case that the available power is low, the performance of the OFDM transmission can be severely degraded if a deep fading occurs in some carrier, so that most of the power is allocated to that frequency band. When comparing the performance of the techniques obtained from Figures 3.14 and 3.15, it is seen that increasing the delay spread also improves the performance due to the increase of the time diversity.

Finally, as far as the computational load is concerned, GEOM is the simplest technique, followed by MAXMIN, HARM, CAP, MMSE, and MEP.

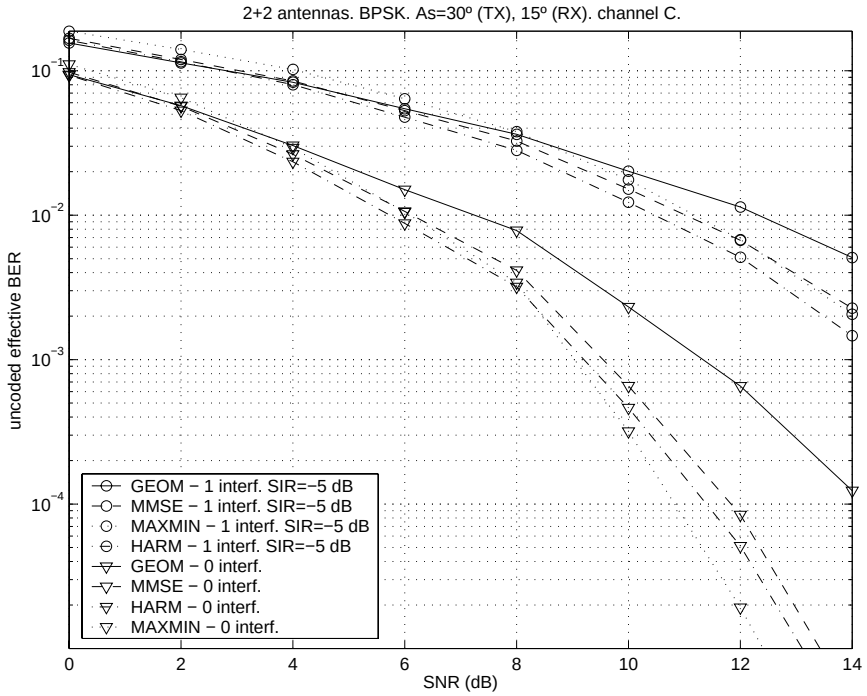


Figure 3.15: Comparison of the performance of the GEOM, MMSE, HARM, and MAXMIN power allocation techniques in terms of the uncoded effective BER vs. the SNR for two interference conditions: 1 interference from another OFDM transmitter with a SIR $= -5$ dB and no interference. In the simulated scenario, 2 transmit and 2 receive antennas are available. The channel power delay profile is that corresponding to model C. The angular spread is 30^{o} at the transmitter and 15^{o} at the receiver.

3.5 Chapter Summary and Conclusions

In this chapter, a design strategy has been presented for a single-user communication system where the multicarrier OFDM modulation is combined with the use of multiple antennas at both the transmitter and the receiver, i.e., with a MIMO channel. The design strategy exploits the full channel knowledge at both sides of the communication system, where this knowledge encompasses not only the channel response, but also the noise plus interferences correlation matrices. Based on this, a joint beamforming solution per carrier has been proposed, where the final objective is the maximization of the SNIR.

After calculating the optimum beamvectors at the transmitter and the receiver, a suitable power allocation among the subcarriers has to be identified. Three different strategies with a low computational load have been proposed: GEOM, HARM, and MAXMIN, and their asymptotic behaviors have been compared to other classical solutions: CAP, MMSE, and MEP.

In the simulations, it has been shown that MAXMIN and HARM have a good performance, even better than the classical solutions requiring more computational effort. Concretely, the MAXMIN approach, which consists in the maximization of the worst SNIR over the subcarriers

leading to a solution where the same SNIR is attained at all the frequencies, has the best trade-off between complexity and transmission quality, since the performance of this technique is improved only by MEP, which minimizes the Chernoff upper-bound on the BER, although MEP has a much higher computational load. From the simulations it is also concluded that increasing the angular and delay spreads improves the performance due to the increase of the space and time diversities.

3.A Appendix: Power Allocation Maximizing the Geometric Mean of the SNIR

The geometric mean of the SNIR is given by

$$\mathcal{G}\left(\{\mathrm{SNIR}_i\}\right) \triangleq \prod_{i=0}^{N-1} \left(\lambda_{\max}(i) P_i\right)^{1/N} = \prod_{i=0}^{N-1} \lambda_{\max}^{1/N}(i) \prod_{i=0}^{N-1} P_i^{1/N} = f\left(\{P_i\}\right) \prod_{i=0}^{N-1} \lambda_{\max}^{1/N}(i), \quad (3.50)$$

where $f\left(\{x_i\}\right) \triangleq \left(\prod_{i=0}^{N-1} x_i\right)^{1/N}$ and it has been assumed that all the maximum eigenvalues are greater than 0, since, otherwise, the geometric mean would be equal to 0 for any power distribution. The function f is concave, since its Hessian matrix is negative semidefinite and, therefore, the geometric mean is also a concave function of the power allocation variables $\{P_i\}$. The proof of the concavity of f is given below.

The expression of the first order derivatives of f is

$$\frac{\partial f\left(\{x_i\}\right)}{\partial x_k} = \frac{1}{N} \frac{\left(\prod_{i=0}^{N-1} x_i\right)^{1/N}}{x_k}, \quad (3.51)$$

whereas the second order derivatives of f are given by

$$\frac{\partial^2 f\left(\{x_i\}\right)}{\partial x_k^2} = -\frac{N-1}{N^2} \frac{\left(\prod_{i=0}^{N-1} x_i\right)^{1/N}}{x_k^2}, \qquad \frac{\partial^2 f\left(\{x_i\}\right)}{\partial x_l \partial x_k} = \frac{1}{N^2} \frac{\left(\prod_{i=0}^{N-1} x_i\right)^{1/N}}{x_l x_k}, \quad k \neq l. \quad (3.52)$$

Using (3.52), the Hessian matrix of f can be written as

$$\nabla^2 f\left(\{x_i\}\right) = -\frac{\left(\prod_{i=0}^{N-1} x_i\right)^{1/N}}{N^2} \left(N \,\mathrm{diag}\left(\{1/x_i^2\}\right) - \mathbf{q}\mathbf{q}^T\right), \quad (3.53)$$

where $\mathbf{q} = \left[1/x_0 \cdots 1/x_{N-1}\right]^T \in I\!R^{N\times 1}$. The Hessian matrix is negative semidefinite $\nabla^2 f \preceq \mathbf{0}$, as proved below:

$$\mathbf{v}^T \left(\nabla^2 f\left(\{x_i\}\right)\right) \mathbf{v} = -\frac{\left(\prod_{i=0}^{N-1} x_i\right)^{1/N}}{N^2} \left(N \sum_{i=0}^{N-1} \frac{v_i^2}{x_i^2} - \left(\sum_{i=0}^{N-1} \frac{v_i}{x_i}\right)^2\right) < 0, \qquad \forall \mathbf{v} \in I\!R^{N\times 1}, \quad (3.54)$$

where $\mathbf{v} = [v_0 \cdots v_{N-1}]$. The expression above results from the Cauchy-Schwarz inequality $((\mathbf{a}^T\mathbf{a})(\mathbf{b}^T\mathbf{b}) \geq (\mathbf{a}^T\mathbf{b})^2)$, applied to the vectors $\mathbf{a} = [1 \cdots 1]^T \in I\!R^{N\times 1}$ and $\mathbf{b} = \left[v_0/x_0 \cdots v_{N-1}/x_{N-1}\right]^T \in I\!R^{N\times 1}$.

Taking into account that the function f is concave, the maximization of the geometric mean of the SNIR can be written as the following convex minimization problem (note that the scalar factor $\prod_{i=0}^{N-1} \lambda_{\max}^{1/N}(i)$ in (3.50) has not been included since it does not modify the value of the

optimum power allocation variables $\{P_i\}$, and that the maximization has been written as the minimization of the same function changing its sign):

$$\begin{array}{ll} \underset{\{P_i\}}{\text{minimize}} & -\prod_{i=0}^{N-1} P_i^{1/N} \\ \text{subject to} & -P_i \leq 0, \qquad i = 0, \ldots, N-1, \\ & \sum_{i=0}^{N-1} P_i - P_0 = 0. \end{array} \tag{3.55}$$

The Lagrangian function is given by

$$L(\{P_i\}; \boldsymbol{\lambda}, \nu) \triangleq -\prod_{i=0}^{N-1} P_i^{1/N} - \sum_{i=0}^{N-1} \lambda_i P_i + \nu \left(\sum_{i=0}^{N-1} P_i - P_0 \right), \tag{3.56}$$

and the corresponding KKT conditions are:

$$P_k^\star \geq 0, \qquad k = 0, \ldots, N-1, \tag{3.57}$$

$$\lambda_k^\star \geq 0, \qquad k = 0, \ldots, N-1, \tag{3.58}$$

$$\sum_{k=0}^{N-1} P_k^\star = P_0, \tag{3.59}$$

$$\lambda_k^\star P_k^\star = 0, \qquad k = 0, \ldots, N-1, \tag{3.60}$$

$$\frac{\partial L(\{P_i^\star\}; \boldsymbol{\lambda}^\star, \nu^\star)}{\partial P_k} = -\frac{1}{N} \frac{\left(\prod_{i=0}^{N-1} P_i^\star \right)^{1/N}}{P_k^\star} - \lambda_k^\star + \nu^\star = 0, \qquad k = 0, \ldots, N-1. \tag{3.61}$$

All the optimum powers allocated to the subcarriers should be greater than 0 since, otherwise, the geometric mean would be equal to 0. Consequently, $\lambda_k^\star = 0,\ k = 0, \ldots, N-1$, and, therefore, from (3.61) it is concluded that all the powers $\{P_k\}$ must be equal. Using this result jointly with (3.59), the following solution is obtained:

$$P_k^\star = \frac{P_0}{N}, \qquad k = 0, \ldots, N-1. \tag{3.62}$$

3.B Appendix: Power Allocation Maximizing the Harmonic Mean of the SNIR

The harmonic mean of the SNIR is given by

$$\mathcal{H}\left(\{\mathrm{SNIR}_i\}\right) \triangleq \frac{N}{\sum_{i=0}^{N-1} \frac{1}{\lambda_{\max}(i) P_i}} = N f\left(\{\lambda_{\max}(i) P_i\}\right), \tag{3.63}$$

where $f\left(\{x_i\}\right) \triangleq 1/\left(\sum_{i=0}^{N-1} 1/x_i\right)$ and it has been assumed that all the maximum eigenvalues are greater than 0, since, otherwise, the harmonic mean would be equal to 0 for any power distribution. The harmonic mean results from the composition of the function f, which is concave (its

Hessian matrix is negative semidefinite), with an affine mapping ($x_i = \lambda_{\max}(i)P_i$); consequently, the harmonic mean is also concave on the variables $\{P_i\}$ [Boy04]. The proof of the concavity of f is shown in the following.

The expression of the first order derivatives of f is

$$\frac{\partial f\left(\{x_i\}\right)}{\partial x_k} = \frac{1}{x_k^2}\frac{1}{\left(\sum_{i=0}^{N-1}\frac{1}{x_i}\right)^2}, \tag{3.64}$$

whereas the second order derivatives of f are given by

$$\frac{\partial^2 f\left(\{x_i\}\right)}{\partial x_k^2} = \left(-\frac{2}{x_k^3}\left(\sum_{i=0}^{N-1}\frac{1}{x_i}\right)+\frac{2}{x_k^4}\right)\frac{1}{\left(\sum_{i=0}^{N-1}\frac{1}{x_i}\right)^3} \tag{3.65}$$

$$\frac{\partial^2 f\left(\{x_i\}\right)}{\partial x_l \partial x_k} = \frac{2}{x_k^2 x_l^2}\frac{1}{\left(\sum_{i=0}^{N-1}\frac{1}{x_i}\right)^3}, \qquad k \neq l, \tag{3.66}$$

and, therefore, the Hessian matrix of f can be written as

$$\nabla^2 f\left(\{x_i\}\right) = -\frac{2}{\left(\sum_{i=0}^{N-1}\frac{1}{x_i}\right)^3}\left(\left(\sum_{i=0}^{N-1}\frac{1}{x_i}\right)\mathrm{diag}\left(\{1/x_i^3\}\right)-\mathbf{q}\mathbf{q}^T\right), \tag{3.67}$$

where $\mathbf{q} = \left[1/x_0^2 \cdots 1/x_{N-1}^2\right]^T \in I\!R^{N\times 1}$. The Hessian matrix is negative semidefinite $\nabla^2 f \preceq \mathbf{0}$, as proved below:

$$\mathbf{v}^T\left(\nabla^2 f\left(\{x_i\}\right)\right)\mathbf{v} = -\frac{2}{\left(\sum_{i=0}^{N-1}\frac{1}{x_i}\right)^3}\left(\left(\sum_{i=0}^{N-1}\frac{1}{x_i}\right)\left(\sum_{i=0}^{N-1}\frac{v_i^2}{x_i^3}\right)-\left(\sum_{i=0}^{N-1}\frac{v_i}{x_i^2}\right)^2\right) < 0, \quad \forall \mathbf{v} \in I\!R^{N\times 1}, \tag{3.68}$$

where $\mathbf{v} = [v_0 \cdots v_{N-1}]$. The expression above results from the Cauchy-Schwarz inequality ($(\mathbf{a}^T\mathbf{a})(\mathbf{b}^T\mathbf{b}) \geq (\mathbf{a}^T\mathbf{b})^2$), applied to the vectors $\mathbf{a} = \left[v_0/x_0^{3/2} \cdots v_{N-1}/x_{N-1}^{3/2}\right]^T$ and $\mathbf{b} = \left[1/\sqrt{x_0} \cdots 1/\sqrt{x_{N-1}}\right]^T$.

Taking into account that the function f is concave, the maximization of the harmonic mean of the SNIR can be written as the following convex minimization problem, in which the maximization of (3.63) has been written as the minimization of the same function changing its sign:

$$\begin{array}{ll} \underset{\{P_i\}}{\text{minimize}} & -\frac{N}{\sum_{i=0}^{N-1}\frac{1}{\lambda_{\max}(i)P_i}} \\ \text{subject to} & -P_i \leq 0, \qquad i = 0,\ldots,N-1, \\ & \sum_{i=0}^{N-1} P_i - P_0 = 0, \end{array} \tag{3.69}$$

whose Lagrangian function is given by

$$L(\{P_i\};\boldsymbol{\lambda},\nu) \triangleq -\frac{N}{\sum_{i=0}^{N-1}\frac{1}{\lambda_{\max}(i)P_i}} - \sum_{i=0}^{N-1}\lambda_i P_i + \nu\left(\sum_{i=0}^{N-1} P_i - P_0\right), \tag{3.70}$$

and the corresponding KKT conditions are:

$$P_k^\star \geq 0, \qquad k = 0, \ldots, N-1, \quad (3.71)$$

$$\lambda_k^\star \geq 0, \qquad k = 0, \ldots, N-1, \quad (3.72)$$

$$\sum_{k=0}^{N-1} P_k^\star = P_0, \quad (3.73)$$

$$\lambda_k^\star P_k^\star = 0, \qquad k = 0, \ldots, N-1, \quad (3.74)$$

$$\frac{\partial L(\{P_i^\star\}; \boldsymbol{\lambda}^\star, \nu^\star)}{\partial P_k} = \frac{-N \frac{1}{\lambda_{\max}(k) P_k^{\star 2}}}{\left(\sum_{i=0}^{N-1} \frac{1}{\lambda_{\max}(i) P_i^\star}\right)^2} - \lambda_k^\star + \nu^\star = 0, \qquad k = 0, \ldots, N-1. \quad (3.75)$$

All the optimum powers allocated to the subcarriers should be greater than 0 since, otherwise, the harmonic mean would be equal to 0. Consequently, $\lambda_k^\star = 0,\ k = 0, \ldots, N-1$, and, therefore, from (3.75) it is concluded that the optimum power $P_k^\star$ is proportional to $1/\sqrt{\lambda_{\max}(k)}$. Using this result jointly with (3.73), the following solution is obtained:

$$P_k^\star = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\sqrt{\lambda_{\max}(l)}}} \frac{1}{\sqrt{\lambda_{\max}(k)}}, \qquad k = 0, \ldots, N-1. \quad (3.76)$$

Chapter 4

Joint Beamforming Design in MIMO-OFDM Multi-User Communications

4.1 Introduction

In Chapter 3, a joint beamforming approach has been proposed to design a MIMO-OFDM single-user communication system. In that case, multiple antennas have been used to increase the diversity and the system performance. Note, however, that the use of multiple transmit and receive antennas provides additional potentials that can be exploited to improve other aspects of a communication system. One of the most important ones is the ability to increase the number of users that can be served simultaneously by the system. This is specially important nowadays, since the capacity of the current communication systems is becoming more limited when taking into account the increase of the demand for wireless services, not only in terms of quantity, but also of quality. In this situation, space diversity plays an important role, since thanks to the use of multiple antennas, the SDMA can be implemented, so that the signals from different users can be transmitted simultaneously in time and using the same frequency band. Thus, by means of spatial processing techniques, these signals can be separated and detected while still guaranteeing a minimum QoS for each communication.

In this chapter, a multi-user system is designed assuming that all the MT's and BS's or AP's have multiple antennas, leading to multiple MIMO channels which are exploited simultaneously in time and using the same frequency. In this kind of scenarios, and depending on the quantity and the quality of the CSI, several designs and architectures are possible. The objective of this chapter is to design an optimum SDMA system assuming a perfect CSI that can be used as a benchmark and comparative framework when designing and evaluating other suboptimum approaches for multi-user MIMO systems. According to this objective, the existence of a centralized

manager is assumed knowing all the channel responses among all the terminals in the network. Obviously, this assumption requires the channel to be slowly time variant, so that the transmitter can have an accurate channel estimate through a feedback channel, for example. Currently, there exist several standards in which this assumption is valid. Among them, some examples can be mentioned, such as the European WLAN HiperLAN/2 [ETS00] and the IEEE 802.11a system [IEE99]. As commented in §3.1, these WLAN's use the OFDM modulation for the physical layer (see §3.2.1) and, therefore, the use of the OFDM by all the terminals is assumed in this chapter.

The design of the multi-user system will be based on a joint beamforming approach, as also in Chapter 3, i.e., all the transmitters and the receivers exploit a beamforming architecture per carrier (see Figure 3.4). Under this consideration, the receiver is based on a bank of single-user detectors and a joint design of all the transmit beamvectors is carried out by the centralized manager, trying to minimize the total transmit power. This will be done subject to several QoS constraints formulated in terms of a maximum BER for each communications and, possibly, constraints regarding the maximum transmit power for some MT's. This optimization problem is very difficult to solve, since the constraint set over which the optimization has to be carried out is not convex. As a solution to this problem, the application of the SA technique will be proposed, a very powerful heuristic optimization tool able to find the global optimum solution to an optimization problem, even when the problem is not convex. This is the main difference when compared with other proposals and works found in the literature, in which the design is based on GS or AM techniques that may find local suboptimum solutions since they are not able to handle non-convex problems. Besides, the notation and networks configuration proposed in this chapter is quite general and encompasses as particular cases, among other configurations, the uplink and downlink transmissions in cellular networks.

There are some papers in the literature considering similar joint beamforming problems. In [Lok00], an uplink MC-CDMA system is presented with one transmit antenna and several antennas at the receiver. The problem consists in the design of the optimum receiver and the transmit frequency signatures for each user. According to this, the obtained notation and the mathematical optimization problem is shown to be equivalent to the one presented in this chapter. There, the QoS constraints are formulated in terms of a minimum SNIR for each user instead of a maximum BER, and no constraints are applied regarding the maximum individual transmit powers. The optimization problem is solved by using a GS technique based on the Lagrange multipliers method and the penalty functions.

In [Won01], a multi-user MIMO-OFDM system based on joint beamforming is also considered. There, the optimization of the transmit beamvectors is based on the application of the AM technique, i.e., when designing the beamvector for one user, all the other beamvectors are assumed to be fixed. Once the design is finished, the optimization of the beamvector for another user is performed. This is applied successively until convergence is attained, although the global

optimum is not guaranteed to be found nor are the QoS constraints in terms of a minimum SNIR guaranteed to be fulfilled.

The case of an uplink flat fading MIMO multi-user channel is analyzed in [Cha02], where both the MT's and the BS's are assumed to have multiple antennas. Two different optimization problems are considered. In the first one, the minimization of the total transmit power is addressed, forcing the SNIR for each user to be higher than a prefixed value. In the second problem, the objective is to maximize the minimum SNIR subject to a total transmit power constraint. In both cases, no individual transmit power constraint is applied. In that paper, several iterative algorithms are proposed to design the beamvectors, although it is shown that those techniques may find a local suboptimum design instead of the global optimum one due to the non-convex behavior of the optimization problem.

There are many other papers and works that analyze different multi-user systems considering the use of multiple antennas. In [Rhe04], an uplink scenario with one BS and several MT's is studied, all of them with multiple antennas. There, the beamforming solution is shown to be optimum in the sense that it achieves the sum capacity for a high number of users, although no QoS can be guaranteed for each user. The same scenario is also considered in [Ser04]. In that paper, the objective is the minimization of the global MSE subject to a transmit power constraint for each MT. The iterative technique proposed to find the design is based on the application of the AM algorithm, which may converge to local suboptimum solutions. A multi-user downlink scenario with one multi-antenna BS and several single-antenna MT's is analyzed in [Boc02]. There, the global optimum design minimizing the total transmit power subject to minimum SNIR constraints is presented based on the duality between the uplink and downlink scenarios and, furthermore, the conditions for the existence of a feasible solution subject to a total transmit power constraint are deduced. The same problem is analyzed in [Vis99], where the scenario is afterwards generalized to the case of several multi-antenna BS's and multi-antenna MT's. The proposed AM iterative algorithm is shown to converge, but not always to the global optimum solution, once again due to the non-convexity of the problem. Finally, in [Ben02], the same scenario with several multi-antenna BS's and MT's is considered. An iterative AM technique for the design of the beamformers is presented to minimize the total transmit power subject to QoS constraints in terms of a minimum SNIR for each user, although it is shown that it may converge to a local suboptimum solution. In all these papers, the channel is assumed to be frequency flat, although in the design presented in this chapter, the scenario has been generalized assuming frequency selective channels and the use of the OFDM modulation.

This chapter is structured as follows. Section 4.2 summarizes the SA algorithm and its analogies with the equivalent physical process. The problem formulation corresponding to the multi-user system design is presented in Section 4.3, whereas the application of the SA algorithm to solve this problem is described in Section 4.4. Other suboptimum techniques based on GS

and AM approaches are proposed in Section 4.5, whereas the comparison of the performances provided by all the solutions are shown in Section 4.6 through several simulation results. Finally, some conclusions and a summary of the chapter are given in Section 4.7.

The publication results corresponding to this chapter can be found in [PI04e, PI02d, PI02c].

4.2 Preliminaries: The Simulated Annealing Algorithm

In this section, some preliminaries are given on the SA algorithm [Aar89, Laa87]. The SA is a powerful randomized and iterative optimization tool able to find the optimum solution to an optimization problem, even when it is not convex. This is the most important difference when compared to other techniques, such as GS and AM algorithms, that may find a local suboptimum design instead of the global optimum one, since they are not able to handle non-convex problems.

The SA algorithm was originally proposed in the early 1980's for the application to large-scale combinatorial problems which were NP-hard or NP-complete, i.e., combinatorial problems that could not be solved in an optimum way in a polynomial time with a finite computational effort. The most paradigmatic and, probably, the most studied example is the *traveling salesman problem*, which consists in finding the optimum order to visit a set of cities so that the total distance of the route is minimized.

In these cases, when the computational effort is limited, as always in practice, the only solution is to use approximation algorithms or heuristics, for which there is usually no guarantee that the obtained solution is optimum, although a polynomial bound on the computation time can be given. These algorithms can define a tradeoff between the quality of the obtained solution and the computational effort.

4.2.1 Similarities of SA with the Annealing Process in Physics

The SA is an approximation algorithm based on randomization and iterative techniques, as will be explained later. It has analogies with the annealing of solids in condensed matter physics, a process in which a solid in a heat bath is heated up by increasing the temperature to a maximum value at which all the particles of the solid randomly arrange themselves in the liquid phase, followed by cooling through slowly lowering the temperature of the heat bath. In this way, all the particles arrange themselves in the lowest energy ground state of a corresponding lattice, provided the maximum temperature is sufficiently high and the cooling is carried out sufficiently slowly. This lowest energy corresponds to the perfect state, in which the particles have no movement and the energy of the links among them is minimized. One of the key points to be successful and attain the optimum minimum energy state is that the lowering of the temperature has to be slow enough, so that the particles have "enough time" to configure their

links in the optimum way minimizing the energy and avoiding a suboptimum state.

In the annealing process, at each temperature value T, the solid is allowed to reach the *thermal equilibrium*, characterized by a probability of being in a state with energy E given by the *Boltzmann distribution*:

$$\Pr(\mathbf{E} = E) = \frac{1}{Z(T)} \exp\left(-\frac{E}{k_B T}\right), \tag{4.1}$$

where $Z(T)$ is a normalization factor that depends on the temperature T and k_B is the Boltzmann constant. The factor $\exp\left(-\frac{E}{k_B T}\right)$ is known as the Boltzmann factor. As the temperature T decreases, the Boltzmann distribution concentrates on the states with the lowest energy and, finally, when the temperature approaches 0, only the minimum energy states have a non-zero probability. This is not true if the system is cooled too fast, since then, the solid is not allowed to reach the thermal equilibrium at each temperature value and, therefore, defects can be "frozen" into the solid and metastable amorphous structures can be reached rather than the lowest energy crystalline lattice structure. The opposite case is known as *quenching*, in which the temperature is instantaneously lowered, resulting also in a metastable state.

In the case of the SA algorithm, the particles state in the annealing process in physics, i.e., the links configuration, corresponds to a possible solution or value of the optimization variables in the optimization problem, whereas the energy of the state is equivalent to the value of a cost function f for the corresponding possible solution, whose minimization is the objective of the optimization problem. In the physical process, the state changes continuously, searching a new configuration with a lower energy, although in some cases, a state with a higher energy may also be adopted. This happens more frequently, i.e., with a higher probability, when the temperature of the heat bath is high, so that the movement of the particles is also high. In the case of the SA algorithm, the iterative technique is continuously searching and checking over the whole space of possible solutions, i.e., over the space of all the possible values of the optimization variables. As in the annealing process, in SA, the ultimate objective is to find the solution with the lowest associated value of the cost function. In some intermediate iterations, however, solutions with a higher value of the cost function may also be accepted. The role of the temperature in the physical annealing process is adopted by a control parameter in SA, such that the acceptance ratio of worse solutions, i.e., solutions with a higher cost function, increases with the control parameter. The mechanism of adopting solutions with a higher value of the cost function is called *hill-climbing*, and is extremely important to be able to escape from local minimums, i.e., local suboptimum solutions. As the value of the control parameter is decreased, i.e., the system is cooled, the acceptance ratio of worse solutions also decreases, so that asymptotically only solutions with a lower value of the cost function are accepted and a minimum of the cost function in the solution space is reached. In case that the decrease of the control parameter is slow enough, the found minimum can be guaranteed to be a global minimum of the cost function in the solution space, i.e., a global optimum is obtained.

4.2.2 Description of the SA Algorithm for Combinatorial Problems

In this subsection, a brief description of the application of the SA algorithm to combinatorial problems is given. Let us give some definitions before explaining the details of the SA algorithm. Let $\mathcal{C}$ be the set of all the possible solutions to the combinatorial problem, and i an element of this set, i.e., a possible solution. The value of the cost function f, which should be minimized, corresponding to the solution i is represented by $f(i)$ (in the case of the traveling salesman problem, the solution i would correspond to a concrete order to visit the cities, and $f(i)$ would be the total distance of the route corresponding to that solution). The control parameter equivalent to the temperature in the annealing process in physics is represented by T, which is assumed to be a non-negative real number. For each solution i, $\mathcal{S}_i$ represents its neighborhood, i.e., the set of solutions from $\mathcal{C}$ that are "close" to i in some sense (the simplest neighborhood structure is given by $S_i = \mathcal{C},\ \forall i \in \mathcal{C}$). Obviously, given a set $\mathcal{C}$, many neighborhood structures can be defined. The neighborhood structure is said to be *suitable* if any two possible solutions in $\mathcal{C}$ can be connected by a sequence of solutions, so that each of them belongs to the neighborhood of the previous solution. More details can be obtained in [Aar89, Laa87] about these definitions, although they are not given here, since they are outside the scope of this dissertation.

The SA algorithm simulates the behavior of a solid until it attains the thermal equilibrium at each temperature T. As commented previously, SA is an iterative random algorithm consisting of a sequence of transitions between possible solutions. Each transition in the algorithm consists of two main steps: the *generation of a new proposed solution* and the *acceptance of this solution.*

Let us assume that in a concrete instant, the solution $i \in \mathcal{C}$ is the *current solution.* Given a current solution i, a new proposed solution j is generated by randomly selecting a solution from the neighborhood of the current solution $\mathcal{S}_i$ ($j \in \mathcal{S}_i$) using an uniform distribution, i.e, all the solutions in the neighborhood $\mathcal{S}_i$ have the same probability of being selected. The new proposed solution j is accepted as the current solution with a probability, whose expression is given by

$$P_{ij} = \Pr(\text{accept } j \text{ from } i) \triangleq \begin{cases} 1, & f(j) \leq f(i) \\ \exp\left(-\frac{f(j)-f(i)}{T}\right), & f(j) > f(i), \end{cases} \tag{4.2}$$

which constitutes a randomized criterion called *Metropolis criterion.* From (4.2) it is concluded that, when the new proposed solution is better that the current one, i.e., the cost function evaluated at the proposed solution is lower than at the current solution ($f(j) \leq f(i)$), then it is always accepted. In case that the proposed solution is worse ($f(j) > f(i)$), then it is accepted with a probability lower than 1. Note also that the acceptance probability increases with the temperature T, which means that when the system is cooled, the algorithm tends to accept only better solution. As commented previously, this algorithm is randomized, which means that the sequence of accepted solutions may be different when running the algorithm more than once.

In [Aar89, Laa87], it is shown under the theory of Markov chains that, if a suitable neigh-

borhood structure is defined, after a sufficiently large number of transitions at a fixed value of T, the algorithm described above will find a solution $i \in \mathcal{C}$ with a probability equal to

$$\Pr(i;T) = \frac{1}{N_0(T)} \exp\left(-\frac{f(i)}{T}\right), \tag{4.3}$$

where $N_0(T)$ is a normalization factor calculated as

$$N_0(T) = \sum_{j \in \mathcal{C}} \exp\left(-\frac{f(j)}{T}\right). \tag{4.4}$$

Note that the probability distribution shown in (4.3) is the same as the probability distribution corresponding to the thermal equilibrium in the annealing process in physics shown in (4.1), i.e., the SA algorithm emulates the annealing process, as explained previously.

The most important property of the SA algorithm is that the equilibrium probability distribution (4.3) tends to select only the optimum solutions of the problem when the temperature tends to 0, i.e., when the system is cooled. Let $\mathcal{S}_{opt}$ be the set of optimum solutions, i.e., the set of solutions at which the cost function f is minimized over the whole set of possible solutions $\mathcal{C}$. Let us calculate the limit of the probability (4.3) as the temperature T tends to 0, as follows:

$$\lim_{T\to 0} \Pr(i;T) = \lim_{T\to 0} \frac{\exp\left(-\frac{f(i)}{T}\right)}{\sum_{j\in\mathcal{C}} \exp\left(-\frac{f(j)}{T}\right)} = \lim_{T\to 0} \frac{1}{\sum_{j\in\mathcal{C}} \exp\left(\frac{f(i)-f(j)}{T}\right)}. \tag{4.5}$$

Note that, in the expression above, all the terms of the sum in the denominator are non-negative. In the case that i is not an optimum solution to the problem, then at least one of the terms in the denominator will have a strictly positive exponent. When $T \to 0$, this term will tend to infinity and, therefore, the probability will tend to 0. In case that i is an optimum solution, then all the terms in the denominator will have negative exponents, except for those terms corresponding to optimum solutions ($j \in \mathcal{S}_{opt}$), which will have an exponent equal to 0 for any value of the temperature. When $T \to 0$, all the terms with a negative exponent will tend to 0 and, therefore, the denominator will tend to $|\mathcal{S}_{opt}|$. These results can be summarized in the following expression of the limit:

$$\lim_{T\to 0} \Pr(i;T) = \begin{cases} \frac{1}{|\mathcal{S}_{opt}|}, & i \in \mathcal{S}_{opt}, \\ 0, & i \notin \mathcal{S}_{opt}, \end{cases} \tag{4.6}$$

which means that, asymptotically, an optimum solution is attained with probability 1. This result is true if the thermal equilibrium is attained at each value of temperature T, concluding that a minimum number of transitions have to be performed at each temperature. The initial value of the temperature has to be high enough such that most of the proposed solutions are accepted. Besides, the temperature has to be lowered slowly, since otherwise, the desired limit behavior is not obtained.

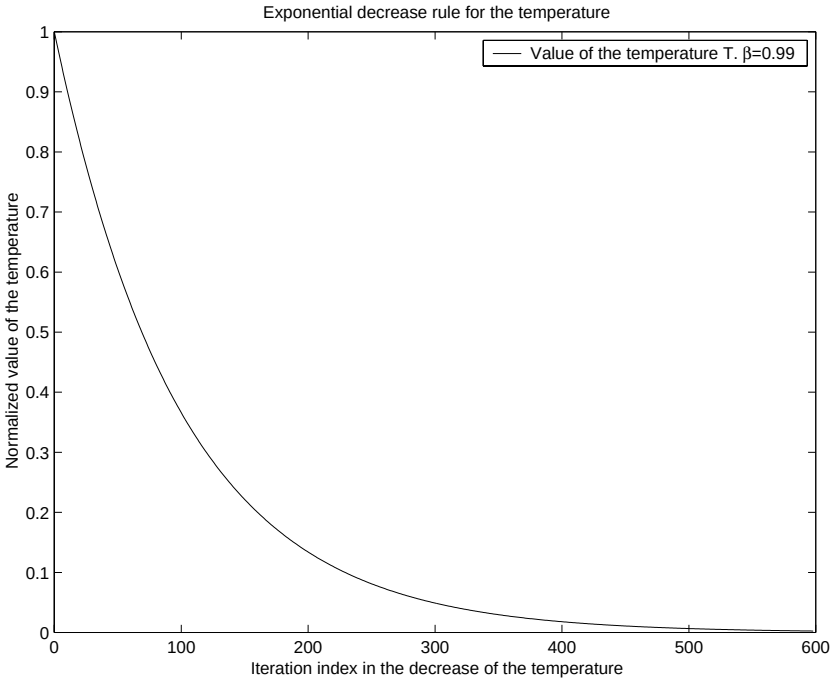


Figure 4.1: Normalized value of the temperature T vs. the iteration index in the decrease of the temperature using an exponential rule with $\beta = 0.99$.

A practical and common way to lower the value of the temperature is to use an exponential decrease rule, which is described as

$$T \leftarrow \beta T, \qquad \beta \simeq 0.99, \tag{4.7}$$

and whose graphical representation is shown in Figure 4.1. In the figure, the value 1 corresponds to the maximum value of the temperature, i.e., the representation corresponds to the evolution of the temperature with respect to the initial value.

One of the main disadvantages of the SA algorithm is that the convergence rate may be low (indeed, this is the main price to be paid in order to be able to find the optimum solution). Variants of the SA have been proposed trying to increase the convergence rate, such as in [Her00], where the modification of the SA algorithm consists in redefining the energy state, i.e., the value of the cost function, by a rescaling given by $\widetilde{f}(i) \triangleq (f(i) - f_t)^2$, where f_t is called the target cost, and whose value is also decreased with the temperature. See [Her00] for more details on this technique.

In Figure 4.3 in §4.4, a flowchart is shown corresponding to the application of the SA algorithm technique to the concrete problem presented in this chapter.

4.2.3 Extension of the SA Algorithm to Continuous Space Solutions

In the previous subsection, the application of the SA algorithm to combinatorial problems has been described. Note, however, that the SA could also be applied to the case of optimization problems where the optimization variables are continuous. In the following, an extension of the SA algorithm is proposed to find the optimum solution in a continuous set of feasible solutions.

Let $\mathbf{x}$ represent the vector of optimization variables and $f(\mathbf{x})$ be the value of the cost function to be minimized at $\mathbf{x}$. In constrained minimization problems, such as generally described in §2.3.3 (see (2.7)), the vector $\mathbf{x}$ has to belong to the constraint set $\mathcal{C}$. This constraint set can be expressed w.l.o.g. as follows:

$$\mathcal{C} \triangleq \{\mathbf{x} : f_i(\mathbf{x}) \leq 0,\ i = 1, \ldots, m,\ h_i(\mathbf{x}) = 0,\ i = 1, \ldots, p\}\,. \tag{4.8}$$

Taking this into account, a possible modified cost function $\widetilde{f}$ could be defined as:

$$\widetilde{f}(\mathbf{x}; T) \triangleq f(\mathbf{x}) + \frac{\alpha}{T} \sum_{i=1}^{m} \left(f_i(\mathbf{x})\right)^{+2} + \frac{\alpha}{T} \sum_{i=1}^{p} \left(h_i(\mathbf{x})\right)^2, \tag{4.9}$$

i.e., the "energy" of the solution $\mathbf{x}$, which is measured by the modified cost function $\widetilde{f}$, takes into account explicitly the constraints by introducing a penalty term greater than 0 if they are not fulfilled, and which is inversely proportional to the temperature (α is a proportional factor). This modified cost function is then used to accept a new proposed solution using the rule described in (4.2) (i.e., in (4.2), the modified cost function $\widetilde{f}$ is used instead of the original one f). When the temperature tends to 0, the penalty term grows without bounds if a non-feasible solution is taken. Since the SA algorithm tends to minimize the cost function, non-feasible solutions are not accepted asymptotically. In a practical implementation, as the temperature is lowered, the accepted solutions tend to fulfill the constraints in a tighter way.

The proposal and generation of new solutions can be done by applying a perturbation to the current solution. This perturbation can be drawn from a zero-mean Gaussian distribution, whose variance measures indirectly the size of the "neighborhood" of $\mathbf{x}$. Mathematically, the proposed solution $\widetilde{\mathbf{x}}$ is generated according to the following expression:

$$\widetilde{\mathbf{x}} = \mathbf{x} + \mathbf{w}, \qquad \mathbf{w} \sim \mathcal{N}\left(\mathbf{0}, \sigma_w^2 \mathbf{I}\right), \tag{4.10}$$

where a complex Gaussian vector is applied instead of a real-valued vector if the optimization variables are complex. The variance of the noise may be decreased as the iterations go on to increase the convergence rate, so that smaller perturbations are applied when an optimum minimum is approached and the temperature is lowered.

There are several references proposing different extensions and variations of the original SA algorithm to the case of continuous space solutions [Aar89]. These references analyze the

convergence properties of the proposed algorithms, showing that the accepted solutions tend to be within a region around the optimum solution, whose volume tends to 0 as the temperature is decreased. In [Van84], the extension of the SA to continuous space solutions is carried out by applying a perturbation drawn from a continuous distribution whose covariance matrix is proportional to the inverse of an estimate of the Hessian matrix of the function to be minimized. The estimate of this Hessian matrix is obtained by using the values of the cost function at the accepted solutions. Other approaches can be found in [Č84, Wil85, Kha86, Szu87], among others, where the main difference is the way the proposed solutions are generated, proposing different and simple alternatives, and the problem to which the algorithm is applied. For example, in [Szu87], the perturbation to be applied is drawn from a Cauchy distribution [Pap91], which guarantees that the optimum solution is found provided an adequate decreasing rule for the temperature is used.

4.3 Extension of the System Model to Multi-User Communications

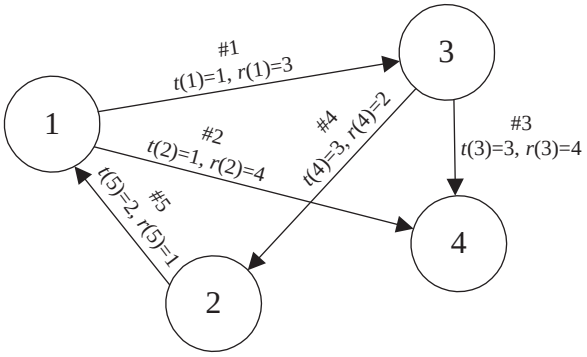
Consider a multi-user wireless scenario in which several terminals coexist in the same area. Among these terminals, U communications or links are established and access the common channel at the same time using the same frequency band. The adopted modulation, as in Chapter 3, is a N-carriers OFDM. All the terminals in the system are allowed to have multiple antennas and each of them is able to transmit and/or receive. Each communication or link is assigned to two terminals, where one of them is the transmitter and the other one is the receiver.

4.3.1 MIMO-OFDM Multi-User System and Signal Models

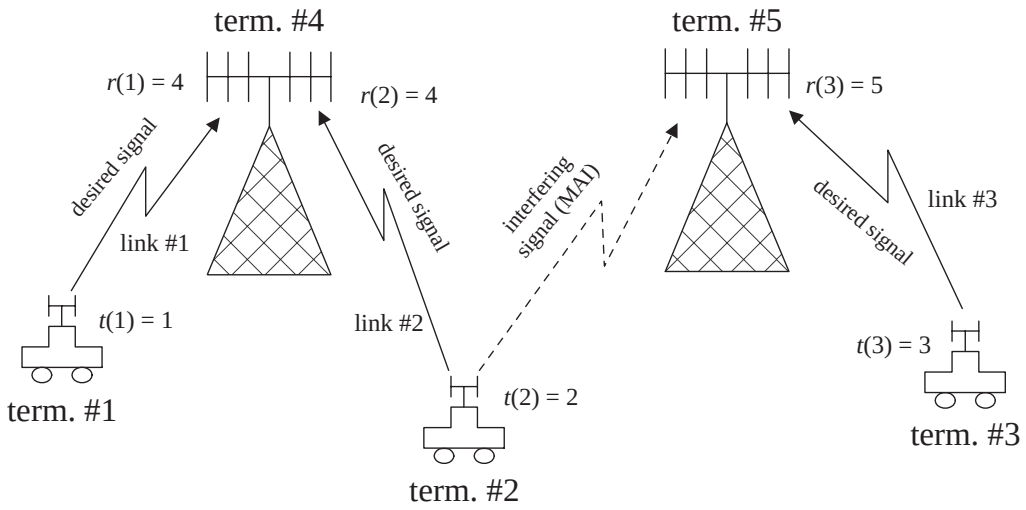
As in Chapter 3, the system design is based on a joint beamforming approach at both the transmitter and the receiver, where the beamvectors corresponding to different communications or links are allowed to be different. In this scenario, there is a set of terminals, where it is not differentiated between MT's and AP's or BS's, since all of them are allowed to transmit and/or receive. All the terminals in the system are numbered and the quantity of terminals may be different from the number of established links (see Figure 4.2).

Let $t(u)$ represent the terminal responsible for transmitting the information corresponding to the uth link, whereas $r(u)$ is the terminal receiving this information. In Figure 4.2, some examples of this kind of systems are given (a generic example and a more concrete one). The number of antennas of the terminal transmitting the information for the uth communication is $n_T^{(u)}$, whereas the receiver for the same link has $n_R^{(u)}$ antennas.

The signal model presented in Chapter 3, corresponding to the MIMO-OFDM single-user



(a)



(b)

Figure 4.2: (a) General configuration of a multi-user system with point-to-point communications. In this example, there are 5 simultaneous communications and 4 terminals. (b) Example of a configuration in a multi-user system with 3 communications or links. The number of terminals is 5, where 3 of them are MT's and the other 2 ones are BS's.

system, can be generalized to the case of a multi-user scenario as shown in the following. The snapshot vector, i.e., the vector containing the received signal samples at all the antennas corresponding to the uth link and the kth subcarrier of the OFDM modulation is

$$\mathbf{y}_k^{(r(u))}(m) = \sum_{l=1}^{U} \mathbf{H}_k^{(t(l),r(u))} \mathbf{b}_k^{(l)} s_k^{(l)}(m) + \mathbf{n}_k^{(r(u))}(m) \in \mathcal{C}^{n_R^{(u)} \times 1}, \tag{4.11}$$

which generalizes expression (3.10) ((3.10) is a particular case of (4.11) when there is only one communication in the scenario, i.e., $U = 1$). The snapshot vector $\mathbf{y}_k^{(r(u))}(m)$ has $n_R^{(u)}$ complex components and contains the contribution of all the transmitted signals for all the communications. The time index for the OFDM symbol is m, as in Chapter 3. The channel matrix $\mathbf{H}_k^{(t(l),r(u))} \in \mathcal{C}^{n_R^{(u)} \times n_T^{(l)}}$ represents the MIMO channel response at the kth subcarrier between the $t(l)$th and the $r(u)$th terminals, i.e., between the terminal transmitting the signal corresponding to the lth communication, and the terminal responsible for detecting the symbols of the uth communication. It is considered that $\mathbf{H}^{(i,i)} = \mathbf{0}$, $\forall i$, which means that each terminal is not receiving the signal transmitted by itself. The transmit beamvector corresponding to the lth communication and the kth subcarrier is represented by $\mathbf{b}_k^{(l)} \in \mathcal{C}^{n_T^{(l)} \times 1}$, whereas $s_k^{(l)}(m)$ is the zero-mean information symbol transmitted at the same carrier and for the same communication during the mth OFDM symbol, and whose energy is assumed to be normalized ($\mathbb{E}\left[|s_k^{(l)}(m)|^2\right] = 1$). Finally, $\mathbf{n}_k^{(r(u))}(m) \in \mathcal{C}^{n_R^{(u)} \times 1}$ represents the zero-mean noise plus interferences contribution at the terminal detecting the symbols corresponding to the uth communication at the kth subcarrier, whose correlation matrix is given by $\mathbf{\Phi}_n^{(r(u))}(k) \triangleq \mathbb{E}\left[\mathbf{n}_k^{(r(u))}(m) \mathbf{n}_k^{(r(u))^H}(m)\right] \in \mathcal{C}^{n_R^{(u)} \times n_R^{(u)}}$. This signal model is quite general and can fit in with many known systems including, but not limited to, the cellular communication systems, both for uplink and downlink.

4.3.2 Optimum Single-User Receiver and Resulting SNIR

The receiver, as in Chapter 3, is based on the application of a beamvector. For every communication u and subcarrier k, a linear combiner or beamformer $\mathbf{a}_k^{(u)}$ is applied to the set of received samples at the same carrier collected in the snapshot vector $\mathbf{y}_k^{(r(u))}(m)$, obtaining the soft-estimate of the transmitted information symbol

$$\begin{aligned} r_k^{(u)}(m) &\triangleq \mathbf{a}_k^{(u)H} \mathbf{y}_k^{(r(u))}(m) \qquad (4.12) \\ &= \mathbf{a}_k^{(u)H} \mathbf{H}_k^{(t(u),r(u))} \mathbf{b}_k^{(u)} s_k^{(u)}(m) + \mathbf{a}_k^{(u)H} \sum_{l=1, l \neq u}^{U} \mathbf{H}_k^{(t(l),r(u))} \mathbf{b}_k^{(l)} s_k^{(l)}(m) + \mathbf{a}_k^{(u)H} \mathbf{n}_k^{(r(u))}(m). \end{aligned}$$

The hard-estimate is finally obtained by demodulating $r_k^{(u)}(m)$, i.e.,

$$\widehat{s}_k^{(u)}(m) \triangleq \mathrm{dec}\left\{ r_k^{(u)}(m) \right\}. \tag{4.13}$$

The quality of the estimate depends of the SNIR at the output of the beamformer. The optimum receive beamvector has been found previously in §3.2.3 and corresponds to the Wiener

filter, also called whitened matched filter:

$$\mathbf{a}_k^{(u)} = \alpha_k^{(u)} {\mathbf{R}_k^{(u)}}^{-1} \mathbf{H}_k^{(t(u),r(u))} \mathbf{b}_k^{(u)}, \tag{4.14}$$

$$\mathbf{R}_k^{(u)} = \mathbf{\Phi}_n^{(r(u))}(k) + \sum_{l=1,l\neq u}^{U} \mathbf{H}_k^{(t(l),r(u))} \mathbf{b}_k^{(l)} {\mathbf{b}_k^{(l)}}^H {\mathbf{H}_k^{(t(l),r(u))}}^H, \tag{4.15}$$

where $\mathbf{R}_k^{(u)}$ is the total noise plus interferences correlation matrix seen at the receiver for the uth link at the kth subcarrier, i.e., including the contribution of the transmitted signals from other communications different from u which are seen as interferences. The coefficient $\alpha_k^{(u)}$ is a scalar factor does not modify of the resulting SNIR and can be calculated to obtain a normalized value of the equivalent channel at the detection stage, i.e., ${\mathbf{a}_k^{(u)}}^H \mathbf{H}_k^{(t(u),r(u))} \mathbf{b}_k^{(u)} = 1$. Note that the design of each receive beamvector can be carried out independently at each carrier and communication, since the SNIR does not depend on the receive beamvectors at other subcarriers and/or for other communications. The resulting SNIR at the kth subcarrier for the uth communication using the design shown in (4.14) is

$$\mathrm{SNIR}_k^{(u)} = {\mathbf{b}_k^{(u)}}^H {\mathbf{H}_k^{(t(u),r(u))}}^H {\mathbf{R}_k^{(u)}}^{-1} \mathbf{H}_k^{(t(u),r(u))} \mathbf{b}_k^{(u)}, \tag{4.16}$$

which is obtained in the same way as deduced in §3.2.3.

As can be seen in (4.14), the optimum receive beamvector $\mathbf{a}_k^{(u)}$ depends on both the transmit beamvector for the same link $\mathbf{b}_k^{(u)}$ and for all the other communications $\left\{\mathbf{b}_k^{(l)}\right\}_{l=1,l\neq u}^{l=U}$ through the correlation matrix $\mathbf{R}_k^{(u)}$. This induces a coupling effect making difficult the optimization and the design of the transmit beamvectors. In §4.4, the problem of the joint design of all the transmit beamvectors is addressed, taking into account explicitly this coupling effect.

When no channel coding is applied and the OFDM modulation is used, the effective error probability for the uth communication $P_{\mathrm{e,eff}}^{(u)}$ is expressed as

$$P_{\mathrm{e,eff}}^{(u)} \triangleq \frac{1}{N} \sum_{k=0}^{N-1} P_e^{(u)}(k), \tag{4.17}$$

where $P_e^{(u)}(k)$ is the error probability associated to the kth subcarrier. In case that the interferences are approximately Gaussian distributed,[1] the error probability for the kth subcarrier can be expressed as

$$P_e^{(u)}(k) \simeq \alpha_m \mathcal{Q}\left(\sqrt{k_m \mathrm{SNIR}_k^{(u)}}\right), \tag{4.18}$$

where $\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} \mathrm{d}t$, α_m and k_m are constants depending on the modulation applied to the subcarriers [Pro95] (for BPSK, $\alpha_m = 1$ and $k_m = 2$), and $\mathrm{SNIR}_k^{(u)}$ corresponds to (4.16) assuming the use of the optimum receiver beamvector (4.14).

[1] This is not true in this signal model, since the interferences from other terminals are OFDM signals and, therefore, they are not Gaussian distributed in the frequency domain. However, if the number of users is high, this assumption is more accurate thanks to the central limit theorem.

4.4 Application of Simulated Annealing to the Optimization Problem

The expression of the optimum receive beamvector (4.14) shows that it depends on the transmit beamvectors. The attention is now focused on the joint design of all the transmit beamvectors for all the communications and all the subcarriers.

When designing the transmit beamvectors, an objective function or optimization criterion has to be identified, as well as a set of design constraints. A desirable objective is the minimization of the total transmit power, since in wireless networks, high transmit powers imply a shorter lifetime of the MT's. As explained in Chapter 3, the power used for transmitting the information symbols corresponding to the kth subcarrier and the uth communication is proportional to $\left\|\mathbf{b}_k^{(u)}\right\|^2$ (see §3.2.3). Taking this into account, a cost function is defined equal to the total transmit power and depending on the transmit beamvectors, which are the optimization variables:

$$f\left(\left\{\mathbf{b}_k^{(u)}\right\}\right) \triangleq P_T\left(\left\{\mathbf{b}_k^{(u)}\right\}\right) = \sum_{u=1}^{U}\sum_{k=0}^{N-1}\left\|\mathbf{b}_k^{(u)}\right\|^2 = \sum_{u=1}^{U}\sum_{k=0}^{N-1}\mathbf{b}_k^{(u)H}\mathbf{b}_k^{(u)}. \tag{4.19}$$

Besides the objective function f, additional constraints have to be included in order to avoid the trivial solution minimizing the total transmit power: $\mathbf{b}_k^{(u)} = \mathbf{0},\ \forall k, u$. Here, two kinds of constraints are proposed. The first ones refer to the minimum QoS for each communication or link and are mandatory, whereas the other ones are related to the maximum individual transmit powers for a concrete set of terminals. This set of terminals can be empty and, therefore, the individual transmit power constraints are optional.

- *QoS constraints:* these constraints are formulated in terms of a maximum effective error probability for each communication and can be expressed as

$$P_{\text{e,eff}}^{(u)} \leq \gamma^{(u)}, \qquad u = 1, \ldots, U, \tag{4.20}$$

  where $\gamma^{(u)}$ is the maximum permitted error probability for the uth communication and, therefore, is an input parameter of the optimization problem. This formulation generalizes the results presented in [Lok00] for a MC-CDMA system, and in [Cha02, Boc02, Vis99, Ben02] for flat fading channels, where the QoS constraints were formulated in terms of the SNIR instead of the effective error probability. In [Won01, Rhe04, Ser04], the transmit power was stated to be a prefixed value and the final objective was the optimization of the mean quality of all the links in terms of the capacity, the MSE, etc. and, therefore, no QoS could be guaranteed for each communication. In all the cases, the corresponding optimization problems were non-convex and the proposed algorithms to solve the problem for the most general scenario comprising several BS's and MT's with multiple antennas,

where shown to be inefficient in the sense that they might find local suboptimum solutions instead of the global optimum one.

- *Individual transmit power constraints:* in addition to the previous QoS constraints, optional constraints can also be included regarding the maximum transmit powers for some terminals. This is specially useful for MT's with a power limited battery in an uplink transmission. Let Υ be the set of terminals to which these constraints are applied. They can be formulated as

$$P_T^{(i)} \triangleq \sum_{u=1,t(u)=i}^{U} \sum_{k=0}^{N-1} \left\| \mathbf{b}_k^{(u)} \right\|^2 \leq P_{\max}^{(i)}, \qquad \forall i \in \Upsilon, \tag{4.21}$$

where $P_{\max}^{(i)}$ represents the maximum transmit power for the ith terminal. These kinds of constraints have not been included in any of the works referenced in this dissertation.

Summarizing, the problem to be solved can be written as

$$\begin{array}{lll} \underset{\{\mathbf{b}_k^{(u)}\}}{\text{minimize}} & \sum_{u=1}^{U} \sum_{k=0}^{N-1} \left\| \mathbf{b}_k^{(u)} \right\|^2 & \\ \text{subject to} & P_{\text{e,eff}}^{(u)} \leq \gamma^{(u)}, & u = 1, \ldots, U, \\ & P_T^{(i)} \leq P_{\max}^{(i)}, & i \in \Upsilon. \end{array} \tag{4.22}$$

Currently, there exists no closed-form solution to this extremely complicated constrained optimization problem, since it is not convex [Boy04]. Although in this case the objective function f is convex in the optimization variables $\left\{\mathbf{b}_k^{(u)}\right\}$, the constraint set is not. In order to prove this last statement, let us consider a simple counterexample corresponding to a scenario with only one communication ($U = 1$) using an OFDM modulation with only one carrier ($N = 1$). In this case, the maximum effective error probability constraint is equivalent to a minimum SNIR constraint. Let us assume that $\mathbf{H} = \mathbf{I}$ and that $\mathbf{\Phi}_n = \mathbf{I}$ (the sub- and super-indexes are obviated to facilitate the notation). According to this, the QoS constraint can be formulated as $\mathbf{b}^H \mathbf{b} \geq \text{SNIR}_{\min}$, where $\text{SNIR}_{\min}$ is the minimum required SNIR to attain the error probability constraint. This constraint can be represented geometrically as the exterior of a sphere in the variable vector $\mathbf{b}$, which, obviously, is not convex. Due to the non-convex behavior of the problem, if a classical GS or AM method is applied to find the optimal design, a local minimum may be found instead of the global optimum one in the constraint set. Some papers have proposed GS techniques, such as in [Lok00], or AM algorithms, as in [Won01, Cha02, Ser04, Vis99, Ben02], among others, all of them suffering from the same problem described above, as was clearly shown in [Cha02] and other works. As the objective is to find the global optimum design minimizing the total transmit power, the SA algorithm can be applied. Besides, in GS and AM techniques, it may be extremely difficult to include any kind of constraint, although in the case of SA, this can be done easily,

as will be shown in the following. Specifically, for the case of GS, the constraint functions are required to be differentiable, although this is not necessary in SA.

The existence of a feasible solution is assumed, i.e., a collection of transmit beamvectors fulfilling all the constraints simultaneously. In case that the optimization problem (4.22) is infeasible, then the SA algorithm will not converge to any acceptable design.

In the following, the main steps to apply the SA algorithm to this optimization problem are described. In each iteration of the SA algorithm, there exists a collection of transmit beamvectors $\left\{\mathbf{b}_k^{(u)}\right\}$, which is called the *current solution*. Given the current solution, which is equivalent to a concrete particles arrangement or state in the annealing process in physics, a *new solution*, i.e., a new collection of transmit beamvectors, is proposed. If it is "better" than the original one, then it is retained as the new current solution. On the contrary, if it is "worse", then the proposed solution is accepted with a certain probability higher than 0. In that case, a *hill-climbing* is performed, which permits the algorithm to escape from local minimums, as explained in §4.2.1. The parameter controlling this acceptance probability is the *temperature* T, as in the case of the annealing in physics. The higher the temperature, the higher the acceptance probability. The temperature has to be lowered, so that asymptotically only "better" solutions are accepted and a minimum is approached. The meaning of "better" and "worse" is related to the definition of a modified cost function $\widetilde{f}$ that depends on the transmit beamvectors and takes into account both the total transmit power (4.19) and whether the constraints in (4.22) are fulfilled or not. This modified cost function corresponds to the energy of a state in physics, whose minimization is the goal of the annealing process. As in the physical process, if the temperature is lowered very slowly, the optimum state with the minimum global energy is attained, i.e., the global minimum of the total transmit power is reached while still fulfilling all the constraints.

The description of the application of SA to the optimization problem (4.22) is presented in the following. First, the modified cost function is defined and, afterwards, the procedures corresponding to the generation and acceptance of new solutions and the system cooling are described.

- *Modified cost function definition:*

$$\widetilde{f}\left(\left\{\mathbf{b}_k^{(u)}\right\};T\right) \triangleq P_T\left(\left\{\mathbf{b}_k^{(u)}\right\}\right) + \frac{\alpha}{T}\sum_{u=1}^{U}\left(\log\frac{P_{\mathrm{e,eff}}^{(u)}}{\gamma^{(u)}}\right)^{+2} + \frac{\alpha}{T}\sum_{i\in\Upsilon}\left(\log\frac{P_T^{(i)}}{P_{\max}^{(i)}}\right)^{+2}. \quad (4.23)$$

  This modified cost function, which also depends on the temperature, is equal to the total transmit power plus and quadratic penalty term. This penalty term takes into account whether the effective error probability for each communication and the individual transmit power for the specified terminals are higher than the maximum permitted values. Besides, this penalty term is inversely proportional to the temperature, so that when the system is

"cooled", the modified cost function increases if the constraints are not fulfilled. Since the algorithm looks for the solution that minimizes $\widetilde{f}$, a collection of transmit beamvectors that does no fulfill the constraints will be avoided. Additionally, in the simulations, it has been shown that this rule performs quite well in terms of convergence speed. The parameter α is a proportional factor for the penalty term and its value has been adjusted to $\alpha = 100$ in order to have good convergence properties. The penalty term is based on relative comparisons of the effective error probabilities and the individual transmit powers with the maximum permitted values by means of the $\log(\cdot)$ function. These kinds of comparisons have been chosen, since it has been observed experimentally that they behave better than absolute comparisons. Note, however, that other kinds of penalty functions could have been used.

- *Proposed solution generation:*

$$\widetilde{\mathbf{b}}_k^{(u)} = \mathbf{b}_k^{(u)} + \mathbf{w}_k^{(u)}, \qquad \mathbf{w}_k^{(u)} \sim \mathcal{CN}(\mathbf{0}, \sigma_b^2 \mathbf{I}), \qquad k = 0, \ldots, N-1, \;\; u = 1, \ldots, U. \tag{4.24}$$

The proposed solution is generated by applying independent complex circularly symmetric Gaussian noise with zero-mean and variance σ_b^2 to the components of the transmit beamvectors. This noise is used to generate any possible collection of transmit beamvectors in a continuous solution space, as explained in §4.2.3. The acceptance ratio is monitored for every value of the temperature T. In case that it is lower than 0.1 for 5 times, then the variance of the Gaussian noise is decreased by means of an exponential rule ($\sigma_b^2 \leftarrow 0.95\sigma_b^2$). This is done in this way as it has been shown experimentally that this rule improves the convergence speed of the algorithm.

- *Probability of acceptance of the proposed solution:*

$$\Pr\left(\text{accept } \left\{\widetilde{\mathbf{b}}_k^{(u)}\right\} \text{ from } \left\{\mathbf{b}_k^{(u)}\right\}\right) \triangleq \begin{cases} 1, & \widetilde{f}\left(\left\{\widetilde{\mathbf{b}}_k^{(u)}\right\}\right) \leq \widetilde{f}\left(\left\{\mathbf{b}_k^{(u)}\right\}\right), \\ \exp\left(-\dfrac{\widetilde{f}\left(\left\{\widetilde{\mathbf{b}}_k^{(u)}\right\}\right) - \widetilde{f}\left(\left\{\mathbf{b}_k^{(u)}\right\}\right)}{T}\right), & \widetilde{f}\left(\left\{\widetilde{\mathbf{b}}_k^{(u)}\right\}\right) > \widetilde{f}\left(\left\{\mathbf{b}_k^{(u)}\right\}\right). \end{cases} \tag{4.25}$$

This acceptance probability corresponds to the Metropolis criterion, as described in [Aar89, Laa87] and explained in §4.2.2, and is related to the Maxwell-Boltzmann approximation of the Fermi-Dirac distribution describing the energy of an electron in different energetic levels. This criterion was initially used in thermodynamics to simulate a thermal equilibrium process showing that, using this criterion, the system could arrive at the minimum possible energy, i.e., the global optimum state, provided the temperature was lowered slowly enough. Afterwards, this philosophy was adopted by the SA algorithm as an efficient method to find the global minimum of non-convex problems.

- *System cooling:*

$$T \leftarrow \beta T, \qquad \beta = 0.99. \tag{4.26}$$

As described in this equation, the temperature is lowered very slowly by means of a decreasing exponential rule, as explained in §4.2.2 (see (4.7) and Figure 4.1). This value of β has been chosen since in the simulations, it has been shown to provide good convergence properties, while still guaranteeing that the global optimum solution is obtained. As seen in (4.25), the hotter the system, the higher the acceptance probability. Consequently, when the temperature is high, most of the proposed transmit beamvectors are accepted, which means that they are searching over the range of all the possible spatial directions. When the temperature is lowered, this range is reduced and the accepted transmit beamvectors begin to look for the best spatial directions, i.e., for the spatial directions coupling the maximum power towards the desired terminal while reducing the interference towards the other ones.

In the SA algorithm, initially the temperature should be high enough so that most of the proposed solutions are accepted. The initial transmit beamvectors are set equal to all-zero vectors. Note, however, that the initialization of the beamvectors is not important since in the first iterations, most of the proposed solutions are accepted and the variance of the noise used to generate and propose new solutions is also high. In the application of the SA algorithm to this optimization problem, 100 iterations are performed for every value of T.

In the initialization step of the algorithm, the variance of the noise σ_b^2 is set equal to the mean power necessary at the transmitters to attain the required QoS in terms of the maximum effective error probability, assuming no interference among users, i.e., assuming single-user links with no MAI, as described in Chapter 3. The temperature T has to be adjusted until the ratio of accepted solutions is between 95 and 100 %. See Figure 4.3(a) for a flowchart corresponding to the initialization of the SA algorithm.

Once the SA has been initialized, 100 generations of new solutions are performed for each value of the temperature. The temperature is then decreased using (4.26). The ratio of non-accepted solutions is monitored for each value of T. In case that this ratio is lower than 10 % for 5 times, then the variance of the Gaussian noise used to generate new solutions is decreased using $\sigma_b^2 \leftarrow 0.95\sigma_b^2$. Thanks to this, the convergence speed of the algorithm increases, since the minimum can be found in a more accurate way when it is approached. The algorithm ends when the ratio of new accepted solutions is equal to 0 and the value of the cost function has converged, i.e., the global minimum has been found. See Figure 4.3(b) for a flowchart corresponding to the main iterations of the SA algorithm.

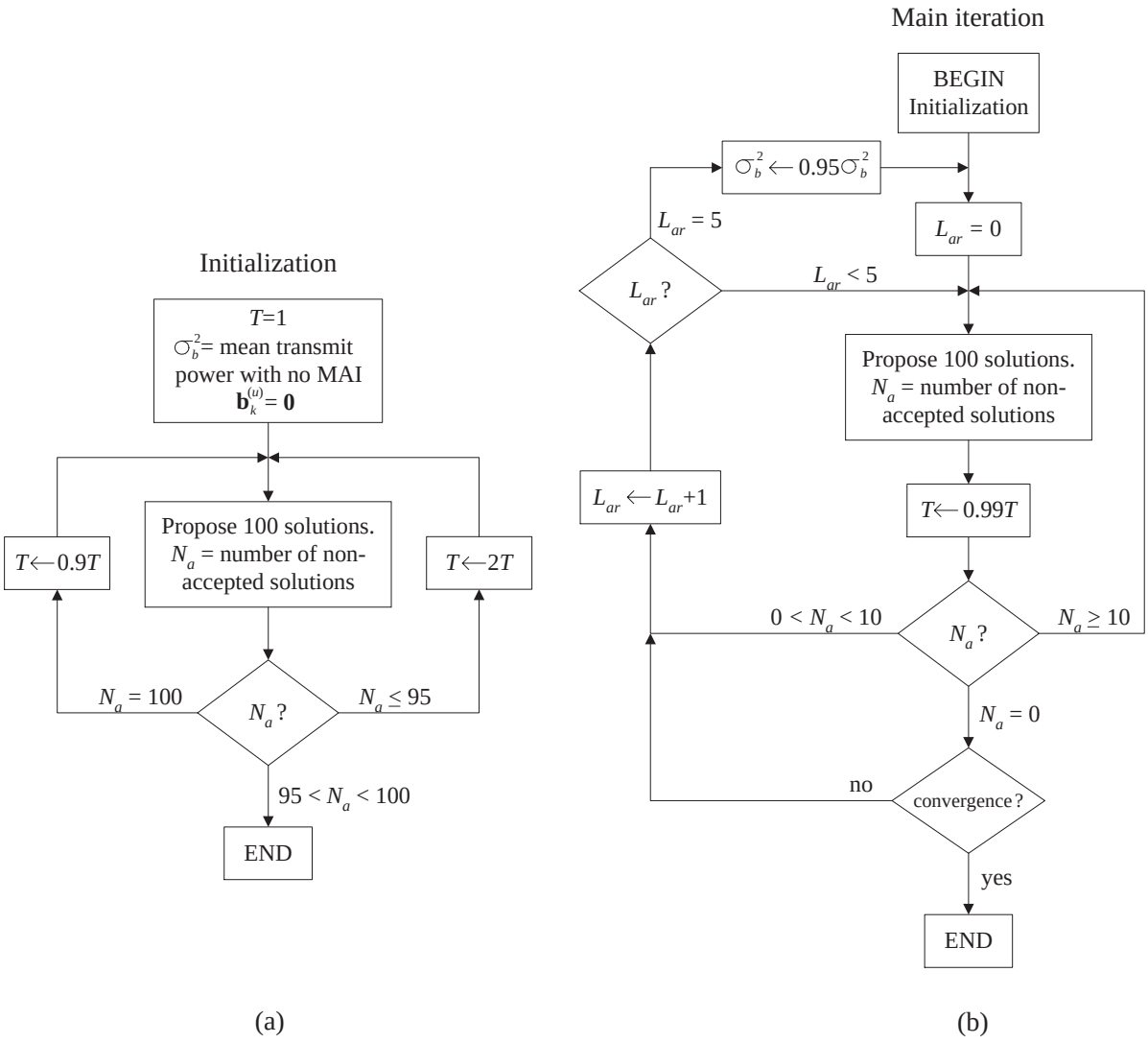


Figure 4.3: Flowchart corresponding to the application of the SA algorithm. (a) Initialization of the parameters of the SA algorithm. (b) Main iteration of the SA algorithm.

4.5 Other Suboptimum Techniques

In §4.4, the SA technique has been proposed to find the global optimum design of the stated constrained optimization problem (4.22). Now, two alternative algorithms are presented based on GS and AM methods.

4.5.1 Gradient Search Algorithm

A classical approach different from the SA consists in the application of a gradient technique although, as has been already commented, the main drawback of this family of algorithms is that they may converge to local suboptimum designs. In order to compare the SA with other classical approaches, in this section, an iterative gradient technique is proposed based on the classical Lagrange multipliers method and the quadratic penalty term [Lok00, Ber82]. This technique is based on the definition of a Lagrangian expression L, in whose formulation, it is taken into account explicitly that the optimal solution implies that the error probability constraints (4.20) are fulfilled with equality. Under this assumption, that can be shown easily, the Lagrangian expression is formulated as

$$L\left(\left\{\mathbf{b}_k^{(u)}\right\};\lambda\right) \triangleq P_T\left(\left\{\mathbf{b}_k^{(u)}\right\}\right) + \lambda\left[\sum_{j=1}^{U}\left(\log\frac{P_{\text{e,eff}}^{(j)}}{\gamma^{(j)}}\right)^2 + \sum_{i\in\Upsilon}\left(\log\frac{P_T^{(i)}}{P_{\max}^{(i)}}\right)^{+2}\right]. \tag{4.27}$$

The equations that show how to update the transmit beamvectors and the penalty factor λ correspond to the well known gradient descent and ascent techniques, as also used in [Lok00]:

$$\mathbf{b}_k^{(u)} \leftarrow \mathbf{b}_k^{(u)} - \mu\nabla_{\mathbf{b}_k^{(u)H}} L, \tag{4.28}$$

$$\lambda \leftarrow \lambda + \mu\left[\sum_{j=1}^{U}\left(\log\frac{P_{\text{e,eff}}^{(j)}}{\gamma^{(j)}}\right)^2 + \sum_{i\in\Upsilon}\left(\log\frac{P_T^{(i)}}{P_{\max}^{(i)}}\right)^{+2}\right], \tag{4.29}$$

where μ is the step-size parameter that has to be adjusted to cope with the tradeoff between the convergence speed and the convergence itself. The initial transmit beamvectors can be calculated assuming that there is no interference among communications, i.e., assuming no MAI, as shown in Chapter 3. The initial value of the penalty factor λ is set equal to 0.

In the following, the necessary expressions to calculate $\nabla_{\mathbf{b}_k^{(u)H}} L$ are provided. In order to facilitate the notation, an uplink scenario is assumed with several MT's transmitting to a single BS, which is responsible for detecting the symbols transmitted by all the MT's. The modulation of the subcarriers is considered to be BPSK. In this scenario, $\mathbf{H}_k^{(u)}$ represents the response of the MIMO channel at the kth subcarrier between the uth MT and the BS. The extension to other kinds of scenarios is quite simple by using very similar expressions. The function δ_Υ is defined

as $\delta_{\Upsilon}(u) = 1,\ u \in \Upsilon$ and $\delta_{\Upsilon}(u) = 0,\ u \notin \Upsilon$. Finally, $\nabla_{\mathbf{b}_k^{(u)H}} L$ can be written as follows:

$$\nabla_{\mathbf{b}_k^{(u)H}} L = \mathbf{b}_k^{(u)} + 2\lambda \sum_{j=1}^{U} \frac{1}{P_{\mathrm{e,eff}}^{(j)}} \left(\log \frac{P_{\mathrm{e,eff}}^{(j)}}{\gamma^{(j)}} \right) \nabla_{\mathbf{b}_k^{(u)H}} P_{\mathrm{e,eff}}^{(j)} + \delta_{\Upsilon}(u) 2\lambda \frac{\mathbf{b}_k^{(u)}}{P_T^{(u)}} \left(\log \frac{P_T^{(u)}}{P_{\max}^{(u)}} \right)^{+}. \quad (4.30)$$

The expression of $\nabla_{\mathbf{b}_k^{(u)H}} P_{\mathrm{e,eff}}^{(j)}$ depends on j. First, it is given for the case $j = u$:

$$\nabla_{\mathbf{b}_k^{(u)H}} P_{\mathrm{e,eff}}^{(u)} = -\frac{1}{N\sqrt{2\pi}} \exp\left(-\mathrm{SNIR}_k^{(u)}\right) \frac{1}{\sqrt{2\mathrm{SNIR}_k^{(u)}}} \mathbf{H}_k^{(u)H} \mathbf{R}_k^{(u)-1} \mathbf{H}_k^{(u)} \mathbf{b}_k^{(u)}. \quad (4.31)$$

For the case $j \neq u$, the expression is as follows, where the matrix inversion lemma [Gol96] has been used:

$$\nabla_{\mathbf{b}_k^{(u)H}} P_{\mathrm{e,eff}}^{(j)} = -\frac{1}{N\sqrt{2\pi}} \exp\left(-\mathrm{SNIR}_k^{(j)}\right) \frac{1}{\sqrt{2\mathrm{SNIR}_k^{(j)}}} \nabla_{\mathbf{b}_k^{(u)H}} \mathrm{SNIR}_k^{(j)}, \quad (4.32)$$

$$\begin{aligned} \nabla_{\mathbf{b}_k^{(u)H}} \mathrm{SNIR}_k^{(j)} = & \frac{\mathbf{H}_k^{(u)H} \mathbf{R}_k^{(j,u)-1} \mathbf{H}_k^{(u)} \mathbf{b}_k^{(u)}}{\left(1 + \mathbf{b}_k^{(u)H} \mathbf{H}_k^{(u)H} \mathbf{R}_k^{(j,u)-1} \mathbf{H}_k^{(u)} \mathbf{b}_k^{(u)}\right)^2} \left| \mathbf{b}_k^{(j)H} \mathbf{H}_k^{(j)H} \mathbf{R}_k^{(j,u)-1} \mathbf{H}_k^{(u)} \mathbf{b}_k^{(u)} \right|^2 \\ & - \frac{\mathbf{b}_k^{(j)H} \mathbf{H}_k^{(j)H} \mathbf{R}_k^{(j,u)-1} \mathbf{H}_k^{(u)} \mathbf{b}_k^{(u)}}{1 + \mathbf{b}_k^{(u)H} \mathbf{H}_k^{(u)H} \mathbf{R}_k^{(j,u)-1} \mathbf{H}_k^{(u)} \mathbf{b}_k^{(u)}} \mathbf{H}_k^{(u)H} \mathbf{R}_k^{(j,u)-1} \mathbf{H}_k^{(j)} \mathbf{b}_k^{(j)}, \end{aligned} \quad (4.33)$$

$$\mathbf{R}_k^{(j,u)} = \mathbf{\Phi}_n(k) + \sum_{l=1, l\neq j, l\neq u}^{U} \mathbf{H}_k^{(l)} \mathbf{b}_k^{(l)} \mathbf{b}_k^{(l)H} \mathbf{H}_k^{(l)H}. \quad (4.34)$$

The complete proof of these expression is given in Appendix 4.A.

As stated previously, one of the main drawbacks of the GS technique is that a local sub-optimum design may be found instead of the global optimum one. This could be solved by using different initial sets of transmit beamvectors selected randomly. Note, however, that this increases the computational load and does not guarantee a successful result.

4.5.2 Alternate & Maximize Algorithm

Finally, another classical solution that has been used previously by many authors in papers such as [Won01, Ser04, Vis99, Ben02], among others, is the AM algorithm. In the problem described in this chapter, the SNIR for a concrete communication and carrier depends not only on the transmit beamvector for the considered link, but also on all the transmit beamvectors for all the other communications through the covariance matrix, as shown in (4.15) and (4.16). The AM algorithm is an iterative technique, so that in each step, the beamvectors associated to a concrete user are designed assuming that the beamvectors for all the other users are fixed, i.e., assuming

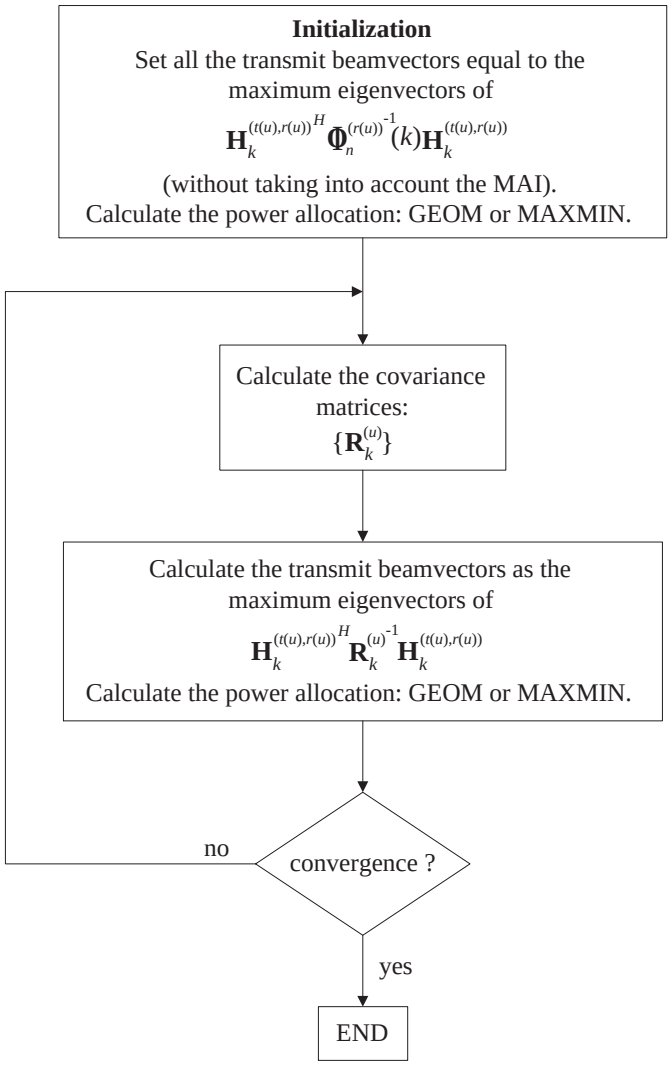


Figure 4.4: Flowchart corresponding to the application of the AM algorithm.

that the noise plus interferences covariance matrix is known. Obviously, when a beamvector for a user is designed, the covariance matrix for the other users change and, therefore, the technique has to be applied iteratively until convergence is attained.

In this subsection, the description of the AM algorithm is provided taking into account only the maximum error probability constraints but not the individual transmit power constraints, since their inclusion in the algorithm is extremely difficult. In each step, the optimum transmit beamvector maximizing the SNIR corresponds to the eigenvector associated to the maximum eigenvalue of the matrix $\mathbf{H}_k^{(t(u),r(u))^H} \mathbf{R}_k^{(u)^{-1}} \mathbf{H}_k^{(t(u),r(u))}$, as shown in §3.2.3 (see (3.22)). Besides this, an adequate power allocation among the subcarriers of the OFDM modulation has to be calculated, such that the QoS constraints in terms of the maximum error probabilities are fulfilled. Two different power allocation strategies are assumed: GEOM (corresponding to a uniform

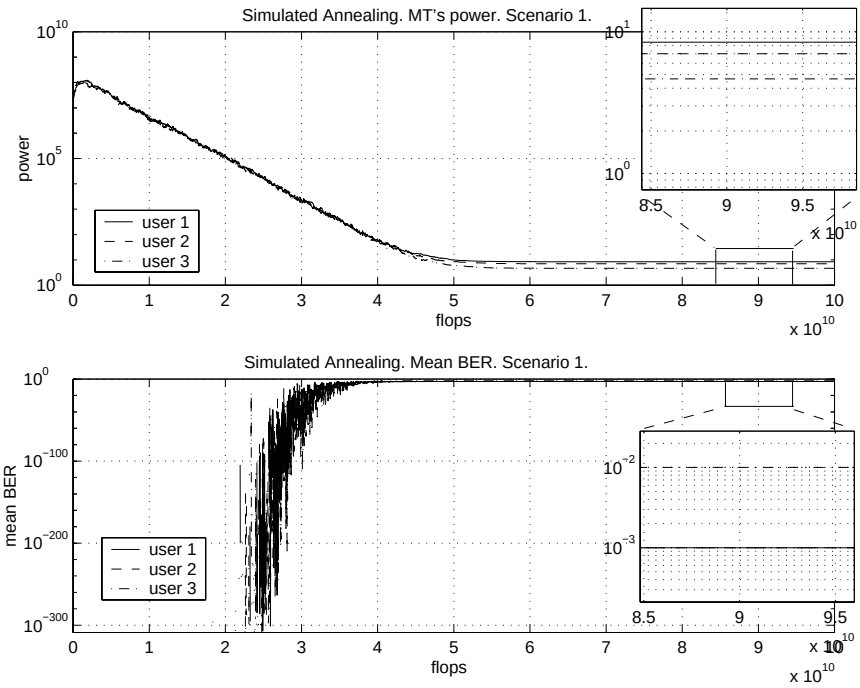


Figure 4.5: Performance of the SA algorithm in scenario 1 with 3 MT's and 1 BS in an uplink configuration, all of them with 5 antennas. The three users have the same path loss. No individual transmit power constraint is applied. OFDM modulation with 16 subcarriers. Evolution of the BER's and the individual transmit powers for the three users vs. the number of floating point operations.

power allocation) and MAXMIN (equivalent to attaining the same SNIR in all the subcarriers) techniques (see §3.3). The initialization of the algorithm is performed by calculating the optimum transmit beamvectors assuming that there is no MAI, i.e., that there is no interference among different users or communications, as explained in Chapter 3. See Figure 4.4 for a flowchart corresponding to the application of the AM algorithm.

The main disadvantage of this algorithm, as commented previously and in papers such as [Ser04, Vis99, Ben02], is that the obtained solution may be a local suboptimum design instead of the global optimum one. Besides, there is no a-priori guarantee of convergence. A possible solution would consist in using different random initializations for the transmit beamvectors. Note, however, that this is an adhoc approach that does not control and guarantee that the global optimum design is obtained.

4.6 Simulation Results

In this section, some simulations results are given considering an uplink scenario with 3 MT's and one BS, all of them with 5 antennas. The OFDM modulation consists of $N = 16$ subcarriers.

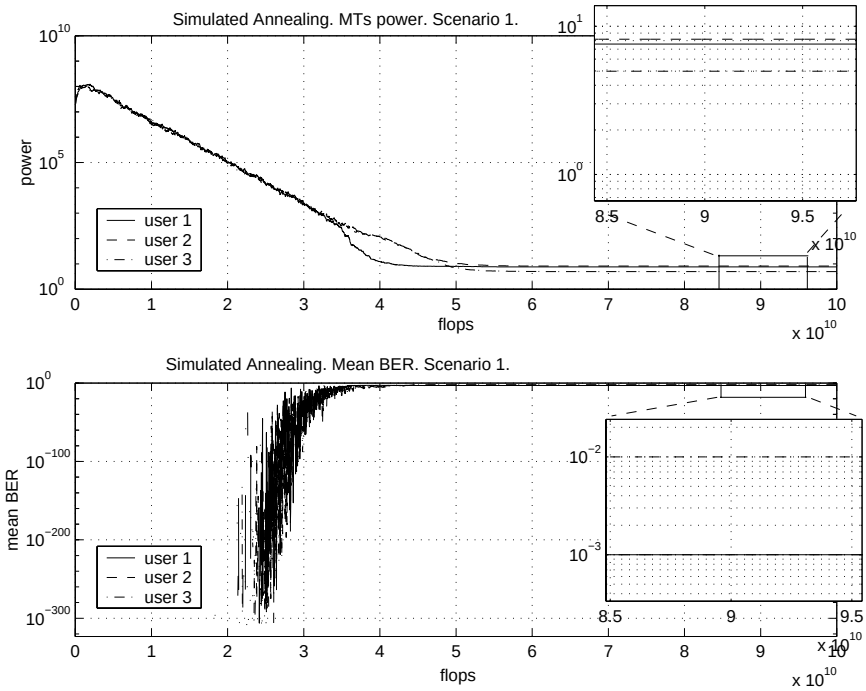


Figure 4.6: Performance of the SA algorithm in scenario 1 with 3 MT's and 1 BS in an uplink configuration, all of them with 5 antennas. The three users have the same path loss. The maximum transmit power for the first user is 8 W. OFDM modulation with 16 subcarriers. Evolution of the BER's and the individual transmit powers for the three users vs. the number of floating point operations.

The BS is responsible for detecting the symbols transmitted by all the MT's, taking as QoS constraints the following maximum effective error probabilities: 10^{-3}, 10^{-3}, and 10^{-2}, respectively. The value of the scalar α in (4.23) is 100. The noise is assumed to be white both in the time and the space domains with a normalized variance equal to 1, i.e., $\mathbf{\Phi}_n^{(r(u))}(k) = \mathbf{I}$, $\forall u, k$. The simulations and the algorithms are applied to a single realization of the multiple MIMO-OFDM channels, although the numerical expressions of the channel matrices are not provided for the sake of clarity.

In the first scenario, it is assumed that the path loss is very similar for all the users and that no individual transmit power constraint is applied to any MT. In Figure 4.5, the evolution of the powers allocated to the three users is shown, and also the mean BER's as the iterations of the SA algorithm go on, concluding that the proposed technique is able to find a design fulfilling the QoS constraints. The optimum power corresponding to the first user is 8.45 W, and the total power is 20.1 W.

If an individual transmit power constraint is applied to the first user equal to 8 W, then the results are those shown in Figure 4.6. The main conclusion is that, in this case, the SA allocates 7.6 W to the first user, whereas the other ones increase their corresponding power consumption.

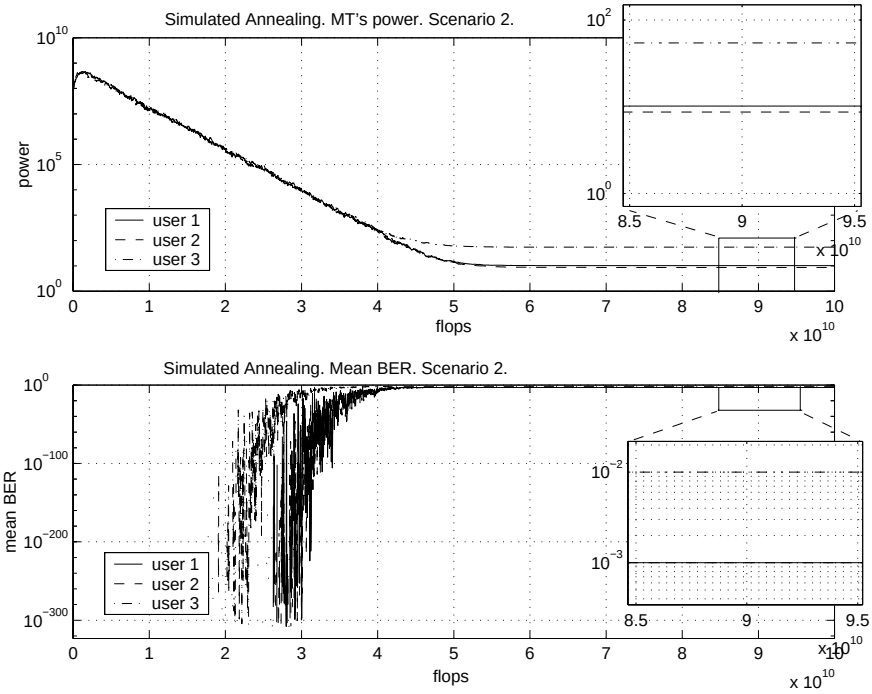


Figure 4.7: Performance of the SA algorithm in scenario 2 with 3 MT's and 1 BS in an uplink configuration, all of them with 5 antennas. The third user has a path loss equal to 12 dB with respect to the other users. OFDM modulation with 16 subcarriers. Evolution of the BER's and the individual transmit powers for the three users vs. the number of floating point operations.

As also shown, the global transmit power has increased up to 20.8 W. This increase of the total transmit power is expected, since in the second example, a more restrictive constraint is applied and, therefore, the optimization has to be carried out over a more limited set of transmit beamvectors fulfilling the constraints.

In Figures 4.7 and 4.8, a set of results are presented for the case of a scenario in which the third user has a path loss with respect to the first two users equal to 12 dB. Figure 4.7 corresponds to the application of the SA algorithm, whereas Figure 4.8 corresponds to the GS algorithm with a step-size equal to $\mu = 0.001$. The main conclusion is that, with the same computational load or number of floating point operations, the SA algorithm can fulfill the constraints, whereas the GS technique decreases importantly the convergence speed as the solution approaches these constraints. This is because the penalty term applied in the Lagrangian expression (4.27) is quadratic and, therefore, when calculating the derivatives in a point near from the fulfillment of the constraints, these derivatives tend to zero.

Simulations concerning the application of the AM algorithm have also been performed in the same scenario for two different power allocation strategies: GEOM and MAXMIN (see §3.3.1 and §3.3.3). From the simulations, it is concluded that AM has a high convergence speed. Table

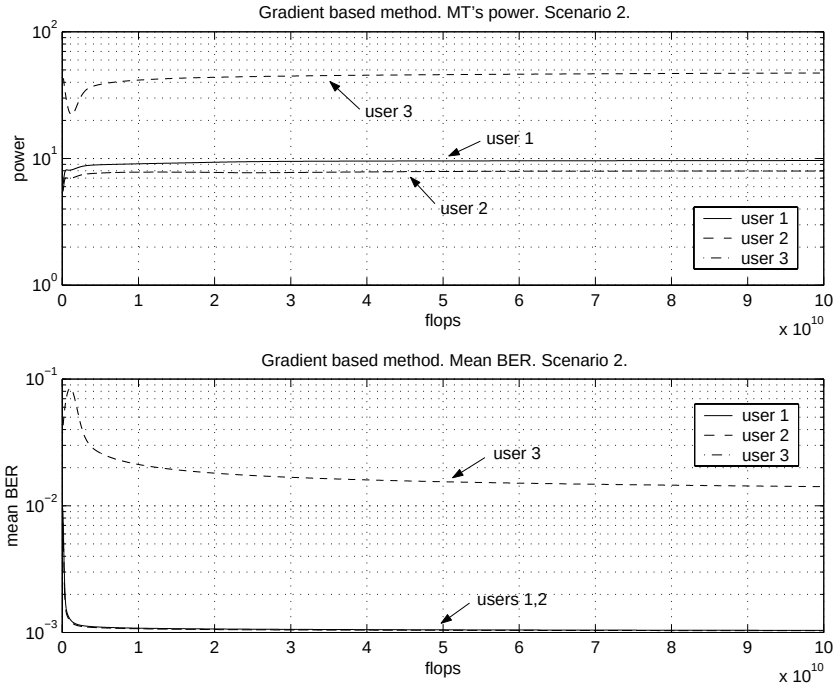


Figure 4.8: Performance of the GS algorithm in scenario 2 with 3 MT's and 1 BS in an uplink configuration, all of them with 5 antennas. The third user has a path loss equal to 12 dB with respect to the other users. OFDM modulation with 16 subcarriers. Evolution of the BER's and the individual transmit powers for the three users vs. the number of floating point operations.

4.1 shows a summary of the results for all the techniques. The conclusion is that GS does not find a solution fulfilling the constraints, whereas AM does not have this problem, as in the case of SA. The main drawback is that the necessary transmit power is higher for AM than for SA, concluding that a local suboptimum design has been found. Indeed, and as explained in [Vis99], the non-convexity and the number of local minima increases as more BS's and MT's are coexisting in the same area.

4.7 Chapter Summary and Conclusions

As a general conclusion, in this chapter, a MIMO-OFDM multi-user system based on a joint beamforming approach has been proposed. The objective has been the joint design of the beamvectors associated to all the established communications or links, taking as the optimization criterion the minimization of the total transmit power subject to maximum mean BER and individual transmit power constraints. It has been shown that this problem is not convex and, therefore, the application of the SA technique has been proposed, in addition to the classical GS and AM methods. The main advantage of the SA algorithm has been shown to be ability to find

Table 4.1: Powers and BER's for the SA, GS, and AM algorithms in scenario 2.

	MT 1 power	**MT 2 power**	**MT 3 power**	**Total power**
SA	10.2 W	8.7 W	54.5 W	73.4 W
GS	9.6 W	8.0 W	46.5 W	64.1 W
AM-MAXMIN	8.3 W	6.8 W	63.6 W	78.7 W
AM-GEOM	9.1 W	7.9 W	65.5 W	82.5 W

	MT 1 BER	**MT 2 BER**	**MT 3 BER**
SA	10^{-3}	10^{-3}	10^{-2}
GS	$1.01 \cdot 10^{-3}$	$1.01 \cdot 10^{-3}$	$1.42 \cdot 10^{-2}$
AM-MAXMIN	10^{-3}	10^{-3}	10^{-2}
AM-GEOM	10^{-3}	10^{-3}	10^{-2}

the global optimum solution. On the other hand, the other presented classical techniques GS and AM may have problems related to the convergence speed and the fact that local suboptimum designs may be obtained. Besides, GS and AM cannot always include every kind of constraint, whereas in SA, this can be done easily by using adequate penalty functions.

Although SA has been shown to be a powerful tool to cope with the optimization of non-convex problems, such as the one presented in this paper, there exist other heuristic approaches that should be also considered as possible strategies. Among these techniques, some examples can be given such as the *genetic algorithms* [Gol88] or *taboo search* [Glo89] approaches. In both cases, the techniques are based on a random generation of possible solutions, such as in the SA algorithm, and are also able to find the optimum solution, even if the problem is not convex. The main difference of genetic algorithms and taboo search when compared to SA is that they have to transform the solution space, i.e., the set of possible transmit beamvectors, into a space composed of "bits" by means of an encoding process. Once this transformation has been performed, the optimization problem is solved in this new transformed solution space. Finally, the solution in terms of the transmit beamvectors is found by transforming or decoding the solution in the coded space. Further work is to be done on the application of these techniques in order to evaluate whether the computational load of the optimization problem can be decreased while still guaranteeing that the global optimum design is found.

4.A Appendix: Derivation of the Gradient of the Effective Error Probability

The objective of this appendix is to deduce the expression of $\nabla_{\mathbf{b}_k^{(u)H}} P_{\text{e,eff}}^{(j)}$. When the modulation of the subcarriers is BPSK, then the effective error probability for the jth user is

$$P_{\text{e,eff}}^{(j)} \triangleq \frac{1}{N}\sum_{k=0}^{N-1} \mathcal{Q}\left(\sqrt{2\text{SNIR}_k^{(j)}}\right) = \frac{1}{N\sqrt{2\pi}}\sum_{k=0}^{N-1}\int_{\sqrt{2\text{SNIR}_k^{(j)}}}^{\infty} e^{-t^2/2}\text{d}t, \tag{4.35}$$

where

$$\text{SNIR}_k^{(j)} = \mathbf{b}_k^{(j)H}\mathbf{H}_k^{(j)H}\mathbf{R}_k^{(j)-1}\mathbf{H}_k^{(j)}\mathbf{b}_k^{(j)}, \tag{4.36}$$

$$\mathbf{R}_k^{(j)} = \mathbf{\Phi}_n(k) + \sum_{l=1,l\neq j}^{U}\mathbf{H}_k^{(l)}\mathbf{b}_k^{(l)}\mathbf{b}_k^{(l)H}\mathbf{H}_k^{(l)H}. \tag{4.37}$$

Using the previous expression, the gradient is calculated as

$$\nabla_{\mathbf{b}_k^{(u)H}} P_{\text{e,eff}}^{(j)} = \frac{1}{N\sqrt{2\pi}}\nabla_{\mathbf{b}_k^{(u)H}}\int_{\sqrt{2\text{SNIR}_k^{(j)}}}^{\infty} e^{-t^2/2}\text{d}t = -\frac{1}{N\sqrt{2\pi}}e^{-\text{SNIR}_k^{(j)}}\frac{\nabla_{\mathbf{b}_k^{(u)H}}\text{SNIR}_k^{(j)}}{\sqrt{2\text{SNIR}_k^{(j)}}}. \tag{4.38}$$

The expression of $\nabla_{\mathbf{b}_k^{(u)H}}\text{SNIR}_k^{(j)}$ is different for $j = u$ and $j \neq u$. For the case $j = u$, the following result is obtained:

$$\nabla_{\mathbf{b}_k^{(u)H}}\text{SNIR}_k^{(u)} = \mathbf{H}_k^{(u)H}\mathbf{R}_k^{(u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}. \tag{4.39}$$

Before calculating the expression of the gradient of $\text{SNIR}_k^{(j)}$ with respect to $\mathbf{b}_k^{(u)H}$ when $j \neq u$, let us calculate a convenient expression of $\mathbf{R}_k^{(j)-1}$ using the matrix inversion lemma [Sch91, Kay93, Gol96]:[2]

$$\mathbf{R}_k^{(j)} = \mathbf{R}_k^{(j,u)} + \mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}\mathbf{b}_k^{(u)H}\mathbf{H}_k^{(u)H}, \tag{4.40}$$

$$\mathbf{R}_k^{(j,u)} = \mathbf{\Phi}_n(k) + \sum_{l=1,l\neq j,l\neq u}^{U}\mathbf{H}_k^{(l)}\mathbf{b}_k^{(l)}\mathbf{b}_k^{(l)H}\mathbf{H}_k^{(l)H}, \tag{4.41}$$

$$\mathbf{R}_k^{(j)-1} = \mathbf{R}_k^{(j,u)-1} - \frac{\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}\mathbf{b}_k^{(u)H}\mathbf{H}_k^{(u)H}\mathbf{R}_k^{(j,u)-1}}{1+\mathbf{b}_k^{(u)H}\mathbf{H}_k^{(u)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}}. \tag{4.42}$$

Using the previous expression, the SNIR can be rewritten as

$$\text{SNIR}_k^{(j)} = \mathbf{b}_k^{(j)H}\mathbf{H}_k^{(j)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(j)}\mathbf{b}_k^{(j)} \tag{4.43}$$

$$- \frac{\mathbf{b}_k^{(j)H}\mathbf{H}_k^{(j)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}\mathbf{b}_k^{(u)H}\mathbf{H}_k^{(u)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(j)}\mathbf{b}_k^{(j)}}{1+\mathbf{b}_k^{(u)H}\mathbf{H}_k^{(u)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}}, \tag{4.44}$$

[2] The general expression of the matrix inversion lemma is $(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B}+\mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$, where a simplified version, called the Woodbury's identity, is given by $(\mathbf{R}+\gamma^2\mathbf{cc}^H)^{-1} = \mathbf{R}^{-1} - \frac{\gamma^2}{1+\gamma^2\mathbf{c}^H\mathbf{R}^{-1}\mathbf{c}}\mathbf{R}^{-1}\mathbf{cc}^H\mathbf{R}^{-1}$.

and, consequently, the gradient of the SNIR can be finally calculated as

$$\begin{aligned}\nabla_{\mathbf{b}_k^{(u)H}}\mathrm{SNIR}_k^{(j)} &= \frac{\mathbf{H}_k^{(u)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}}{\left(1+\mathbf{b}_k^{(u)H}\mathbf{H}_k^{(u)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}\right)^2}\left|\mathbf{b}_k^{(j)H}\mathbf{H}_k^{(j)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}\right|^2 \\ &\quad -\frac{\mathbf{b}_k^{(j)H}\mathbf{H}_k^{(j)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}}{1+\mathbf{b}_k^{(u)H}\mathbf{H}_k^{(u)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(u)}\mathbf{b}_k^{(u)}}\mathbf{H}_k^{(u)H}\mathbf{R}_k^{(j,u)-1}\mathbf{H}_k^{(j)}\mathbf{b}_k^{(j)}. \qquad (4.45)\end{aligned}$$

Chapter 5

Sources of Imperfections in the CSI and Robustness Strategies

5.1 Introduction

In Chapters 3 and 4, a joint transmitter-receiver design has been proposed for the general case of MIMO communications and single-user and multi-user scenarios. In those cases, a perfect channel knowledge has been assumed to be available at both sides of the system and, therefore, the obtained solutions are optimum. Note, however, that in a realistic deployment, the CSI cannot be assumed to be perfect, i.e., the estimates of the channel and the noise plus interferences correlation matrices will have some error or imperfection, specially at the transmitter side. The main consequence is that the system performance will be always lower than that obtained in the case of an optimum design using a perfect CSI. The main objective of this chapter is to evaluate the different sources of errors and their impact on the system performance, as well as different techniques to obtain robust designs less sensitive to these errors.

The structure of this chapter is explained in the following. The different sources of errors in the CSI are identified and classified in Section 5.2. Afterwards, the impact of these errors on the system performance is analyzed in Section 5.3 for the case of the design presented in Chapter 3 and assuming that the CSI is perfect, despite not being true. Finally, the Bayesian and the maximin robustness strategies are described in Section 5.4. These robustness strategies are tools that can be used to design systems less sensitive to the imperfections in the CSI. The chapter ends with two examples of Bayesian robust designs in Section 5.5, and a summary and some conclusions in Section 5.6.

The results corresponding to the analysis of the degradation of the system performance due to the errors in the CSI have been published in [PI04f, PI03d], whereas the two examples of Bayesian robust designs have been published in [PI03c, PI03b, PI03a].

5.2 Description of the Sources of Imperfections in the CSI

There exist several possible sources of errors in the CSI available during the design stage of the communication system. The error level and the imperfections depend heavily on the physical phenomenon producing the error and also on the way the CSI is acquired at the transmitter and/or the receiver. In this section, the objective is to identify the main sources of errors and describe quantitatively the impact on the quality of the CSI.

The CSI acquisition method can be significantly different depending on the duplexing mode of the communication system. In the case of TDD systems, the channel reciprocity principle states that the channel can be assumed to be equal in the forward and the reverse links, if the variability of the channel is slow enough. This property can be exploited, since the channel estimate to be used in the design of the transmitter in the forward link can be obtained through the received signals in the reverse link. Note, however, that although the channel reciprocity principle is expected to hold, the RF chains including, for example, the high power amplifiers and the mixers, are expected not to be reciprocal, which implies that a calibration mechanism has to be included. In FDD systems, the channel reciprocity principle does not longer hold and, consequently, a feedback channel has to be implemented from the receiver to the transmitter, if the design of the transmitter is to be carried out according to the channel state. By means of the feedback channel, the receiver can send to the transmitter the channel estimate corresponding to the forward link.

The following points show a coarse classification of the different kinds of errors that may be present in the CSI. In the case of FDD systems, all these sources of errors have to be taken into account, whereas in the case of TDD, only the two first reasons have to be considered.

- Channel estimation errors.
- Estimation and feedback delay errors.
- Quantization errors.
- Feedback errors.

In the following, a more detailed analysis is given for each source of error in order to evaluate its impact on the quality of the CSI and, consequently, also on the performance of the system designed according to the imperfect CSI.

For simplicity in the notation and the sake of clarity in the presentation, in this section, only the imperfections in the channel estimate will be analyzed, i.e., the errors in the estimates of the noise plus interferences correlation matrices are not considered, although a similar analysis could also be applied. In the following analysis, the matrix $\mathbf{H}$ represents the actual channel, whereas

$\widehat{\mathbf{H}}$ corresponds to the channel estimate, i.e., the CSI. This is a generic representation of the CSI, which means that, for example, in the case of the systems presented in Chapters 3 and 4, the matrix $\mathbf{H}$ encompasses the channel responses at each subcarrier of the OFDM modulation.

5.2.1 Channel Estimation Errors

Most of the standards for wireless communication systems define training signals, i.e., sequences of pilot symbols that are sent by the transmitter and are also known at the receiver side. Usually, these training signals can be used to obtain the channel estimate. The received signals during the transmission of the pilot symbols contain not only the contribution of the transmitted signal, but also the contribution of the noise plus interferences. The most common approach to calculate the channel estimate is to apply the MSE criterion, which looks for the channel that best fits the received signal in a LS sense. This approach is the optimum one in the sense that it corresponds to the ML estimation, if the noise and interferences are Gaussian and white. Obviously, there will always be errors in the obtained channel estimate due to the presence of the noise and interferences.

The generic LS channel estimation can be carried out as follows. Let $\mathbf{X}_T \in \mathbb{C}^{n_T \times L_{\text{train}}}$ be the matrix containing the n_T training signal sequences corresponding to the n_T transmit antennas with length equal to L_{train} samples. As explained previously, this matrix is assumed to be also known at the receiver. During the transmission of the pilot symbols, the matrix $\mathbf{X}_R$ containing the received signals can be expressed as

$$\mathbf{X}_R = \mathbf{H}\mathbf{X}_T + \mathbf{N} \in \mathbb{C}^{n_R \times L_{\text{train}}}, \tag{5.1}$$

where $\mathbf{N} \in \mathbb{C}^{n_R \times L_{\text{train}}}$ is the matrix representing the contribution of the noise plus interferences at the receiver. The channel estimate can be obtained by looking for the channel matrix that minimizes the MSE, i.e., the LS channel estimate $\widehat{\mathbf{H}}$ can be formulated as

$$\widehat{\mathbf{H}} \triangleq \arg\min_{\mathbf{H}} \|\mathbf{X}_R - \mathbf{H}\mathbf{X}_T\|_F^2, \tag{5.2}$$

which leads to a solution that can be expressed in closed-form as follows, assuming that the matrix $\mathbf{X}_T\mathbf{X}_T^H$ is full-rank, as usual:

$$\widehat{\mathbf{H}} = \mathbf{X}_R\mathbf{X}_T^H\left(\mathbf{X}_T\mathbf{X}_T^H\right)^{-1} \in \mathbb{C}^{n_R \times n_T}. \tag{5.3}$$

This is an unbiased channel estimate, where the error $\boldsymbol{\Delta}$, defined as

$$\boldsymbol{\Delta} \triangleq \widehat{\mathbf{H}} - \mathbf{H}, \tag{5.4}$$

can be expressed as

$$\boldsymbol{\Delta} = \mathbf{N}\mathbf{X}_T^H\left(\mathbf{X}_T\mathbf{X}_T^H\right)^{-1}. \tag{5.5}$$

Usually, the standards define orthogonal training sequences to be transmitted through each antenna, so that the channel estimation for each transmit-receive antenna pair can be decoupled. Under this assumption, the transmission matrix $\mathbf{X}_T$ has the property $\mathbf{X}_T\mathbf{X}_T^H = \alpha\mathbf{I}$, which simplifies the channel estimation, since no matrix inversion has to be calculated:

$$\widehat{\mathbf{H}} = \frac{1}{\alpha}\mathbf{X}_R\mathbf{X}_T^H. \tag{5.6}$$

The constant α is related to the transmit power during the transmission of the training sequences. The relationship between the transmit power P_0^{train} and the constant α is given by $\alpha = P_0^{\text{train}}L_{\text{train}}/n_T$, where $P_0^{\text{train}}L_{\text{train}}$ is the total energy used during the transmission of the pilot symbols through all the antennas in the training period.

Based on the orthogonality of the training signals, the error in the channel estimate can be rewritten as

$$\mathbf{\Delta} = \frac{1}{\alpha}\mathbf{N}\mathbf{X}_T^H. \tag{5.7}$$

If all the components of the noise plus interferences matrix are i.i.d. with zero-mean and variance σ_n^2, then also all the components of $\mathbf{\Delta}$ are also i.i.d. with zero-mean and power σ_n^2/α, i.e., the power of the estimation noise in this case is given by

$$\mathbb{E}\left[\left|[\mathbf{\Delta}]_{i,j}\right|^2\right] = \frac{\sigma_n^2 n_T}{P_0^{\text{train}}L_{\text{train}}}. \tag{5.8}$$

5.2.2 Estimation and Feedback Delay Errors

Usually, there exists a delay between the instants in which the channel estimate is obtained and the transmission. In case that the Doppler spread is low, i.e., if the channel variability is not significant, the errors due to the delay are also low; however, when the Doppler spread cannot be neglected, the errors due to the channel variability have to be taken into account.

In the following, the errors due to the channel variability are analyzed isolated from other sources of errors. Let $\mathbf{H}$ be the actual channel in the instant in which the transmission is performed. Note, however, that the design is carried out according to an outdated CSI denoted by $\widehat{\mathbf{H}}$, which represents the channel response in a previous instant t_{del} seconds before the transmission.

As previously, the error can be expressed as

$$\mathbf{\Delta} \triangleq \widehat{\mathbf{H}} - \mathbf{H}. \tag{5.9}$$

In case that the channel is Rayleigh, i.e., all the components of $\mathbf{H}$ and $\widehat{\mathbf{H}}$ are complex and Gaussian distributed with zero-mean and a variance equal to σ_h^2, and the variation of the channel follows the Jake's model [Ste99], then the cross-correlation between any component of both matrices is given by

$$\mathbb{E}\left[\left[\widehat{\mathbf{H}}\right]_{i,j}\left[\mathbf{H}\right]_{i,j}^*\right] = \sigma_h^2 J_0(2\pi f_D t_{\text{del}}), \tag{5.10}$$

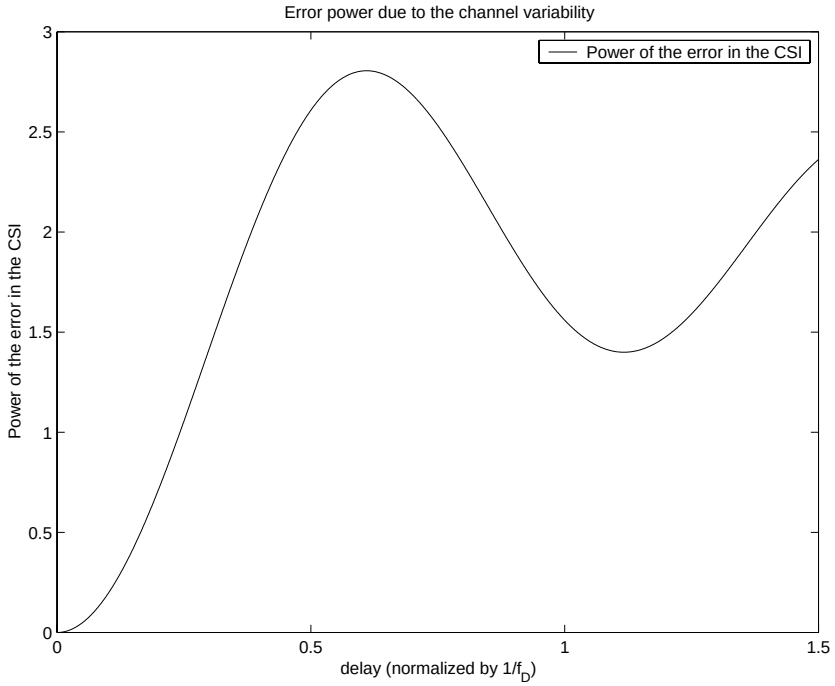


Figure 5.1: Power of the error in the CSI vs. the delay between the channel acquisition and the transmission instants. In the figure, the delay is normalized by the inverse of the maximum Doppler frequency f_D. The variance of the channel is assumed to be normalized, i.e., $\sigma_h^2 = 1$.

where f_D is the maximum Doppler frequency and $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. According to this, the power of the (i,j)th component of the error matrix $\mathbf{\Delta}$ can be finally calculated as follows:

$$\mathbb{E}\left[\left|[\mathbf{\Delta}]_{i,j}\right|^2\right] = 2\sigma_h^2 - 2\sigma_h^2 J_0(2\pi f_D t_{\text{del}}). \tag{5.11}$$

In Figure 5.1, the power of the error in the CSI is shown as a function of the delay between the instant in which the channel is acquired and the transmission. From the figure it is concluded that the delay should be much lower than the inverse of the maximum Doppler frequency to have an acceptable error power. In some papers such as [Cho02b], a channel prediction based on the MMSE criterion is applied to obtain a more accurate estimate of the current channel based on several previous channel measurements, showing that the error power can be decreased significantly.

The delay effects have to be taken into account, specially in FDD systems, where the delay is produced not only by the estimation process itself, but also by the feedback delay due to the transmission of the CSI from the receiver to the transmitter.

5.2.3 Quantization Errors

When the channel state has to be sent from the receiver to the transmitter through a feedback channel, which is expected to be digital, the channel matrix $\mathbf{H}$ has to be quantized, obtaining $\widehat{\mathbf{H}}$. Obviously, the quantization of the channel matrix produces an error which can be expressed as previously, i.e.,

$$\boldsymbol{\Delta} \triangleq \widehat{\mathbf{H}} - \mathbf{H}. \tag{5.12}$$

The simplest technique consists in quantizing each component of the channel matrix independently. Let us assume that all the components of the channel matrix are i.i.d. according to the distribution $[\mathbf{H}]_{i,j} \sim \mathcal{CN}(0, \sigma_h^2)$. Let us also assume that the real and imaginary parts of each component of $\mathbf{H}$, with a variance equal to $\sigma_h^2/2$, are quantized independently using n_b bits. Consequently, the total number of bits used in the quantization is $N_b = 2n_T n_R n_b$. The quantization strategy that minimizes the MSE can be obtained using the rate-distortion theory [Cov91, Cho02b], giving as a result independent errors in each quantization of the real and imaginary parts with a power equal to $\sigma_h^2 2^{-2n_b}/2$, leading to

$$\mathbb{E}\left[\left|[\boldsymbol{\Delta}]_{i,j}\right|^2\right] = \sigma_h^2 2^{-2n_b} = \sigma_h^2 2^{-N_b/n_T n_R}. \tag{5.13}$$

Other quantization strategies could also be applied, such as vector quantization, or differential encoding. The number of possibilities is large and increasing, since currently, much research work is being carried out on this topic. See [Lov04], and references thereon, for a general overview of these issues.

5.2.4 Feedback Errors

When a feedback channel is used, additional errors can arise due to the presence of noise in this channel. The effect of these errors on the quality of the received CSI can be reduced by using appropriate channel coding strategies in the transmission of the quantized channel matrix. Note, however, that an error-free transmission can be never guaranteed and, therefore, the quality of the received CSI will be always lower than the quality of the transmitted CSI. The relationship between the quality of the feedback link and the final quality of the channel estimate available during the design stage is highly nonlinear and has to be calculated by means of simulations that are outside the scope of this dissertation.

5.3 Degradation of the MIMO-OFDM Single-User System Due to Errors in the CSI

The imperfections and errors in the CSI generally imply a decrease of the system performance, specially when the system is designed according to the CSI and without taking into account these errors, i.e., in a non-robust way. There are several works and papers where this performance degradation is studied. In [Cho02b] and [Cho02c], the degradation of a Rayleigh flat fading MISO channel is analyzed, where the transmitter is based on beamforming and designed assuming a perfect CSI. This performance degradation is studied in terms of the mean value of the Chernoff upper-bound on the BER, concluding that the errors produce a reduction in the diversity order and that, when the error level is high, a space-time coding approach is more suitable than a beamforming solution. In [Bha02], a similar analysis is carried out, also for a Rayleigh flat fading MISO channel. There, the degradation in terms of the decrease of the outage mutual information is studied in three different situations depending on the quality of the CSI at the transmitter and the receiver: imperfect CSI at the receiver and no CSI at the transmitter, perfect CSI at the receiver and quantized CSI at the transmitter, and imperfect CSI at the receiver and quantized CSI at the transmitter.

In this section, the objective is to study the performance degradation of a single-user MIMO-OFDM system due to an imperfect CSI at the transmitter. The system is designed assuming that the CSI is perfect, despite not being true, and according to the technique presented in Chapter 3. Before presenting the analysis, some mathematical preliminaries are given related to the theory of the eigenvectors and eigenvalues sensitivity to matrix perturbations, on which the analysis of the performance degradation is based.

5.3.1 Preliminaries: Eigenvector/Eigenvalue Sensitivity to Matrix Perturbations

Definition 1 *Given a square matrix* $\mathbf{A} \in \mathcal{C}^{n\times n}$*, the eigenvector decomposition of this matrix, if it exists, is given by a nonsingular matrix* $\mathbf{X} = \left[\mathbf{x}_1 \cdots \mathbf{x}_n\right] \in \mathcal{C}^{n\times n}$ *(i.e., its inverse matrix* $\mathbf{X}^{-1}$ *exists) and a diagonal matrix* $\mathbf{\Lambda} = \mathrm{diag}\left(\{\lambda_1, \ldots, \lambda_n\}\right) \in \mathcal{C}^{n\times n}$ *such that*

$$\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}, \tag{5.14}$$

$$\lambda_i \mathbf{x}_i = \mathbf{A}\mathbf{x}_i, \qquad i = 1, \ldots, n, \tag{5.15}$$

where $\{\mathbf{x}_i\}$ *and* $\{\lambda_i\}$ *are the eigenvectors and the eigenvalues of* $\mathbf{A}$*, respectively, and* $\mathbf{x}_i$ *is said to be the eigenvector associated to the eigenvalue* λ_i*.*

The perturbation theory studies the variation of the eigenvectors and the eigenvalues of a matrix when the matrix changes. One of the main results is given by the *Bauer-Fike theorem*,

which is formulated as follows:

Theorem 1 *[Gol96, Theor. 7.2.2.] If μ is an eigenvalue of the matrix $\mathbf{A}+\mathbf{E} \in \mathbb{C}^{n\times n}$ and the eigenvector decomposition of $\mathbf{A}$ is given by $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$, where $\mathbf{\Lambda} = \mathrm{diag}\left(\{\lambda_1,\ldots,\lambda_n\}\right) \in \mathbb{C}^{n\times n}$, then*

$$\min_{\lambda\in\lambda(\mathbf{A})} |\lambda-\mu| \leq k_p(\mathbf{X})\|\mathbf{E}\|_p, \tag{5.16}$$

where $\|\cdot\|_p$ denotes any of the p-norms of a matrix and $k_p(\mathbf{X})$ is the p-condition number of the nonsingular matrix $\mathbf{X}$.

The previous theorem provides an upper-bound on the variation of the eigenvalues when a perturbation is applied to the original matrix. In the following, a corollary of the previous theorem is given for the case of Hermitian matrices:

Corollary 1 *If $\mathbf{A} \in \mathbb{C}^{n\times n}$ is an Hermitian matrix ($\mathbf{A} = \mathbf{A}^H$), then there exists an unitary matrix $\mathbf{U}$ ($\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$) and a diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}\left(\{\lambda_1,\ldots,\lambda_n\}\right) \in \mathbb{R}^{n\times n}$ such that $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, corresponding to the eigenvector decomposition of $\mathbf{A}$. If an Hermitian perturbation matrix $\mathbf{E} = \mathbf{E}^H \in \mathbb{C}^{n\times n}$ is applied to $\mathbf{A}$, then, for any eigenvalue μ of $\mathbf{A}+\mathbf{E}$, the following upper-bound holds:*

$$\min_{\lambda\in\lambda(\mathbf{A})} |\lambda-\mu| \leq \|\mathbf{E}\|_2, \tag{5.17}$$

which is obtained by taking $p=2$ and taking into account that $k_2(\mathbf{U}) = 1$ in Theorem 1.

5.3.2 Degradation of the SNIR in the MIMO-OFDM Single-User System

Let us consider the system and signal models corresponding to the MIMO-OFDM single-user system presented in Chapter 3, in which the transmitter and the receiver were designed jointly based on a joint beamforming approach per carrier and according to the knowledge of the channel and the noise plus interferences correlation matrix at both sides of the communication system.

The optimum design of the transmit $\{\mathbf{b}_k\}$ and the receive $\{\mathbf{a}_k\}$ beamvectors according to a perfect knowledge of the channel $\{\mathbf{H}_k\}$ and the noise plus interferences correlation $\{\mathbf{R}_n(k)\}$ matrices is (see §3.2.3 for more details)

$$\mathbf{b}_k = \sqrt{P_k}\mathbf{u}_{\max}\left(\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k\right), \tag{5.18}$$

$$\mathbf{a}_k = \alpha_k\mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{b}_k, \tag{5.19}$$

$$\lambda_{\max}(k) = \lambda_{\max}\left(\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k\right), \tag{5.20}$$

resulting in the following SNIR:

$$\mathrm{SNIR}_k = \mathbf{b}_k^H\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{b}_k = \lambda_{\max}(k)P_k, \tag{5.21}$$

where P_k is the power allocated to the kth subcarrier.

In a realistic system, however, the assumption of having a perfect knowledge of the channel and correlation matrices is not longer true and, therefore, a degradation of the performance is expected. This degradation is now studied obtaining a closed-form expression of an upper-bound on the worst relative SNIR reduction for the case in which the error in the estimates of the matrices is complex white Gaussian noise. In the simulations subsection, the tightness of the upper-bound is shown and evaluated, and additional results corresponding to the degradation in terms of the uncoded BER are also presented.

Let us introduce, as a first step, the error model in the channel and correlation matrices. The estimates of these matrices at the transmitter for each subcarrier are assumed to be

$$\widehat{\mathbf{H}}_k \triangleq \mathbf{H}_k + \mathbf{\Delta}_k, \tag{5.22}$$

$$\widehat{\mathbf{R}}_n(k) \triangleq \mathbf{R}_n(k) + \mathbf{T}_k, \tag{5.23}$$

where the matrices $\mathbf{\Delta}_k$ and $\mathbf{T}_k$ represent the errors in the estimates corresponding to the kth subcarrier. As a result of differential matrix theory [Rog80, Mag99], it can be verified that the inverse of the estimate of the correlation matrix can be approximated as follows, using the first order Taylor expansion of $\widehat{\mathbf{R}}_n^{-1}(k)$:

$$\widehat{\mathbf{R}}_n^{-1}(k) = \mathbf{R}_n^{-1}(k) + \mathbf{S}_k, \tag{5.24}$$

$$\mathbf{S}_k \simeq -\mathbf{R}_n^{-1}(k)\mathbf{T}_k\mathbf{R}_n^{-1}(k), \tag{5.25}$$

where $\mathbf{S}_k$ represents the error in the estimate of the inverse of the correlation matrix. The matrix $\mathbf{T}_k$ is assumed to be Hermitian, i.e., $\mathbf{T}_k = \mathbf{T}_k^H$, considering also that it can be expressed as $\mathbf{T}_k \triangleq \mathbf{B}_k + \mathbf{B}_k^H$, introducing the new matrix $\mathbf{B}_k$. The components of the matrices $\mathbf{\Delta}_k$ and $\mathbf{B}_k$ are modeled as i.i.d. complex and circularly symmetric Gaussian random variables, with zero-mean and variances proportional to σ_H^2 and σ_R^2, respectively. The assumed model for $\mathbf{\Delta}_k$ is known and corresponds to a ML estimation of the channel in which orthogonal training sequences are used at the transmitter (in [Li02], the orthogonal training sequences are defined as cyclicly delayed sequences). The Gaussian model assumed for $\mathbf{T}_k$ is based on the central limit theorem, as explained in [Del01], for example. Besides, and for simplicity reasons, the components of the matrix $\mathbf{B}_k$ are assumed to be independent, although the results obtained in this subsection could be extended to the correlated case.

According to the estimates of the channel and correlation matrices available at the transmitter, the transmit beamvectors are designed as follows (the superscript $'$ is used to indicate that the design is based on the estimates of the channel and correlation matrices instead of the actual matrices):

$$\mathbf{b}_k' = \sqrt{P_k'}\,\mathbf{u}_{\max}\big(\widehat{\mathbf{H}}_k^H\widehat{\mathbf{R}}_n^{-1}(k)\widehat{\mathbf{H}}_k\big), \tag{5.26}$$

$$\lambda_{\max}'(k) = \lambda_{\max}\big(\widehat{\mathbf{H}}_k^H\widehat{\mathbf{R}}_n^{-1}(k)\widehat{\mathbf{H}}_k\big) = \big\|\widehat{\mathbf{H}}_k^H\widehat{\mathbf{R}}_n^{-1}(k)\widehat{\mathbf{H}}_k\big\|_2, \tag{5.27}$$

where, in the following, and for simplicity in the notation, $\lambda'(k)$ is used instead of $\lambda'_{\max}(k)$ and $\lambda(k)$ is used instead of $\lambda_{\max}(k)$.

In case that the MAXMIN criterion is used to distribute the total available transmit power P_0 among the subcarriers, the power allocated to the kth subcarrier P'_k is calculated as

$$P'_k = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\lambda'(l)}} \frac{1}{\lambda'(k)}, \tag{5.28}$$

as shown in (3.42).

The objective is to evaluate the new SNIR at each subcarrier when using the design of the transmitter according to the estimates of the channel and correlation matrices instead of the actual matrices. When calculating the new SNIR, it will be assumed that the receiver has a perfect knowledge of the actual channel and correlation matrices, in addition to the designed transmit beamvectors $\{\mathbf{b}'_k\}$. Consequently, the optimum receiver is given by the whitened matched filter, as deduced in §3.2.3 and shown in (3.17):

$$\mathbf{a}'_k = \alpha_k \mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{b}'_k. \tag{5.29}$$

According to this receiver design, the resulting SNIR at the kth subcarrier at the detection stage assuming errors in the CSI at the transmitter is given by

$$\mathrm{SNIR}'_k = {\mathbf{b}'_k}^H \mathbf{H}_k^H \mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{b}'_k = P'_k \frac{{\mathbf{b}'_k}^H \mathbf{H}_k^H \mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{b}'_k}{{\mathbf{b}'_k}^H\mathbf{b}'_k}, \tag{5.30}$$

which results from using the design of the receiver shown in (5.29) in (3.16).

A first order approximation of $\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k$ can be obtained as a function of the estimates of the channel and correlation matrices, as expressed in the following equation:

$$\begin{aligned}
\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k &= \left(\widehat{\mathbf{H}}_k - \boldsymbol{\Delta}_k\right)^H \left(\widehat{\mathbf{R}}_n^{-1}(k) - \mathbf{S}_k\right)\left(\widehat{\mathbf{H}}_k - \boldsymbol{\Delta}_k\right) && (5.31)\\
&\simeq \widehat{\mathbf{H}}_k^H\widehat{\mathbf{R}}_n^{-1}(k)\widehat{\mathbf{H}}_k - \mathbf{A}_k, && (5.32)\\
\mathbf{A}_k &= \mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\boldsymbol{\Delta}_k + \mathbf{H}_k^H\mathbf{S}_k\mathbf{H}_k + \boldsymbol{\Delta}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k, && (5.33)
\end{aligned}$$

where $\mathbf{A}_k$ is an Hermitian matrix. Based on these definitions, the SNIR according to the imperfect CSI can be rewritten as

$$\begin{aligned}
\mathrm{SNIR}'_k &= P'_k \frac{{\mathbf{b}'_k}^H \mathbf{H}_k^H \mathbf{R}_n^{-1}(k)\mathbf{H}_k\mathbf{b}'_k}{{\mathbf{b}'_k}^H\mathbf{b}'_k} \simeq P'_k \left(\frac{{\mathbf{b}'_k}^H \widehat{\mathbf{H}}_k^H \widehat{\mathbf{R}}_n^{-1}(k)\widehat{\mathbf{H}}_k\mathbf{b}'_k}{{\mathbf{b}'_k}^H\mathbf{b}'_k} - \frac{{\mathbf{b}'_k}^H\mathbf{A}_k\mathbf{b}'_k}{{\mathbf{b}'_k}^H\mathbf{b}'_k} \right)\\
&= P'_k\lambda'(k) - P'_k \frac{{\mathbf{b}'_k}^H\mathbf{A}_k\mathbf{b}'_k}{{\mathbf{b}'_k}^H\mathbf{b}'_k} = \frac{P_0}{\sum_{l=0}^{N-1}\frac{1}{\lambda'(l)}} - \frac{P_0}{\sum_{l=0}^{N-1}\frac{1}{\lambda'(l)}} \frac{1}{\lambda'(k)} \frac{{\mathbf{b}'_k}^H\mathbf{A}_k\mathbf{b}'_k}{{\mathbf{b}'_k}^H\mathbf{b}'_k}, \qquad (5.34)
\end{aligned}$$

where the expression of the power allocation corresponding to the MAXMIN strategy (5.28) has been used.

When the errors in the CSI are small when compared to the actual channel and correlation matrices, it is possible to obtain a worst-case bound on the SNIR assuming that the second term in (5.34) is positive and maximum. The maximum value of $\frac{\mathbf{b}_k'^H \mathbf{A}_k \mathbf{b}_k'}{\mathbf{b}_k'^H \mathbf{b}_k'}$ is known to be upper-bounded by $\|\mathbf{A}_k\|_2$ [Gol96], leading to

$$\mathrm{SNIR}_k' \geq \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\lambda'(l)}} - \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\lambda'(l)}} \frac{1}{\lambda'(k)} \|\mathbf{A}_k\|_2. \tag{5.35}$$

Based on Corollary 1 resulting from the Bauer-Fike theorem (see Theorem 1) and taking into account equation (5.32), it is possible to obtain an expression of the interval in which the maximum eigenvalue $\lambda'(k)$ of $\widehat{\mathbf{H}}_k^H \widehat{\mathbf{R}}_n^{-1}(k) \widehat{\mathbf{H}}_k$ is located as a function of the maximum eigenvalue $\lambda(k)$ with no error in the CSI and the 2-norm of the matrix $\mathbf{A}_k$:

$$\lambda'(k) \in \left[\lambda(k) - \delta_k, \lambda(k) + \delta_k\right], \qquad \delta_k = \|\mathbf{A}_k\|_2, \tag{5.36}$$

$$\lambda'(k) = \lambda_{\max}\left(\widehat{\mathbf{H}}_k^H \widehat{\mathbf{R}}_n^{-1}(k) \widehat{\mathbf{H}}_k\right) = \left\|\widehat{\mathbf{H}}_k^H \widehat{\mathbf{R}}_n^{-1}(k) \widehat{\mathbf{H}}_k\right\|_2, \tag{5.37}$$

$$\lambda(k) = \lambda_{\max}\left(\mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k\right) = \left\|\mathbf{H}_k^H \mathbf{R}_n^{-1}(k) \mathbf{H}_k\right\|_2. \tag{5.38}$$

The transmitter allocates the available power P_0 among the subcarriers according to $\lambda'(k)$. According to the MAXMIN criterion (as shown in §3.3.3), the carriers with a lower gain, i.e., with a lower eigenvalue, are given more power than the carriers with a higher gain. Since the objective is to evaluate the maximum degradation of the SNIR at the kth subcarrier, let us assume the following worst-case situation derived from the uncertainty interval (5.36):

$$\lambda'(l) = \lambda(l) - \delta_l, \qquad \forall l \neq k, \tag{5.39}$$

$$\lambda'(k) = \lambda(k) + \delta_k. \tag{5.40}$$

Taking into account that the SNIR for the MAXMIN strategy with no errors in the CSI is the same for all the carriers (as deduced in §3.3.3 and expressed in (3.43)):

$$\mathrm{SNIR}_k^{\mathrm{MAXMIN}} = \mathrm{SNIR}_{\mathrm{MAXMIN}} = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\lambda(l)}}, \qquad \forall k, \tag{5.41}$$

it is possible to deduce the maximum degradation of the SNIR at the kth subcarrier as

$$\mathrm{SNIR}_{\mathrm{MAXMIN}} - \mathrm{SNIR}_k' \leq \frac{P_0}{\sum_{l \neq k} \frac{1}{\lambda(l)} + \frac{1}{\lambda(k)}} - \frac{P_0}{\sum_{l \neq k} \frac{1}{\lambda(l)-\delta_l} + \frac{1}{\lambda(k)+\delta_k}} \tag{5.42}$$

$$+ \frac{P_0}{\sum_{l \neq k} \frac{1}{\lambda(l)-\delta_l} + \frac{1}{\lambda(k)+\delta_k}} \frac{\delta_k}{\lambda(k) + \delta_k} \tag{5.43}$$

$$= \frac{P_0}{\sum_{l \neq k} \frac{1}{\lambda(l)} + \frac{1}{\lambda(k)}} - \frac{P_0}{\sum_{l \neq k} \frac{1}{\lambda(l)-\delta_l} + \frac{1}{\lambda(k)+\delta_k}} \frac{\lambda(k)}{\lambda(k) + \delta_k}. \tag{5.44}$$

At this point, the relative degradation parameter $d(k)$ at the kth subcarrier is defined as the maximum degradation of the SNIR related to the original SNIR assuming a perfect CSI, i.e.,

$$d(k) \triangleq \left.\frac{\mathrm{SNIR}_{\mathrm{MAXMIN}} - \mathrm{SNIR}'_k}{\mathrm{SNIR}_{\mathrm{MAXMIN}}}\right|_{\max} = \frac{\frac{P_0}{\sum_{l\neq k}\frac{1}{\lambda(l)}+\frac{1}{\lambda(k)}} - \frac{P_0}{\sum_{l\neq k}\frac{1}{\lambda(l)-\delta_l}+\frac{1}{\lambda(k)+\delta_k}}\frac{\lambda(k)}{\lambda(k)+\delta_k}}{\frac{P_0}{\sum_l \frac{1}{\lambda(l)}}}. \tag{5.45}$$

A first order approximation of the previous expression can be calculated, leading to (see Appendix 5.A-A for a complete derivation of the first order approximation):

$$d(k) \simeq \frac{\sum_{l\neq k}\frac{\delta_l}{\lambda^2(l)} - \frac{\delta_k}{\lambda^2(k)}}{\sum_l \frac{1}{\lambda(l)}} + \frac{\delta_k}{\lambda(k)} = \frac{1}{\sum_l \frac{1}{\lambda(l)}}\left(\sum_{l\neq k}\frac{\delta_l}{\lambda^2(l)} + \frac{\delta_k}{\lambda(k)}\sum_{l\neq k}\frac{1}{\lambda(l)}\right). \tag{5.46}$$

The objective is to obtain an upper-bound on the mean value of the parameter $d(k)$ averaged over the statistics of the error in the channel and correlation matrices. In order to obtain this expression, it is useful to calculate an upper-bound on $\mathbb{E}[\delta_k]$ as follows:

$$\mathbb{E}[\delta_k] = \mathbb{E}[\|\mathbf{A}_k\|_2] = \mathbb{E}\left[\sqrt{\lambda_{\max}(\mathbf{A}_k^H\mathbf{A}_k)}\right] \leq \mathbb{E}\left[\sqrt{\mathrm{Tr}(\mathbf{A}_k^H\mathbf{A}_k)}\right] \tag{5.47}$$

$$\leq \sqrt{\mathbb{E}\left[\mathrm{Tr}(\mathbf{A}_k^H\mathbf{A}_k)\right]} = \sqrt{\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{A}_k^H\mathbf{A}_k\right]\right)}, \tag{5.48}$$

where the last inequality corresponds to the Jensen's inequality for concave functions (the function $\sqrt{x}$ is concave).[1] Using the previous inequality, the following upper-bound is obtained for $\mathbb{E}[\delta_k]$:

$$\mathbb{E}[\delta_k] \leq \sqrt{\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{A}_k^H\mathbf{A}_k\right]\right)} = \sqrt{2\sigma_H^2 t(k) + 2\sigma_R^2 t^2(k)}, \tag{5.49}$$

where $t(k) = \mathrm{Tr}\left(\mathbf{H}_k^H\mathbf{R}_n^{-2}(k)\mathbf{H}_k\right)$ (see Appendix 5.A-B for a complete proof).

By making use of (5.49), the final value of the upper-bound on the maximum relative SNIR degradation $\mathbb{E}[d(k)]$ averaged over the statistics of the errors in the channel and correlation matrices is given by

$$\begin{aligned}\mathbb{E}[d(k)] \leq & \left(\sum_l \frac{1}{\|\mathbf{H}_l^H\mathbf{R}_n^{-1}(l)\mathbf{H}_l\|_2}\right)^{-1} \\ & \times\left(\sum_{l\neq k}\frac{\sqrt{2\sigma_H^2\,\mathrm{Tr}\left(\mathbf{H}_l^H\mathbf{R}_n^{-2}(l)\mathbf{H}_l\right) + 2\sigma_R^2\,\mathrm{Tr}^2\left(\mathbf{H}_l^H\mathbf{R}_n^{-2}(l)\mathbf{H}_l\right)}}{\|\mathbf{H}_l^H\mathbf{R}_n^{-1}(l)\mathbf{H}_l\|_2^2} + \right. \\ & \left.\frac{\sqrt{2\sigma_H^2\,\mathrm{Tr}\left(\mathbf{H}_k^H\mathbf{R}_n^{-2}(k)\mathbf{H}_k\right) + 2\sigma_R^2\,\mathrm{Tr}^2\left(\mathbf{H}_k^H\mathbf{R}_n^{-2}(k)\mathbf{H}_k\right)}}{\|\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k\|_2}\sum_{l\neq k}\frac{1}{\|\mathbf{H}_l^H\mathbf{R}_n^{-1}(l)\mathbf{H}_l\|_2}\right),\end{aligned} \tag{5.50}$$

which is valid for low error levels in the CSI.

[1] The Jensen's inequality states that, in case that f is a convex function, then $f(\mathbb{E}[\mathbf{x}]) \leq \mathbb{E}[f(\mathbf{x})]$. If the function f is concave, then $f(\mathbb{E}[\mathbf{x}]) \geq \mathbb{E}[f(\mathbf{x})]$ (see [Boy04]).

One of the most important conclusions is that the relative SNIR degradation depends linearly on the standard deviation of the error in the channel σ_H and correlation matrices σ_R, when this error is low.

Although the previous results have been obtained assuming that the MAXMIN power allocation is applied at the transmitter, they can be extended to other power allocation strategies. For example, in case that the HARM power allocation strategy is applied, the power allocated to the kth subcarrier is calculated as (see §3.3.2 and (3.32))

$$P'_k = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\sqrt{\lambda'(l)}}} \frac{1}{\sqrt{\lambda'(k)}}, \tag{5.51}$$

and, therefore, a lower-bound on the SNIR can be calculated in the same way as in the case of the MAXMIN strategy, obtaining

$$\mathrm{SNIR}'_k \geq \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\sqrt{\lambda'(l)}}} \sqrt{\lambda'(k)} - \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\sqrt{\lambda'(l)}}} \frac{1}{\sqrt{\lambda'(k)}} \delta_k. \tag{5.52}$$

As the MAXMIN strategy, the HARM approach allocates more power to the carriers with a lower gain. This can be used to calculate the maximum SNIR degradation at the kth subcarrier, which is obtained by taking the expressions (5.39) and (5.40). Taking into account that the SNIR in case of having a perfect CSI is (see §3.3.2 and (3.33))

$$\mathrm{SNIR}_k^{\mathrm{HARM}} = \frac{P_0}{\sum_{l=0}^{N-1} \frac{1}{\sqrt{\lambda(l)}}} \sqrt{\lambda(k)}, \tag{5.53}$$

the maximum SNIR degradation can be upper-bounded by

$$\begin{aligned} \mathrm{SNIR}_k^{\mathrm{HARM}} - \mathrm{SNIR}'_k \;\leq\; & \frac{P_0}{\sum_{l\neq k} \frac{1}{\sqrt{\lambda(l)}} + \frac{1}{\sqrt{\lambda(k)}}} \sqrt{\lambda(k)} - \frac{P_0}{\sum_{l\neq k} \frac{1}{\sqrt{\lambda(l)-\delta_l}} + \frac{1}{\sqrt{\lambda(k)+\delta_k}}} \frac{\lambda(k)+\delta_k}{\sqrt{\lambda(k)+\delta_k}} \\ & + \frac{P_0}{\sum_{l\neq k} \frac{1}{\sqrt{\lambda(l)-\delta_l}} + \frac{1}{\sqrt{\lambda(k)+\delta_k}}} \frac{\delta_k}{\sqrt{\lambda(k)+\delta_k}} \qquad (5.54) \\ = \; & \frac{P_0}{\sum_{l\neq k} \frac{1}{\sqrt{\lambda(l)}} + \frac{1}{\sqrt{\lambda(k)}}} \sqrt{\lambda(k)} - \frac{P_0}{\sum_{l\neq k} \frac{1}{\sqrt{\lambda(l)-\delta_l}} + \frac{1}{\sqrt{\lambda(k)+\delta_k}}} \frac{\lambda(k)}{\sqrt{\lambda(k)+\delta_k}}. \end{aligned}$$

In the following expression, the degradation parameter $d(k)$ is defined and a first order approximation is given as follows (see Appendix 5.A-C for a complete derivation):

$$d(k) \triangleq \left. \frac{\mathrm{SNIR}_k^{\mathrm{HARM}} - \mathrm{SNIR}'_k}{\mathrm{SNIR}_k^{\mathrm{HARM}}} \right|_{\max} \simeq \frac{1}{2} \frac{\sum_{l\neq k} \frac{\delta_l}{\sqrt{\lambda^3(l)}} - \frac{\delta_k}{\sqrt{\lambda^3(k)}}}{\sum_l \frac{1}{\sqrt{\lambda(l)}}} + \frac{1}{2} \frac{\delta_k}{\lambda(k)}. \tag{5.55}$$

The main conclusion is that, also in this case, the maximum relative degradation depends linearly on δ_k and, therefore, the mean value $\mathbb{E}[d(k)]$ also depends linearly on the standard deviation of the error in the channel σ_H and correlation matrices σ_R, for high estimation SNR.

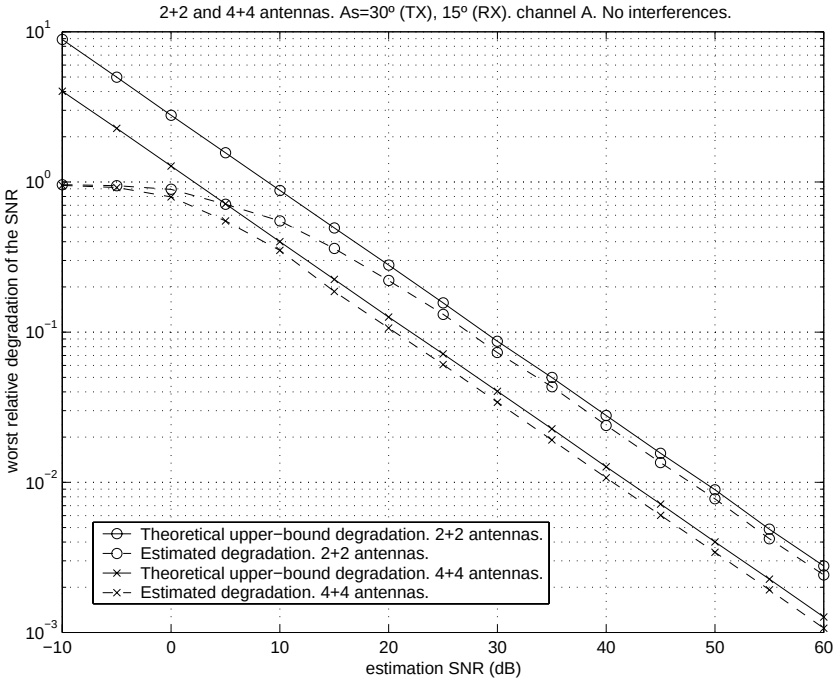


Figure 5.2: Mean value of the worst relative degradation of the SNR vs. the estimation SNR for the MAXMIN power allocation strategy. Comparison between the theoretical upper-bound on the relative degradation and the actual one obtained by simulations. In the simulations, 2+2 and 4+4 antennas are available. The channel power delay profile is that corresponding to model A. The angular spread is 30º at the transmitter and 15º at the receiver. There are no interferences.

5.3.3 Simulation Results

In this subsection, some simulation results are provided to analyze the degradation of the system performance as a consequence of the errors in the CSI. As in §3.4 in Chapter 3, the parameters defined for the physical layer of the HiperLAN/2 system [ETS00] have been taken.

In the performed simulations, the channel power delay profile is that corresponding to model A as described in [ETS98], with a rms delay spread equal to 50 ns, i.e., an indoor scenario. The channel impulse responses are normalized, i.e., $\mathbb{E}\big[\sum_{n=0}^{L-1}|h_{q,p}(n)|^2\big] = 1,\ \forall p,q$ (as in §3.4, the channel is assumed to have L taps). The simulated angular spread at the transmitter is 30º, and at the receiver, 15º. In the considered scenario, no interference has been assumed. According to this, the SNR is defined as in §3.4, i.e., $\text{SNR} = P_0 n_T n_R/\sigma_n^2$, where σ_n^2 is the mean power of the AWGN at the receive antennas.

The estimates of the channel matrices are considered to be imperfect, where the errors in these estimates follow a Gaussian distribution with a variance equal to σ_H^2. Using this parameter, the estimation SNR is defined as $\text{SNR}_{\text{est}} \triangleq 1/\sigma_H^2$. As far as the correlation matrices is concerned,

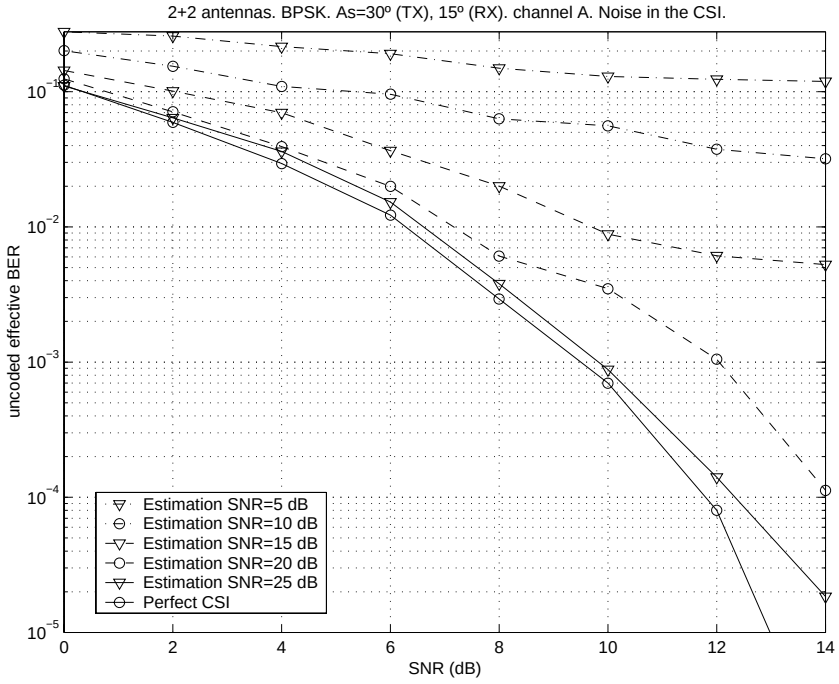


Figure 5.3: Degradation of the system performance due to the errors in the CSI in terms of the uncoded effective BER vs. the SNR for the MAXMIN power allocation strategy. Different values of the estimation SNR have been simulated: 5, 10, 15, 20, and 25 dB, and a scenario with perfect CSI. In the simulations, 2 transmit and 2 receive antennas are available. The channel power delay profile is that corresponding to model A. The angular spread is 30º at the transmitter and 15º at the receiver. There are no interferences.

it is assumed that they are perfectly known at both sides of the communication system. This approach has been taken since the results are more clearly presented and the conclusions that are obtained from the simulations do not change even if interferences are considered in the scenario and the errors in the correlation matrices are also taken into account.

In Figure 5.2, the worst relative SNR degradation is shown averaged over the channel statistics and also over the statistics of the errors in the channel estimate. These results correspond to the application of the MAXMIN power allocation strategy, although the same conclusions can be obtained when using the HARM power distribution. The number of antennas is 2+2 and 4+4. In the figure, the relative degradation obtained theoretically in §5.3.2 (see (5.50)) is shown, in addition to the realistic results obtained from the simulations. As can be seen in the figure, the obtained upper-bound is very tight, specially for high estimation SNR. Besides, the slopes of both curves tend to be $-1/2$ when the estimation SNR increases, as expected from the theoretical analysis, where it was shown that the performance degradation parameter $d(k)$ was proportional to the standard deviation of the estimation noise.

In Figure 5.3, the degradation in terms of the uncoded effective BER is shown for a 2+2

antenna system with no interferences and for different values of the estimation SNR. As can be observed, the degradation is more important as the estimation SNR decreases, specially for values lower than 20-25 dB. For these values of estimation SNR, the worst relative degradation of the SNR is around 0.1, as shown in Figure 5.2.

By means of analysis such as the ones presented in Figures 5.2 and 5.3, important conclusions about the realistic application of MIMO techniques can be obtained. As an example, using the curves in Figure 5.2, a minimum quality threshold in the degradation can be identified, and the corresponding transmit power during the estimation period can be calculated. In case of using a feedback channel to transmit the CSI from the receiver to the transmitter, the channel and correlation matrices have to be quantized. Taking into account the results on the performance degradation, the minimum number of bits to carry out the quantization could be calculated.

5.4 Robustness Strategies

In the previous section, and in some references such as [Cho02b, Cho02c, Zho04], the performance degradation of a communication system is studied, when the transmitter and/or the receiver is designed according to the available imperfect CSI. The designs presented in these references, and also in the previous section, are *non-robust*, in the sense that they consider that the errors in the CSI are negligible, despite not being true. The main conclusion is that this performance degradation increases rapidly with the error level.

In case that there exist errors in the CSI, i.e., in the estimates of the channel and correlation matrices, the optimum design strategy should take into account these errors, obtaining a *robust* solution. Note that, interestingly, the first applications of robust solutions were not for wireless communications, but for control theory (see [Zho96, Has99] and references therein). Indeed, the concepts of signal state space and MIMO were originally used in this area. Afterwards, all these techniques and concepts were extended to other fields due to their potential benefits.

In a communication system, and as commented previously in §5.2, the receiver usually acquires the channel estimate using a training sequence, also known as pilot symbols. Although this is the most usual procedure, note that there also exist some blind algorithms that can obtain a channel estimate without any training sequence (see some examples of blind channel estimation strategies based on the exploitation of the constant modulus property of some signals [Ste99], or higher order statistics [Gia89, Gar94], among others). At the transmitter side, the CSI can be obtained through a feedback channel from the receiver to the transmitter or from previous received signals, exploiting the channel reciprocity principle in a TDD system (see [Ben01] for an overview of different channel estimation strategies).

Different sources of errors in the CSI can be identified depending on the acquisition method.

In case of exploiting the channel reciprocity, the Gaussian noise from the estimation process and the outdated estimate due to the channel variability have to be considered. Note that the errors due to the channel variability can be reduced by using a channel predictor, although also in this case, a residual prediction error still exists (see [Cho02b] and [Zho04], among other references, for the expression of the residual error corresponding to a channel predictor in a multi-antenna system designed under the MMSE criterion). If a feedback channel is used, additional effects arise, such as the quantization of the channel estimate and the errors of the communication through the feedback channel.

According to the way the error in the channel estimate is modeled, the robust techniques can be classified into two families: the *Bayesian (or stochastic)* and the *maximin (or worst-case)* approaches [Boy04]. In the Bayesian philosophy, the statistics of the error are assumed to be known and a stochastic measure of the system performance is optimized, such as the mean or outage values. On the other hand, the maximin approach considers that the error belongs to a predefined uncertainty region and the final objective is to look for the design that optimizes the worst system performance for any error in this region.

The Bayesian philosophy has been considered in works such as [Nar98], where a multi-antenna transmitter is designed to maximize the SNR and the mutual information assuming two sources of errors in the CSI: the Gaussian noise from the estimation, and the quantization errors. The minimization of the BER instead of the maximization of the SNR has been addressed in [Wit95]. The more general case of MIMO flat fading channels has been considered in [Jön02], where the proposed transmitter architecture is composed of an OSTBC stage and a matrix performing a linear transformation of the outputs of the OSTBC. This matrix is designed to minimize an upper-bound on the BER assuming Gaussian errors. The same objective has been taken in [Rey03] and [Xia04] for a MIMO frequency selective channels using a multicarrier modulation: in [Rey03], the transmitter and the receiver are based on matrices performing a linear transformation, whereas in [Xia04], the transmitter is composed of the Alamouti's code [Ala98] combined with two beamformers.

Regarding the maximin approach, [Kas85] and [Ver84] provide a general insight using a game theoretic [Osb94] formulation and describing several applications in signal processing. See also [Roc71a] for a reference on the theory of saddle-points and maximin. This approach has been recently used in the classical problem of designing a receive beamformer under mismatches in the presumed model, as in [Vor03], where the errors are assumed to be in the estimated steering vector and to belong to a spherical uncertainty region. This has been afterwards generalized in [Sha03] to embrace uncertainties both in the array response and the correlation matrix. The classical Capon's beamformer [Mon80] has been extended to its robust version in [Lor04], [Li03], and [Sto03], taking generic uncertainty regions and different formulations. In some of these examples, the robustness is obtained by minimizing the output power of the beamformer while

guaranteeing a minimum gain for any direction modeled by the uncertainty region. In [Pal03b], the maximin approach has been applied to design a MIMO system in which the channel is completely unknown at the transmitter and taking as objective the maximization of the mutual information, leading to a solution in which a uniform power allocation is performed among the transmit dimensions. Finally, several applications of this robust approach to multi-user systems with multi-antenna base stations can also be found in [Ben01], [Ben99], and [Big03].

5.4.1 Mathematical Description

A generic formulation can be stated for both the Bayesian and the maximin approaches. Let $\mathbf{H}$ represent the actual channel response (in the case of the signal models presented in Chapters 3 and 4, and also in §5.3, the "channel" $\mathbf{H}$ would embrace not only the channel matrices, but also the correlation matrices). The imperfect CSI, which is represented by $\widehat{\mathbf{H}}$, can be expressed by

$$\widehat{\mathbf{H}} \triangleq \mathbf{H} + \boldsymbol{\Delta}, \tag{5.56}$$

where $\boldsymbol{\Delta}$ is the error in the CSI.

The system performance is usually measured by a cost function f, whose minimization is the objective of the design (usual cost functions are based on the BER, the MSE, or the SNIR, among others). In the following, a generic formulation is given both for the Bayesian and the maximin approaches in order to minimize the cost function.

The Bayesian Approach

The Bayesian philosophy has been applied classically in estimation theory. The main characteristic of this approach is that both the parameters to be estimated and the observations are considered random variables, as opposed to ML, in which the parameters to be estimated are assumed to be deterministic but unknown (see [Kay93]).

The Bayesian approach can also be applied to obtain a robust design. In this case, the error and the actual channel are modeled statistically through their pdf's $p_{\boldsymbol{\Delta}}(\boldsymbol{\Delta})$ and $p_{\mathbf{H}}(\mathbf{H})$, respectively, which are assumed to be known. Note that knowing these pdf's is equivalent to knowing the pdf of the actual channel conditioned to the channel estimate $p_{\mathbf{H}|\widehat{\mathbf{H}}}\left(\mathbf{H}|\widehat{\mathbf{H}}\right)$, which is equal to $p_{\boldsymbol{\Delta}}\left(\widehat{\mathbf{H}}-\mathbf{H}\right)p_{\mathbf{H}}(\mathbf{H})/p_{\widehat{\mathbf{H}}}\left(\widehat{\mathbf{H}}\right)$ by the Bayes rule [Pap91], since $p_{\widehat{\mathbf{H}}|\mathbf{H}}\left(\widehat{\mathbf{H}}|\mathbf{H}\right) = p_{\boldsymbol{\Delta}}\left(\widehat{\mathbf{H}}-\mathbf{H}\right)$ when the actual channel and the error are independent. A possible design strategy consists in the minimization of the mean value of the cost function f. Note, however, that other stochastic measures of the performance could also be used, such as the outage performance. If the criterion of the mean value is adopted, the following expression has to be minimized:

$$\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left[f\left(\mathbf{H},\mathbf{A},\mathbf{B}\right)\right] = \int f\left(\mathbf{H},\mathbf{A},\mathbf{B}\right)p_{\boldsymbol{\Delta}}\left(\widehat{\mathbf{H}}-\mathbf{H}\right)\frac{p_{\mathbf{H}}(\mathbf{H})}{p_{\widehat{\mathbf{H}}}\left(\widehat{\mathbf{H}}\right)}\mathrm{d}\mathbf{H}, \tag{5.57}$$

where $\mathbf{B}$ and $\mathbf{A}$ represent the transmitter and the receiver, respectively. Obviously, the minimization of (5.57) has to be performed subject to a power constraint at the transmitter. Note that, in case of applying this criterion, no guarantee can be given in terms of the instantaneous performance of the system, but only in terms of the mean performance; besides, a full statistical characterization is necessary, however, this is not always possible.

The Maximin Approach

In the maximin approach, instead of modeling the error statistically, it is assumed that it is unknown but belongs to a predefined uncertainty region $\mathcal{R}$, i.e., $\mathbf{\Delta} \in \mathcal{R}$. The objective of the maximin strategy is to look for the design that optimizes the worst performance of the system for any error in the uncertainty region. This worst performance, that has to be minimized subject to the transmit power constraint, can be formulated as

$$\sup_{\mathbf{\Delta} \in \mathcal{R}} f\left(\mathbf{H}, \mathbf{A}, \mathbf{B}\right), \tag{5.58}$$

where, as in the case of the Bayesian formulation, $\mathbf{B}$ and $\mathbf{A}$ represent the transmitter and the receiver, respectively. It is important to remark that, in this case, a full statistical characterization is not necessary. Besides, this approach guarantees a minimum instantaneous performance for any error modeled by the uncertainty region, i.e., when the actual error behaves as expected (in a realistic situation, this will be satisfied with a high probability, declaring an outage otherwise). Note that this guarantee cannot be provided by the Bayesian approach optimizing the average performance. The performance of the maximin designs is directly related to the definition of the uncertainty region, i.e., it is important to define uncertainty regions that represent the physical phenomenon that is producing the imperfections in the channel knowledge; otherwise, the performance of the system may be degraded.

5.5 Some Examples of Bayesian Designs

In this section, two different examples of Bayesian designs are shown. In §5.5.1, this robustness approach is applied to derive a power allocation strategy that is robust to channel uncertainties in a SISO-ODFM system, where the objective is the minimization of the Chernoff upper-bound on the mean BER. Afterwards, in §5.5.2, a MISO channel is considered in a system exploiting a single-carrier modulation. At the transmitter, a bank of FIR filters is designed to maximize the mean received SNR and minimize the MSE.

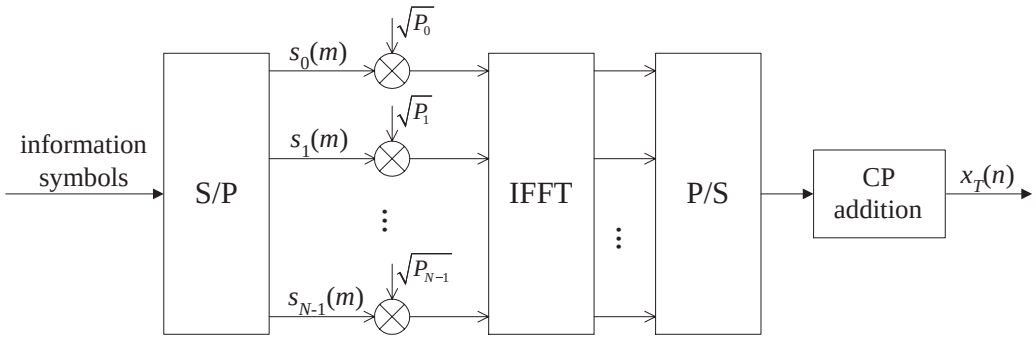


Figure 5.4: Transmitter scheme applying a power allocation among the subcarriers of the OFDM modulation.

5.5.1 Minimum BER Power Allocation in SISO-OFDM Communications

In Chapters 3 and 4, the use of the OFDM modulation and MIMO channels has been considered, i.e., channels with multiple transmit and receive antennas. In the designs obtained in those cases, the channel and correlation matrices have been assumed to be perfectly known at both sides of the communication system. In the following, the system is assumed to have only one transmit and one receive antenna, i.e., a SISO channel is considered. The objective is to derive a power allocation strategy for the OFDM modulation that is robust to the errors in the available channel estimate.

The notation used in the previous chapters can be simplified for the case of a SISO channel combined with a N-carriers OFDM modulation. Let $\{h(n)\}_{n=0}^{L-1}$ be the L-taps time impulse response of the channel that can be expressed in a compact way by using the vector notation $\mathbf{h} \triangleq \left[h(0)\cdots h(L-1)\right]^T \in \mathcal{C}^{L\times 1}$. The complex response of the channel at the kth subcarrier, which is denoted by $H(k)$, is equal to the Fourier transform of the L-taps time response of the channel $h(n)$, i.e.,

$$H(k) \triangleq \sum_{n=0}^{L-1} h(n) e^{-j\frac{2\pi}{N}kn} = \mathbf{f}_k^H \mathbf{h}, \tag{5.59}$$

where $\mathbf{f}_k \triangleq \left[1 \;\; e^{j\frac{2\pi}{N}k} \;\; \cdots \;\; e^{j\frac{2\pi}{N}k(L-1)}\right]^T \in \mathcal{C}^{L\times 1}$.

At the transmitter side, a power allocation is applied among the subcarriers, as shown in Figure 5.4, where P_k is the power allocated to the kth subcarrier. When designing the power distribution, a global transmit power constraint has to be applied, which is formulated as

$$\sum_{k=0}^{N-1} P_k = P_0, \tag{5.60}$$

in the same way as described in Chapter 3 (see (3.24)).

Taking into account this transmission scheme, the received signal sample $y_k(m)$ at the kth

subcarrier during the transmission of the mth OFDM symbol can be written as

$$y_k(m) = H(k)\sqrt{P_k}s_k(m) + n_k(m), \tag{5.61}$$

where $s_k(m)$ is the information symbol transmitted at the kth subcarrier during the mth OFDM symbol with a normalized energy ($\mathbb{E}\left[|s_k(m)|^2\right] = 1$) and $n_k(m)$ denotes the noise plus interferences contribution at the same carrier. In the following, it is assumed that, at the receiver side, there is only AWGN with power σ_n^2 at all the subcarriers, i.e., $\mathbb{E}\left[|n_k(m)|^2\right] = \sigma_n^2$. Based on this signal model, the received SNR at the kth subcarrier can be expressed as

$$\mathrm{SNR}_k = \frac{\mathbb{E}\left[|H(k)\sqrt{P_k}s_k(m)|^2\right]}{\mathbb{E}\left[|n_k(m)|^2\right]} = \frac{|H(k)|^2 P_k}{\sigma_n^2} = \frac{P_k \mathbf{h}^H \mathbf{f}_k \mathbf{f}_k^H \mathbf{h}}{\sigma_n^2}. \tag{5.62}$$

Note that this signal model and the resulting expression of the SNR are obtained as the particularization of the general signal model corresponding to MIMO channels presented in §3.2.2 for the case of a SISO channel.

The effective error probability, assuming that no channel coding is applied to the information symbols, is defined as (see (3.44))

$$P_{\mathrm{e,eff}} \triangleq \frac{\alpha_m}{N} \sum_{k=0}^{N-1} \mathcal{Q}\left(\sqrt{k_m \mathrm{SNR}_k}\right), \tag{5.63}$$

where this expression results from the assumption that the receiver has a perfect channel knowledge and, therefore, the decision boundaries are correctly defined (note that for BPSK and QPSK only the channel phase is necessary).

The direct minimization of the effective error probability is quite complicated and, consequently, the minimization of the Chernoff upper-bound is proposed, which is a tight approximation for high values of the SNR. This Chernoff upper-bound is given by

$$P_{\mathrm{e,eff}} = \frac{\alpha_m}{N} \sum_{k=0}^{N-1} \mathcal{Q}\left(\sqrt{k_m \mathrm{SNR}_k}\right) \leq \frac{\alpha_m}{N} \sum_{k=0}^{N-1} e^{-\frac{1}{2} k_m \mathrm{SNR}_k} = \frac{\alpha_m}{N} \sum_{k=0}^{N-1} e^{-\frac{k_m}{2\sigma_n^2} P_k \mathbf{h}^H \mathbf{f}_k \mathbf{f}_k^H \mathbf{h}}, \tag{5.64}$$

which is a convex function with respect to the power allocation variables $\{P_k\}$.[2]

In the following, two different techniques are proposed to design the power allocation. First, a non-robust power distribution is deduced assuming that the channel estimate is perfect. Afterwards, a robust solution is obtained taking into account explicitly the errors in the channel estimate under the Bayesian philosophy.

[2]This upper-bound is convex since it is the summation of negative exponential functions, which are also convex (the second-order derivatives are positive: $\frac{\partial^2}{\partial x^2} e^{-x} = e^{-x} > 0$).

Non-Robust Power Allocation

The non-robust power allocation is obtained by assuming that the available channel estimate, which is represented by $\widehat{H}(k)$, is perfect, i.e., $\widehat{H}(k) = H(k)$, despite not being true. Hence, the optimization problem to be solved is formulated as

$$\begin{array}{ll} \underset{\{P_k\}}{\text{minimize}} & \frac{\alpha_m}{N} \sum\limits_{k=0}^{N-1} \exp\left(-\frac{k_m}{2\sigma_n^2} P_k |\widehat{H}(k)|^2\right) \\ \text{subject to} & P_k \geq 0, \qquad\qquad k = 0, \ldots, N-1, \\ & \sum\limits_{k=0}^{N-1} P_k = P_0. \end{array} \tag{5.65}$$

This optimization problem is the same as that solved in §3.3.3, whose solution is given by (3.47). Using that result, the non-robust power allocation is expressed as

$$P_k = \frac{2\sigma_n^2}{k_m} \frac{\max\left\{0, \log\left(|\widehat{H}(k)|^2\right) + \mu\right\}}{|\widehat{H}(k)|^2}, \tag{5.66}$$

where μ is a constant calculated to satisfy the global transmit power constraint (5.60).

Bayesian Robust Power Allocation

When the errors in the channel estimate are not negligible, the system performance can decrease importantly if the non-robust power allocation is used. In this case, the optimum design should take into account the errors explicitly, obtaining a robust solution less sensitive to these errors. In the following, a robust power distribution is obtained under the Bayesian philosophy, i.e., using a pure statistical approach.

Let $\widehat{\mathbf{h}} \in \mathcal{C}^{L\times 1}$ be the vector containing the available channel estimate in the time domain, which is related to the actual channel by

$$\widehat{\mathbf{h}} \triangleq \mathbf{h} + \boldsymbol{\delta}, \tag{5.67}$$

and where $\boldsymbol{\delta}$ represents the error in the channel estimate. The channel estimate in the frequency domain $\widehat{H}(k)$ can be expressed as the Fourier transform of the channel estimate in the time domain, i.e., $\widehat{H}(k) \triangleq \mathbf{f}_k^H \widehat{\mathbf{h}} = H(k) + \mathbf{f}_k^H \boldsymbol{\delta}$. The actual channel $\mathbf{h}$ is assumed to be zero-mean and circularly symmetric Gaussian distributed, with a correlation matrix $\boldsymbol{\Lambda} \triangleq \mathbb{E}\left[\mathbf{h}\mathbf{h}^H\right]$. If the taps of the channel are uncorrelated, the matrix $\boldsymbol{\Lambda}$ is diagonal and its elements describe the power delay profile of the channel. Consequently, the pdf of the channel vector $p_{\mathbf{h}}(\mathbf{h})$ is

$$p_{\mathbf{h}}(\mathbf{h}) = \frac{1}{\pi^L \det(\boldsymbol{\Lambda})} \exp\left(-\mathbf{h}^H \boldsymbol{\Lambda}^{-1} \mathbf{h}\right). \tag{5.68}$$

The error $\boldsymbol{\delta}$ is also assumed to be zero-mean and circularly symmetric Gaussian distributed with correlation matrix $\boldsymbol{\Sigma} \triangleq \mathbb{E}\left[\boldsymbol{\delta}\boldsymbol{\delta}^H\right]$ and independent from the actual channel $\mathbf{h}$; hence, the pdf of

the error is given by

$$p_{\boldsymbol{\delta}}(\boldsymbol{\delta}) = \frac{1}{\pi^L \det(\boldsymbol{\Sigma})} \exp\left(-\boldsymbol{\delta}^H \boldsymbol{\Sigma}^{-1} \boldsymbol{\delta}\right). \tag{5.69}$$

Using the pdf's of the actual channel and the error, the pdf of the actual channel conditioned to the channel estimate $p_{\mathbf{h}|\widehat{\mathbf{h}}}\left(\mathbf{h}|\widehat{\mathbf{h}}\right)$ can be calculated as shown in §5.4.1, obtaining:

$$\begin{aligned} p_{\mathbf{h}|\widehat{\mathbf{h}}}\left(\mathbf{h}|\widehat{\mathbf{h}}\right) &= \frac{1}{p_{\widehat{\mathbf{h}}}\left(\widehat{\mathbf{h}}\right)} p_{\boldsymbol{\delta}}\left(\widehat{\mathbf{h}} - \mathbf{h}\right) p_{\mathbf{h}}(\mathbf{h}) \qquad (5.70) \\ &= \frac{1}{p_{\widehat{\mathbf{h}}}\left(\widehat{\mathbf{h}}\right)} \frac{1}{\pi^L \det(\boldsymbol{\Sigma})} \exp\left(-\left(\widehat{\mathbf{h}} - \mathbf{h}\right)^H \boldsymbol{\Sigma}^{-1} \left(\widehat{\mathbf{h}} - \mathbf{h}\right)\right) \frac{1}{\pi^L \det(\boldsymbol{\Lambda})} \exp\left(-\mathbf{h}^H \boldsymbol{\Lambda}^{-1} \mathbf{h}\right), \end{aligned}$$

where

$$p_{\widehat{\mathbf{h}}}\left(\widehat{\mathbf{h}}\right) = \int p_{\widehat{\mathbf{h}}|\mathbf{h}}\left(\widehat{\mathbf{h}}|\mathbf{h}\right) p_{\mathbf{h}}(\mathbf{h}) \mathrm{d}\mathbf{h} = \int p_{\boldsymbol{\delta}}\left(\widehat{\mathbf{h}} - \mathbf{h}\right) p_{\mathbf{h}}(\mathbf{h}) \mathrm{d}\mathbf{h}. \tag{5.71}$$

The objective is to minimize the following expression related to the Chernoff upper-bound on the effective error probability averaged over the statistics of the actual channel conditioned to the channel estimate:

$$\frac{1}{p_{\widehat{\mathbf{h}}}\left(\widehat{\mathbf{h}}\right)} \int \frac{\alpha_m}{N} \sum_{k=0}^{N-1} e^{-cP_k \mathbf{h}^H \mathbf{f}_k \mathbf{f}_k^H \mathbf{h}} p_{\boldsymbol{\delta}}\left(\widehat{\mathbf{h}} - \mathbf{h}\right) p_{\mathbf{h}}(\mathbf{h}) \mathrm{d}\mathbf{h}, \tag{5.72}$$

where $c = k_m / 2\sigma_n^2$. The expression above is convex with respect to $\{P_k\}$ since it is a linear combination (an integral indexed by $\mathbf{h}$) of the set of convex functions $\frac{\alpha_m}{N} \sum_{k=0}^{N-1} e^{-cP_k \mathbf{h}^H \mathbf{f}_k \mathbf{f}_k^H \mathbf{h}}$. The minimization of the expression above subject to the global transmit power constraint (5.60) is equivalent to the following convex optimization problem (see Appendix 5.B for a complete proof of this equivalence):

$$\begin{array}{lll} \underset{\{P_k\}}{\text{minimize}} & \sum\limits_{k=0}^{N-1} \frac{\exp\left(-c \frac{t_k P_k}{1 + c s_k P_k}\right)}{1 + c s_k P_k} & \\ \text{subject to} & -P_k \leq 0, & k = 0, \ldots, N-1, \\ & \sum\limits_{k=0}^{N-1} P_k = P_0, & \end{array} \tag{5.73}$$

where

$$s_k = \mathbf{f}_k^H \left(\boldsymbol{\Lambda}^{-1} + \boldsymbol{\Sigma}^{-1}\right)^{-1} \mathbf{f}_k, \tag{5.74}$$

$$t_k = \left| \mathbf{f}_k^H \left(\boldsymbol{\Lambda}^{-1} + \boldsymbol{\Sigma}^{-1}\right)^{-1} \boldsymbol{\Sigma}^{-1} \widehat{\mathbf{h}} \right|^2. \tag{5.75}$$

In Appendix 5.C, a numerical algorithm is proposed to calculate the Bayesian robust power allocation based on the KKT conditions for the convex problem (5.73).

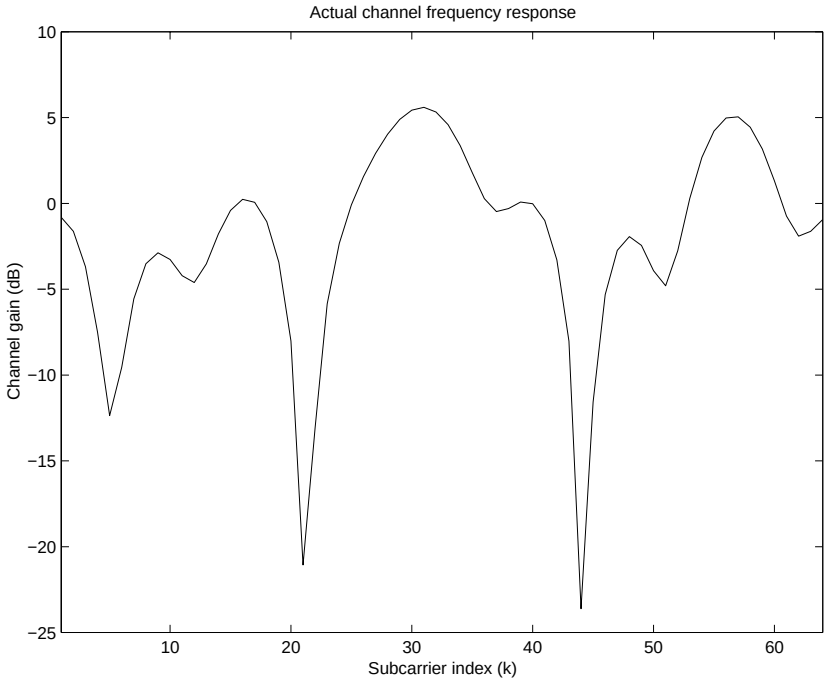
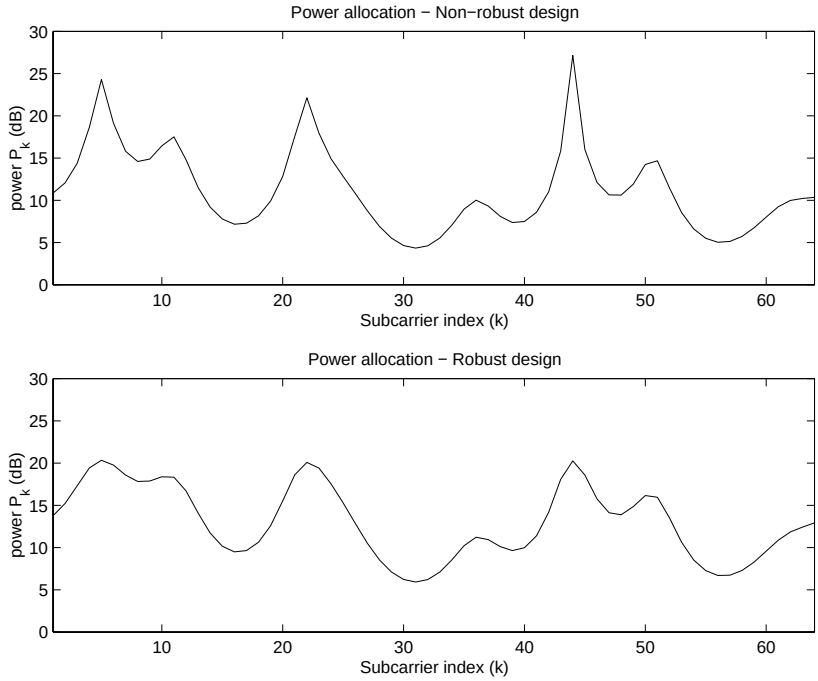


Figure 5.5: Frequency response of the actual channel, i.e., the channel realization, vs. the subcarrier index k.

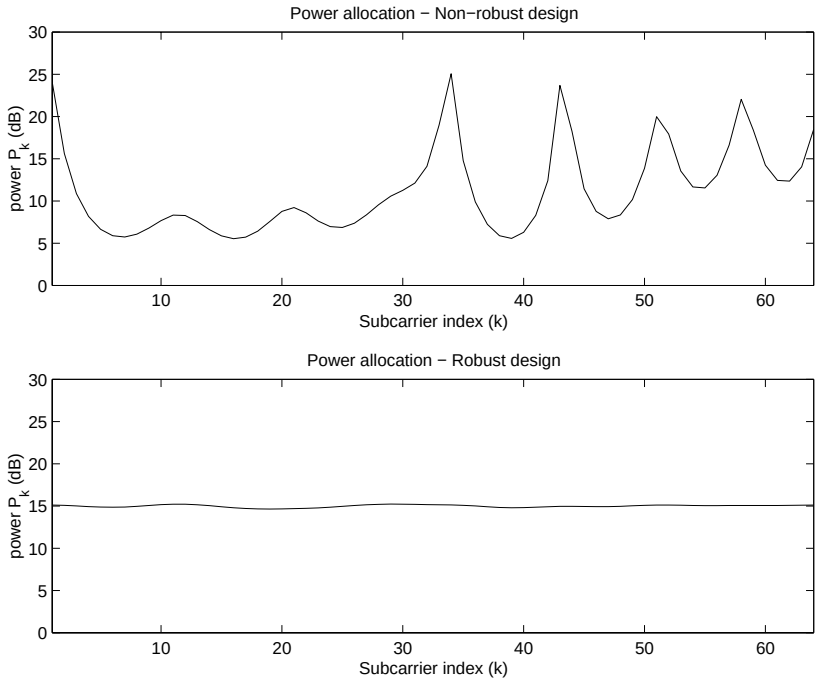
Simulation Results

Some simulation results are now given to show the benefits of the robust approach when compared to the non-robust design. The simulated SISO channel is assumed to have a normalized energy, i.e., $\mathbb{E}\big[\sum_{n=0}^{L-1}|h(n)|^2\big] = 1$. The power delay profile of the channel is exponential and the delay spread normalized to the sampling period is equal to 3. The number of subcarriers of the OFDM modulation is $N = 64$.

Figure 5.5 shows the actual frequency response of a channel realization with 10 taps vs. the subcarrier index k. In Figures 5.6(a) and 5.6(b), the power allocations for such a channel realization are shown assuming a SNR equal to 15 dB and for two different conditions regarding the power of the estimation noise represented by $\mathbf{\Sigma} = 0.01\mathbf{I}$ and $\mathbf{\Sigma} = 0.5\mathbf{I}$, respectively. In the first case, it is seen that the non-robust and the robust techniques behave almost in the same way, since the channel estimate is very similar to the actual channel response, concluding that both techniques tend to be equivalent when having a reliable CSI. In the other case corresponding to a very high estimation noise power, it is seen that the robust approach tends to allocate the same power to all the subcarriers, i.e., an almost uniform power allocation is carried out. The reason for this behavior is that no confidence can be put in the channel estimate and, therefore, it is preferred to transmit the same power in all the transmission bandwidth as if no CSI was



(a)



(b)

Figure 5.6: Robust and non-robust power allocations. (a) High estimation SNR ($\mathbf{\Sigma} = 0.01\mathbf{I}$). (b) Low estimation SNR ($\mathbf{\Sigma} = 0.5\mathbf{I}$).

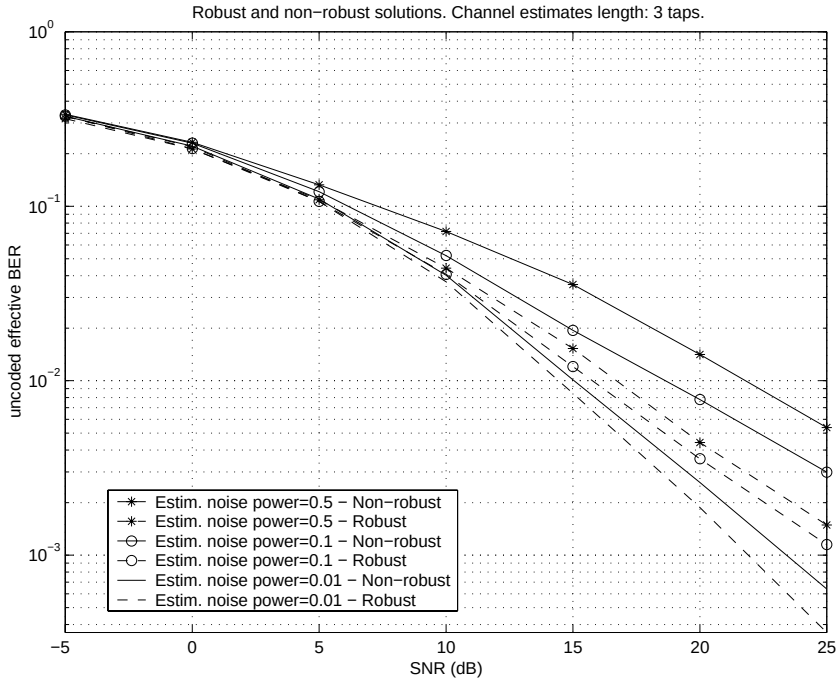


Figure 5.7: Comparison of the performance of the robust and non-robust power allocation techniques in terms of the uncoded effective BER vs. the SNR assuming channels with a normalized energy, a delay spread equal to 3 times the sampling period, and for three different conditions regarding the power of the estimation noise: 0.5, 0.1, and 0.01. The length of the channel estimates is equal to 3 taps.

available at the transmitter.

Figures 5.7 and 5.8 show the performance gain of the robust design when compared to the non-robust power allocation in terms of the uncoded effective BER and assuming BPSK modulated subcarriers. The channel has the same power delay profile as described previously. In the first figure, the channel estimates have 3 taps, whereas in the second example, the assumed number of taps is 10. For each case, three different estimation noise powers are considered: 0.01, 0.1, and 0.5, where this noise is assumed to be white. From the comparison of both figures, it is concluded that the performance gain of the robust power allocation when compared to the non-robust solution is specially significant when the noise power is high, as expected. As far as the non-robust design is concerned, the performance is degraded when the channel estimate has a higher length. This is due to the fact that when there are more taps to be estimated, the global estimation noise has a higher level, even when the additional taps are not significant. The robust approach is less sensitive to this problem, since the power delay profile of the channel is taken into account explicitly in the design of the power allocation (the matrix $\mathbf{\Lambda}$ in (5.68) contains the power delay profile). If the number of taps of the estimate is increased but these additional taps are not significant, the robust solution avoids increasing the estimation noise level by using the

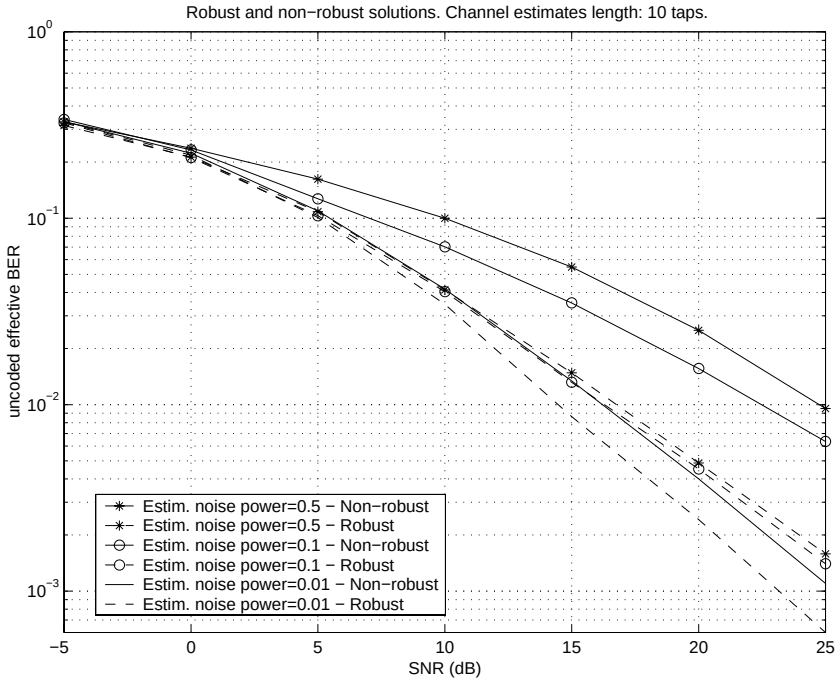


Figure 5.8: Comparison of the performance of the robust and non-robust power allocation techniques in terms of the uncoded effective BER vs. the SNR assuming channels with a normalized energy, a delay spread equal to 3 times the sampling period, and for three different conditions regarding the power of the estimation noise: 0.5, 0.1, and 0.01. The length of the channel estimates is equal to 10 taps.

a-priori information provided by the power delay profile.

5.5.2 FIR Filters Design in MISO Frequency Selective Channels

The robustness strategies mentioned previously, and specifically the Bayesian approach, can be applied to other scenarios and systems where the OFDM modulation is not used. In this section, the robust Bayesian approach is taken to design the transmitter in a MISO channel assuming a single-carrier modulation.

The considered system is composed of a transmitter with n_T antennas and a single-antenna receiver. The transmitter is based on a bank of FIR filters, each one corresponding to a different transmit antenna. These filters are designed according to an imperfect channel estimate, whose error is modeled in a statistical way. At the receiver, two different detectors can be applied: an optimum ML detector based on the application of the Viterbi algorithm, or a symbol-by-symbol detector designed under the MMSE criterion. Figure 5.9 shows the general architecture for such a system, where the channel estimate at the transmitter is obtained through a feedback channel from the receiver to the transmitter, which is a possible solution in FDD systems. Note, however,

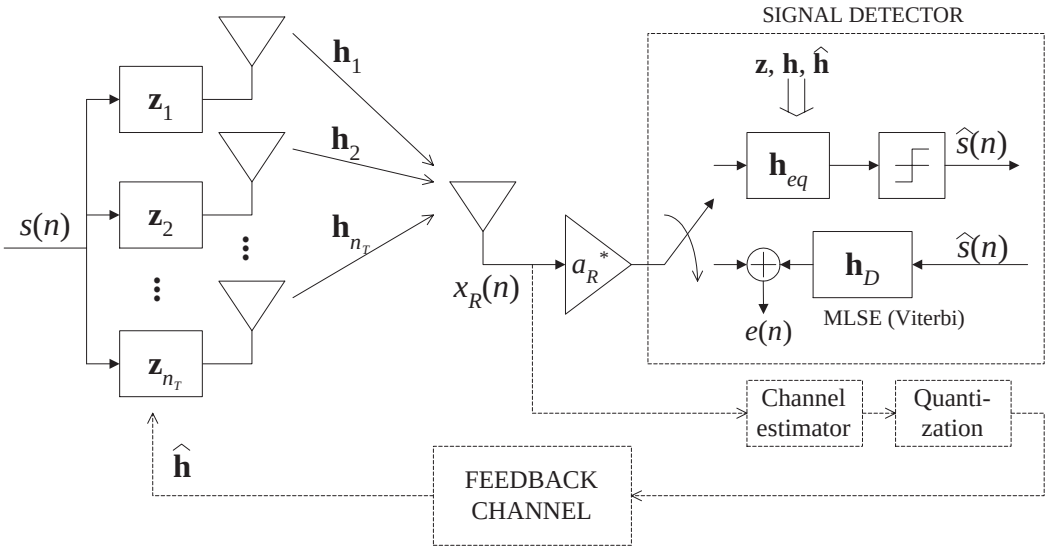


Figure 5.9: General scheme for a transmit diversity system using a single-carrier modulation and based on a bank of transmit FIR filters.

that other ways of acquiring the channel estimate at the transmitter could also be employed, such as the exploitation of the channel reciprocity principle in TDD systems.

Let $\{h_p(n)\}_{n=0}^{L-1}$ be the L-taps time response of the channel between the pth transmit and the receive antennas, which can be represented in a compact way using the vector notation: $\mathbf{h}_p \triangleq \left[h_p(0) \cdots h_p(L-1)\right]^T \in \mathcal{C}^{L\times 1}$. It is possible to collect all the channel responses for all the transmit antennas by defining the vector $\mathbf{h}$ as $\mathbf{h} \triangleq \left[\mathbf{h}_1^T \cdots \mathbf{h}_{n_T}^T\right]^T \in \mathcal{C}^{n_T L\times 1}$. As in the previous example of a Bayesian design (see §5.5.1), in this case, the channel vector is modeled as a complex circularly symmetric Gaussian vector with mean value $\mathbf{m} \triangleq \mathbb{E}[\mathbf{h}]$ and covariance matrix $\mathbf{\Lambda} \triangleq \mathbb{E}\left[(\mathbf{h}-\mathbf{m})(\mathbf{h}-\mathbf{m})^H\right]$, which collects both the spatial correlation and the power delay profile of the channel. Consequently, the pdf of the channel vector can be expressed as

$$p_{\mathbf{h}}(\mathbf{h}) = \frac{1}{\pi^{n_T L}\det(\mathbf{\Lambda})}\exp\left(-(\mathbf{h}-\mathbf{m})^H\mathbf{\Lambda}^{-1}(\mathbf{h}-\mathbf{m})\right). \tag{5.76}$$

At the transmitter, only a channel estimate $\widehat{\mathbf{h}}$ is available, which is related to the actual channel $\mathbf{h}$ by

$$\widehat{\mathbf{h}} = \mathbf{h} + \boldsymbol{\delta}, \tag{5.77}$$

and where $\boldsymbol{\delta}$ represents the error in the estimate, which is assumed to be statistical independent from the actual channel $\mathbf{h}$, and whose pdf is given by $p_{\boldsymbol{\delta}}(\boldsymbol{\delta})$. Consequently, the pdf of the channel estimate conditioned to the actual channel can be expressed as $p_{\widehat{\mathbf{h}}|\mathbf{h}}\left(\widehat{\mathbf{h}}|\mathbf{h}\right) = p_{\boldsymbol{\delta}}\left(\widehat{\mathbf{h}}-\mathbf{h}\right)$. In the following, the error $\boldsymbol{\delta}$ will be assumed to be circularly symmetric Gaussian distributed, with zero-mean and covariance matrix $\mathbf{\Sigma}$, i.e., $\boldsymbol{\delta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$:

$$p_{\boldsymbol{\delta}}(\boldsymbol{\delta}) = \frac{1}{\pi^{n_T L}\det(\mathbf{\Sigma})}\exp\left(-\boldsymbol{\delta}^H\mathbf{\Sigma}^{-1}\boldsymbol{\delta}\right). \tag{5.78}$$

As indicated previously, the objective is to design the transmit FIR filters according to a robust Bayesian approach. Let $\{z_p(n)\}_{n=0}^{M-1}$ be the M-taps time impulse response of the FIR filter corresponding to the pth transmit antenna. This time response can be expressed in a compact way as $\mathbf{z}_p \triangleq \left[z_p(0)\cdots z_p(M-1)\right]^T \in \mathcal{C}^{M\times 1}$, whereas all the time responses for all the filters can be represented jointly as $\mathbf{z} \triangleq \left[\mathbf{z}_1^T\cdots\mathbf{z}_{n_T}^T\right]^T \in \mathcal{C}^{n_T M\times 1}$. These filters are applied to the information symbol stream represented by $s(n)$, where the symbols are assumed to have a normalized energy ($\mathbb{E}\left[|s(n)|^2\right]=1$). Using this notation, the received signal $x_R(n)$ (see Figure 5.9) can be expressed as

$$x_R(n) = s(n) * \sum_{p=1}^{n_T} h_p(n) * z_p(n) + n(n), \tag{5.79}$$

where $n(n)$ is the AWGN with power σ_n^2. This received signal model can be simplified using a matrix-vector notation as

$$x_R(n) = \mathbf{s}^T(n)\sum_{p=1}^{n_T}\mathbf{H}_p\mathbf{z}_p + n(n) = \mathbf{s}^T(n)\mathbf{H}\mathbf{z} + n(n), \tag{5.80}$$

where $\mathbf{s}(n) = \left[s(n)\cdots s(n-L-M+2)\right]^T$, $\mathbf{H} = \left[\mathbf{H}_1\cdots\mathbf{H}_{n_T}\right] \in \mathcal{C}^{(L+M-1)\times n_T M}$, and $\mathbf{H}_p$ is the Toeplitz convolution matrix corresponding to the channel $\mathbf{h}_p$, i.e.,

$$\mathbf{H}_p = \begin{bmatrix} h_p(0) & 0 & \cdots & 0 \\ h_p(1) & h_p(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_p(L-1) & h_p(L-2) & \cdots & 0 \\ 0 & h_p(L-1) & \ddots & h_p(0) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_p(L-1) \end{bmatrix} \in \mathcal{C}^{(L+M-1)\times M}. \tag{5.81}$$

As in the previous examples, a power constraint has to be applied when designing the transmit filters, which, in this case, can be formulated in the time domain as

$$\|\mathbf{z}\|^2 = \mathbf{z}^H\mathbf{z} = \sum_{p=1}^{n_T}\sum_{n=0}^{L-1}|z_p(n)|^2 = P_0. \tag{5.82}$$

Bayesian Robust Transmit Filters

In the following, two different criteria are applied to design the filters: the minimum MSE and the maximum SNR techniques, combined with a symbol-by-symbol and a ML detector, respectively.

Symbol-by-Symbol Detector

When applying a symbol-by-symbol detector at the receiver (see the upper branch at the receiver in Figure 5.9), an adequate design criterion is the MMSE, since it takes into account both

the noise power and the received signal distortion, i.e., the ISI. In case that the CSI is perfect, then the equivalent channel at the receiver resulting from the convolution of the transmit filters with the channel responses, i.e., $\sum_{p=1}^{n_T} z_p(n) * h_p(n)$, would be almost equalized so that the ISI is minimized. As in a realistic system the CSI is not perfect, it is expected that the channel is not equalized by the transmit filters and, therefore, another filter $\mathbf{h}_{eq}$ should be added at the receiver to equalize the residual ISI. Note that this equalizer filter could also be designed under the MMSE criterion. Obviously, this can be performed only when the receiver knows perfectly the equivalent channel, which will be assumed in the following. The gain factor a_R^* at the receiver is responsible for adjusting the dynamic range of the signal to the decision boundaries of the symbol-by-symbol detector.

Using all the notation described previously, the MSE can be expressed as

$$\begin{aligned} \xi(\mathbf{h}, a_R, \mathbf{z}) &\triangleq \mathbb{E}\left[|a_R^* x_R(n) - s(n-l)|^2\right] && (5.83)\\ &= \|a_R^* \mathbf{H}\mathbf{z} - \mathbf{1}_l\|^2 + |a_R|^2 \sigma_n^2 && (5.84)\\ &= |a_R|^2 \mathbf{z}^H \mathbf{H}^H \mathbf{H}\mathbf{z} - a_R^* \mathbf{1}_l^T \mathbf{H}\mathbf{z} - a_R \mathbf{z}^H \mathbf{H}^H \mathbf{1}_l + |a_R|^2 \sigma_n^2 + 1, \end{aligned}$$

where l is the global delay in the signal transmission due to the transmit filters and the channel and $\mathbf{1}_l = [0 \cdots 0 \ 1 \ 0 \cdots 0]^T \in I\!R^{(L+M-1)\times 1}$ is the all-zero vector, except the component 1 in the $(l+1)$th position. The robust Bayesian design is obtained by minimizing the mean value of the MSE, ξ, averaged over the statistics of the actual channel conditioned to the channel estimate, i.e., the following expression has to be minimized with respect to the gain factor a_R at the receiver and the transmit filters vector $\mathbf{z}$:

$$\mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\xi(\mathbf{h}, a_R, \mathbf{z})\right] = |a_R|^2 \mathbf{z}^H \mathbf{X}\mathbf{z} - a_R^* \mathbf{1}_l^T \mathbf{M}\mathbf{z} - a_R \mathbf{z}^H \mathbf{M}^H \mathbf{1}_l + |a_R|^2 \sigma_n^2 + 1, \quad (5.85)$$

where

$$\begin{aligned} \mathbf{M} &= \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\mathbf{H}\right] \in C\!\!\!\!I^{(L+M-1)\times n_T M}, && (5.86)\\ \mathbf{X} &= \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\mathbf{H}^H \mathbf{H}\right] \in C\!\!\!\!I^{n_T M \times n_T M}. && (5.87) \end{aligned}$$

A closed-form expression of the matrices $\mathbf{M}$ and $\mathbf{X}$ is deduced in Appendix 5.D.

Note that the objective function (5.85) to be minimized is convex with respect to the optimization vector $\mathbf{z}$, and also with respect to the variable a_R. Given a concrete value of the vector $\mathbf{z}$, the optimum gain factor a_R should minimize (5.85). This optimum value can be easily obtained by calculating the derivative of $\mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\xi(\mathbf{h}, a_R, \mathbf{z})\right]$ with respect to a_R^* and setting it equal to 0. The solution to this problem is given by

$$\begin{aligned} \frac{\partial \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\xi(\mathbf{h}, a_R, \mathbf{z})\right]}{\partial a_R^*} &= a_R \mathbf{z}^H \mathbf{X}\mathbf{z} - \mathbf{1}_l^T \mathbf{M}\mathbf{z} + a_R \sigma_n^2 = 0, && (5.88)\\ a_R &= \frac{\mathbf{1}_l^T \mathbf{M}\mathbf{z}}{\mathbf{z}^H \mathbf{X}\mathbf{z} + \sigma_n^2}. && (5.89) \end{aligned}$$

On the other hand, given a concrete value of the factor gain a_R, the optimum transmit filters vector $\mathbf{z}$ should minimize (5.85) subject to the power constraint (5.82). This is a convex problem, defined by the following Lagrangian expression (the power constraint has been formulated as the inequality constraint $\mathbf{z}^H\mathbf{z} - P_0 \leq 0$ w.l.o.g., since the optimum solution is attained when the constraint is fulfilled with equality):

$$L(\mathbf{z};\lambda) = |a_R|^2\mathbf{z}^H\mathbf{X}\mathbf{z} - a_R^*\mathbf{1}_l^T\mathbf{M}\mathbf{z} - a_R\mathbf{z}^H\mathbf{M}^H\mathbf{1}_l + |a_R|^2\sigma_n^2 + 1 + \lambda(\mathbf{z}^H\mathbf{z} - P_0). \tag{5.90}$$

The KKT conditions for this problem are:

$$\nabla_{\mathbf{z}^H} L(\mathbf{z};\lambda) = |a_R|^2\mathbf{X}\mathbf{z} - a_R\mathbf{M}^H\mathbf{1}_l + \lambda\mathbf{z} = \mathbf{0}, \tag{5.91}$$

$$\mathbf{z}^H\mathbf{z} \leq P_0, \tag{5.92}$$

$$\lambda \geq 0, \tag{5.93}$$

$$\lambda(\mathbf{z}^H\mathbf{z} - P_0) = 0. \tag{5.94}$$

Finally, using the partial result (5.89) in (5.91), the following result is obtained, as proved in Appendix 5.E:

$$\mathbf{z} = \mu\left(\mathbf{X} + \frac{\sigma_n^2}{P_0}\mathbf{I}\right)^{-1}\mathbf{M}^H\mathbf{1}_l, \tag{5.95}$$

where μ is a constant calculated to satisfy the transmit power constraint $\mathbf{z}^H\mathbf{z} = P_0$ (5.82).

The non-robust design corresponding to the symbol-by-symbol detector is obtained by assuming that the channel estimate is perfect, leading to a solution in which the previous expressions also hold, but using the matrices $\widehat{\mathbf{H}}$ and $\widehat{\mathbf{H}}^H\widehat{\mathbf{H}}$ instead of $\mathbf{M}$ and $\mathbf{X}$, respectively.

Maximum Likelihood Detector

When the computational complexity of the receiver can be increased, it is possible to use other kind of detectors different from the previous one, attaining a better performance. One possibility consists in the application of the ML criterion and the Viterbi algorithm, as shown in the lower branch at the receiver in Figure 5.9. This detector is the optimum one if the channel is known at the receiver and the transmitted symbols are independent. The performance of the ML detector is directly related to the SNR, which can be expressed as

$$\mathrm{SNR}(\mathbf{h},\mathbf{z}) = \frac{1}{\sigma_n^2}\mathbf{z}^H\mathbf{H}^H\mathbf{H}\mathbf{z}, \tag{5.96}$$

whose maximization with respect to the transmit filters $\mathbf{z}$ is the objective of the design.

The robust Bayesian design of the transmitter is obtained by maximizing the mean value of the SNR averaged over the statistics of the actual channel conditioned to the channel estimate, i.e., by maximizing

$$\mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\mathrm{SNR}(\mathbf{h},\mathbf{z})\right] = \frac{1}{\sigma_n^2}\mathbf{z}^H\mathbf{X}\mathbf{z}, \tag{5.97}$$

where the matrix $\mathbf{X}$ is defined as shown in (5.87), and whose expression is deduced in Appendix 5.D. The maximization of (5.97) subject to the transmit power constraint (5.82) is similar to the problem solved in §3.2.3 corresponding to the maximization of the SNIR with respect to the transmit beamvectors in a MIMO system. Using that result, the optimum transmit filters are calculated as follows:

$$\mathbf{z} = \sqrt{P_0}\mathbf{u}_{\max}(\mathbf{X}). \tag{5.98}$$

It can be shown that this design criterion not only maximizes the SNR, but simultaneously minimizes the error power $\sigma_e^2 = \mathbb{E}\big[|e(n)|^2\big]$, where the error is defined as shown in Figure 5.9, and where $\mathbf{h}_D \in \mathcal{C}^{(L+M-1)\times 1}$ is the vector containing the time response of the equivalent channel to be used when applying the Viterbi algorithm and estimating the transmitted symbols, i.e., $\mathbf{h}_D = a_R^*\mathbf{Hz}$ [LH00]. The gain factor a_R is arbitrary and does not affect the performance of the system. a_R could be calculated, for example, so that the mean power at the input of the MLSE is normalized to be equal to 1. It is important to remark that the computational complexity and the memory requirements increase exponentially with the length of $\mathbf{h}_D$.

The non-robust version of this design criterion is obtained by assuming that the channel estimate is perfect, which leads to the calculation of the maximum eigenvector of $\widehat{\mathbf{H}}^H\widehat{\mathbf{H}}$ instead of the matrix $\mathbf{X}$.

Simulation Results

Some simulations are now presented to show the performance gains provided by the robust Bayesian design of the transmit filters when compared to the non-robust approach. In all the simulations, normalized MISO channels are assumed, i.e., $\mathbb{E}\big[\sum_{n=0}^{L-1}|h_p(n)|^2\big] = \mathbb{E}\big[\|\mathbf{h}_p\|^2\big] = 1,\ \forall p$, with an exponential power delay profile and a delay spread equal to 3 sampling periods. The angular spread at the transmitter is 30^o and the symbols are assumed to be BPSK modulated. The assumed length of the channel is $L = 5$ for the case of the simulations using the symbol-by-symbol detector, whereas is $L = 3$ for the case of the Viterbi detector.

In Figure 5.10, some results are shown for the case of the MMSE design. The transmit filters are assumed to have $M = 7$ taps, whereas the filter $\mathbf{h}_{eq}$ at the receiver responsible for equalizing the residual ISI has 10 taps. In all the simulations in which the robust approach has been taken, the error in the channel estimate has been assumed to be Gaussian, i.e., the result (5.95) and the closed-form expressions of the matrices $\mathbf{M}$ and $\mathbf{X}$ deduced in Appendix 5.D have been used. Two different situations have been analyzed. The first one corresponds to an error which is actually Gaussian, whereas in the second case, the error is due to the quantization of the channel time response; therefore, in this last case, the statistical model of the error is not exact. Different number of bits in the quantization of each tap of the channel, and also different powers of the Gaussian error have been simulated, always assuming that it is white, i.e., $\mathbf{\Sigma} = \sigma_\delta^2\mathbf{I}$, where σ_δ^2 is

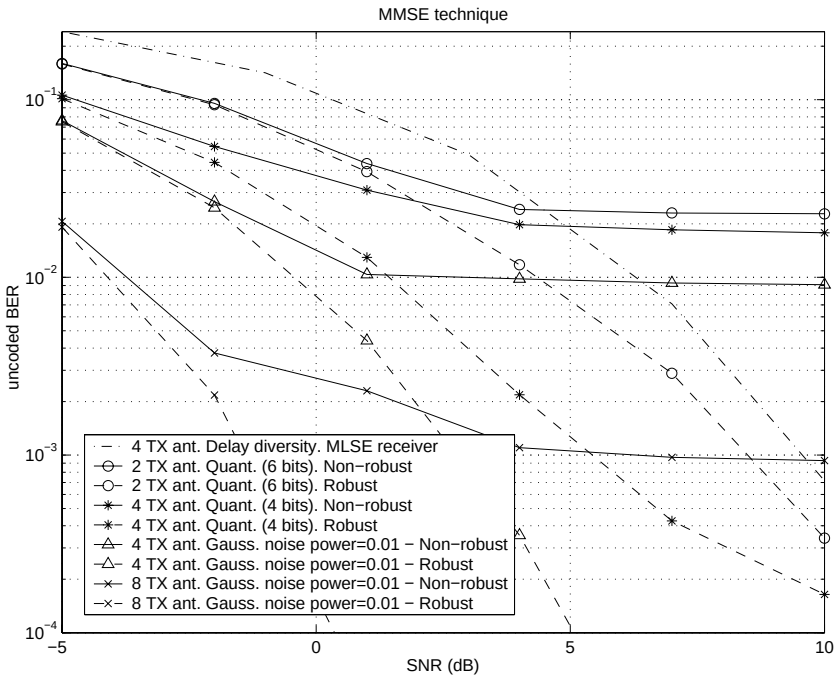


Figure 5.10: Comparison of the performance of delay diversity and the robust and non-robust MMSE transmit filters designs in terms of the uncoded BER vs. the SNR assuming channels with a normalized energy, a delay spread equal to 3 times the sampling period, and for three different conditions regarding the quality of the channel estimate: Gaussian errors with a power equal to 0.01, and quantization errors corresponding to 4 and 6 bits. The assumed length of the channel is equal to 5 taps, whereas the transmit and the residual ISI equalizer filters have 7 and 10 taps, respectively. The angular spread at the transmitter is 30^o, whereas the number of transmit antennas is 2, 4, and 8.

the power of the estimation noise. As can be seen in the figure, the non-robust solution, which does not take into account the error in the CSI and assumes that the channel estimate is perfect, is not able to decrease the BER although the SNR is increased, whereas in the case of the robust Bayesian design, the BER decreases as the SNR increases. It can also be concluded that, even in the case of quantization errors, where the statistical model is not accurate, the Bayesian design is also able to improve the performance when compared to the non-robust design. In Figure 5.10, *delay diversity* (see [Ses93] for an example) represents a simple linear coding approach, so that each transmit filter applies a different delay to the input sequence. The performance of this technique has been included as a benchmark for comparison with a design that does not use any channel knowledge at the transmitter.

In Figure 5.11, the equivalent results are presented for the MLSE technique assuming 4 transmit antennas and a Gaussian estimation error. The transmit filters have only one tap, i.e., a narrowband transmit beamvector is implemented. The results corresponding to the robust and non-robust approaches have been compared with the case of having only one antenna, and also

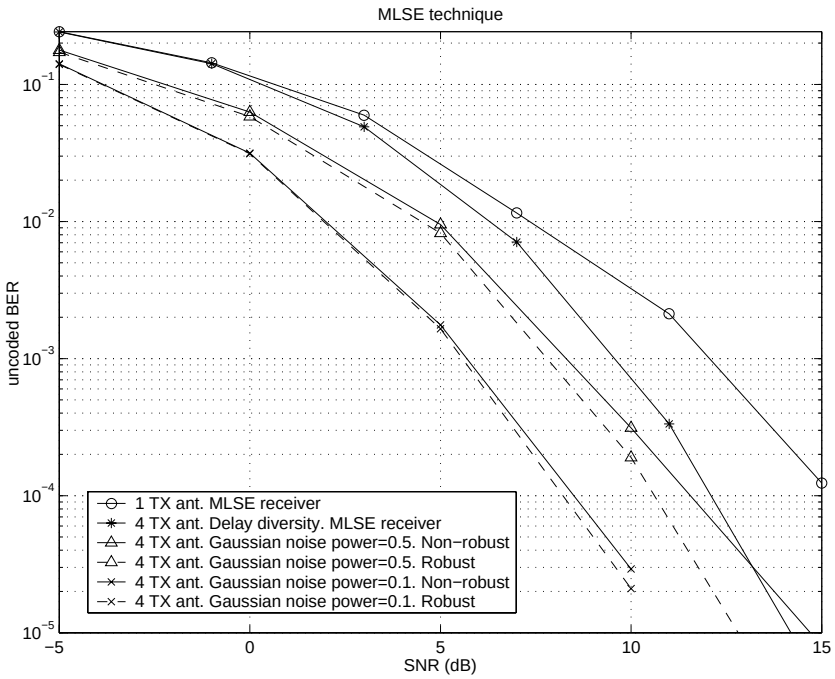


Figure 5.11: Comparison of the performance of delay diversity, a one transmit antenna system, and the robust and non-robust maximum SNR transmit narrowband beamvector designs in terms of the uncoded BER vs. the SNR. The detection technique is based on the Vitberi algorithm and the ML criterion. The channels assumed in the simulations have a normalized energy, a delay spread equal to 3 times the sampling period. Two different powers of the Gaussian errors in the channel estimate have been considered: 0.5 and 0.1. The assumed length of the channel is equal to 3 taps. The angular spread at the transmitter is 30^o, whereas the number of transmit antennas is 4.

with the case of the delay diversity scheme. As can be seen in the figure, in this case, the gains obtained from the use of the Bayesian robust design are less important than those obtained in the MMSE solution. The reason is that the Viterbi detector is less sensitive to non-equalized channels than the MMSE strategy, and that the gains in terms of mean SNR obtained by the Bayesian solution are not extremely important and do not have a direct impact on the BER. It can also be concluded that, even in the case where the noise power in the CSI is very high and the quality of the channel estimate is very bad, the solution based on the MLSE detector has a much better performance than the delay diversity scheme.

5.6 Chapter Summary and Conclusions

In this chapter, the effects of the errors and imperfections in the CSI available during the design stage of the communication system have been analyzed. A classification of the different sources of errors has been performed, stating that they may be different depending on the duplexing

mode of the system (TDD or FDD), i.e., depending on whether a feedback channel from the receiver to the transmitter is necessary or not.

The impact of the errors in the CSI on the system performance has also been studied for the same case as the one presented in Chapter 3, i.e., a single-user MIMO-OFDM channel, where the transmitter is designed in a non-robust way according to the available estimates of the channel and correlation matrices assuming that they are perfect, despite not being true. The main conclusion is that the performance can be severely degraded if the errors in the estimates are not taken into account explicitly.

Two different robustness strategies to design the system have been summarized: the Bayesian (or stochastic) and the maximin (or worst-case) approaches. In both cases, the objective is to obtain a robust design of the system taking into account explicitly the errors in the estimates and, consequently, being less sensitive to them. The difference between the two approaches is based on the cost function to be optimized. In the case of the Bayesian solution, the objective is to optimize the mean value of the performance function averaged over the statistics of the channel and the errors, whereas in the maximin approach, the objective is to optimize the worst system performance for any possible error in the estimate.

Finally, two simple examples of Bayesian designs have been shown: a power allocation strategy for SISO-OFDM systems where the objective is the minimization of the Chernoff upper-bound on the error probability, and a MISO system based on a bank of transmit FIR filters where the objective is either the maximization of the mean SNR or the minimization of the MSE. In both cases, the final conclusion is that the robust solution provides a better performance than the non-robust approach, in which the errors are not taken into account in the design.

5.A Appendix: Derivation of the Expression of the Relative SNIR Degradation

A - Derivation of the Expression (5.46) for the MAXMIN Technique

In order to obtain (5.46), first, a first order approximation of the maximum SNIR degradation at the kth subcarrier (5.44) $\Delta\text{SNIR}_k|_{\max} = \text{SNIR}_{\text{MAXMIN}} - \text{SNIR}'_k|_{\max}$ is calculated as follows, taking into account the first order Taylor expansion $\frac{1}{1\pm x} \simeq 1\mp x,\ x \ll 1$, and assuming $\delta_k/\lambda(k) \ll 1$:

$$\Delta\text{SNIR}_k|_{\max} = \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} - \frac{P_0}{\sum_{l\neq k} \frac{1}{\lambda(l)-\delta_l} + \frac{1}{\lambda(k)+\delta_k}} \frac{\lambda(k)}{\lambda(k)+\delta_k} \tag{5.99}$$

$$= \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} - \frac{P_0}{\sum_{l\neq k} \frac{1}{\lambda(l)} \frac{1}{1-\delta_l/\lambda(l)} + \frac{1}{\lambda(k)} \frac{1}{1+\delta_k/\lambda(k)}} \frac{1}{1+\delta_k/\lambda(k)} \tag{5.100}$$

$$\simeq \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} - \frac{P_0}{\sum_{l\neq k} \frac{1}{\lambda(l)}\left(1+\frac{\delta_l}{\lambda(l)}\right) + \frac{1}{\lambda(k)}\left(1-\frac{\delta_k}{\lambda(k)}\right)} \left(1-\frac{\delta_k}{\lambda(k)}\right) \tag{5.101}$$

$$= \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} - \frac{P_0}{\sum_l \frac{1}{\lambda(l)} + \left(\sum_{l\neq k} \frac{\delta_l}{\lambda^2(l)} - \frac{\delta_k}{\lambda^2(k)}\right)} \left(1-\frac{\delta_k}{\lambda(k)}\right) \tag{5.102}$$

$$= \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} - \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} \frac{1}{1+\left(\sum_{l\neq k} \frac{\delta_l}{\lambda^2(l)} - \frac{\delta_k}{\lambda^2(k)}\right) / \sum_l \frac{1}{\lambda(l)}} \left(1-\frac{\delta_k}{\lambda(k)}\right) \tag{5.103}$$

$$\simeq \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} - \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} \left(1 - \frac{\sum_{l\neq k} \frac{\delta_l}{\lambda^2(l)} - \frac{\delta_k}{\lambda^2(k)}}{\sum_l \frac{1}{\lambda(l)}}\right) \left(1-\frac{\delta_k}{\lambda(k)}\right) \tag{5.104}$$

$$= \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} \frac{\delta_k}{\lambda(k)} + \frac{P_0}{\left(\sum_l \frac{1}{\lambda(l)}\right)^2} \left(\sum_{l\neq k} \frac{\delta_l}{\lambda^2(l)} - \frac{\delta_k}{\lambda^2(k)}\right) \left(1-\frac{\delta_k}{\lambda(k)}\right) \tag{5.105}$$

$$\simeq \frac{P_0}{\sum_l \frac{1}{\lambda(l)}} \frac{\delta_k}{\lambda(k)} + \frac{P_0}{\left(\sum_l \frac{1}{\lambda(l)}\right)^2} \left(\sum_{l\neq k} \frac{\delta_l}{\lambda^2(l)} - \frac{\delta_k}{\lambda^2(k)}\right). \tag{5.106}$$

Using the previous approximation, the expression (5.46) can be finally obtained as

$$d(k) \triangleq \frac{\Delta\text{SNIR}_k|_{\max}}{\text{SNIR}_{\text{MAXMIN}}} \simeq \frac{\frac{P_0}{\sum_l \frac{1}{\lambda(l)}} \frac{\delta_k}{\lambda(k)} + \frac{P_0}{\left(\sum_l \frac{1}{\lambda(l)}\right)^2} \left(\sum_{l\neq k} \frac{\delta_l}{\lambda^2(l)} - \frac{\delta_k}{\lambda^2(k)}\right)}{\frac{P_0}{\sum_l \frac{1}{\lambda(l)}}} \tag{5.107}$$

$$= \frac{\delta_k}{\lambda(k)} + \frac{\sum_{l\neq k} \frac{\delta_l}{\lambda^2(l)} - \frac{\delta_k}{\lambda^2(k)}}{\sum_l \frac{1}{\lambda(l)}} \tag{5.108}$$

$$= \frac{1}{\sum_l \frac{1}{\lambda(l)}} \left(\sum_{l\neq k} \frac{\delta_l}{\lambda^2(l)} + \frac{\delta_k}{\lambda(k)} \sum_{l\neq k} \frac{1}{\lambda(l)}\right). \tag{5.109}$$

B - Derivation of the Upper-Bound on $\mathbb{E}[\delta_k]$

As shown in (5.48), $\mathbb{E}[\delta_k] \leq \sqrt{\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{A}_k^H\mathbf{A}_k\right]\right)}$, where $\mathbf{A}_k = \mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{\Delta}_k + \mathbf{H}_k^H\mathbf{S}_k\mathbf{H}_k + \mathbf{\Delta}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k$ and $\mathbf{S}_k \simeq -\mathbf{R}_n^{-1}(k)\left(\mathbf{B}_k + \mathbf{B}_k^H\right)\mathbf{R}_n^{-1}(k)$. In order to evaluate $\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{A}_k^H\mathbf{A}_k\right]\right)$, the following assumptions are used regarding the statistical distribution of the errors in the channel and correlation matrices:

$$\mathbb{E}\left[\mathbf{\Delta}_k\mathbf{\Delta}_k^H\right] = \sigma_H^2\mathbf{I}, \tag{5.110}$$

$$\mathbb{E}\left[\,\mathrm{vec}(\mathbf{\Delta}_k)(\mathrm{vec}(\mathbf{\Delta}_k))^T\right] = \mathbb{E}\left[\,\mathrm{vec}(\mathbf{\Delta}_k^H)(\mathrm{vec}(\mathbf{\Delta}_k^H))^T\right] = \mathbf{0}, \tag{5.111}$$

$$\mathbb{E}\left[\,\mathrm{vec}(\mathbf{B}_k)(\mathrm{vec}(\mathbf{B}_k))^H\right] = \sigma_R^2\mathbf{I}, \tag{5.112}$$

$$\mathbb{E}\left[\,\mathrm{vec}(\mathbf{B}_k)(\mathrm{vec}(\mathbf{B}_k))^T\right] = \mathbb{E}\left[\,\mathrm{vec}(\mathbf{B}_k^H)(\mathrm{vec}(\mathbf{B}_k^H))^T\right] = \mathbf{0}, \tag{5.113}$$

which results from taking into account that $\mathrm{vec}(\mathbf{\Delta}_k)$ and $\mathrm{vec}(\mathbf{B}_k)$ are circularly symmetric Gaussian random vectors.

The expression $\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{A}_k^H\mathbf{A}_k\right]\right)$ can be expanded as

$$\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{A}_k^H\mathbf{A}_k\right]\right) = \mathrm{Tr}\left(2\mathbb{E}\left[\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{\Delta}_k\mathbf{\Delta}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k\right]\right) + \mathrm{Tr}\left(\mathbb{E}\left[(\mathbf{H}_k^H\mathbf{S}_k\mathbf{H}_k)^2\right]\right), \tag{5.114}$$

where the cross-products between terms containing $\mathbf{\Delta}_k$ and $\mathbf{B}_k$ have not been included since $\mathbf{\Delta}_k$ and $\mathbf{B}_k$ are assumed to be zero-mean and independent.

The first term in (5.114) can be easily calculated as

$$\mathrm{Tr}\left(2\mathbb{E}\left[\mathbf{H}_k^H\mathbf{R}_n^{-1}(k)\mathbf{\Delta}_k\mathbf{\Delta}_k^H\mathbf{R}_n^{-1}(k)\mathbf{H}_k\right]\right) = 2\sigma_H^2\,\mathrm{Tr}\left(\mathbf{H}_k^H\mathbf{R}_n^{-2}(k)\mathbf{H}_k\right), \tag{5.115}$$

whereas the second term $\mathrm{Tr}\left(\mathbb{E}\left[(\mathbf{H}_k^H\mathbf{S}_k\mathbf{H}_k)^2\right]\right)$ can be expanded as

$$\begin{aligned}
\mathrm{Tr}\left(\mathbb{E}\left[(\mathbf{H}_k^H\mathbf{S}_k\mathbf{H}_k)^2\right]\right) &= \mathbb{E}\left[\mathrm{Tr}\left(2\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}^H\mathbf{R}_n^{-1}\mathbf{H}\right.\right. && (5.116)\\
&\quad +\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}\mathbf{R}_n^{-1}\mathbf{H} && (5.117)\\
&\quad \left.\left.+\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}^H\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}^H\mathbf{R}_n^{-1}\mathbf{H}\right)\right], && (5.118)
\end{aligned}$$

where the dependence with the index k has not been included to facilitate the notation. Taking into account (5.113) and using the matrix identity $\mathrm{Tr}(\mathbf{ABCD}) = \left(\mathrm{vec}(\mathbf{D}^H)\right)^H\left(\mathbf{C}^T\otimes\mathbf{A}\right)\mathrm{vec}(\mathbf{B})$ [Mag99], the following result is obtained:

$$\begin{aligned}
\mathrm{Tr}\left(\mathbb{E}\left[(\mathbf{H}_k^H\mathbf{S}_k\mathbf{H}_k)^2\right]\right) &= 2\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}^H\mathbf{R}_n^{-1}\mathbf{H}\right)\right] && (5.119)\\
&= 2\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{B}^H\right)\right] && (5.120)\\
&= 2\mathbb{E}\left[\left(\mathrm{vec}(\mathbf{B})\right)^H\left(\left(\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\right)^T\otimes\left(\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\right)\right)\mathrm{vec}(\mathbf{B})\right]\\
&= 2\,\mathrm{Tr}\left(\left(\left(\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\right)^T\otimes\left(\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\right)\right)\mathbb{E}\left[\,\mathrm{vec}(\mathbf{B})(\mathrm{vec}(\mathbf{B}))^H\right]\right)\\
&= 2\sigma_R^2\,\mathrm{Tr}^2\left(\mathbf{R}_n^{-1}\mathbf{H}\mathbf{H}^H\mathbf{R}_n^{-1}\right) && (5.121)\\
&= 2\sigma_R^2\,\mathrm{Tr}^2\left(\mathbf{H}^H\mathbf{R}_n^{-2}\mathbf{H}\right). && (5.122)
\end{aligned}$$

Collecting these partial results, the upper-bound on $\mathbb{E}[\delta_k]$ can be finally written as

$$\mathbb{E}[\delta_k] \leq \sqrt{2\sigma_H^2 \,\mathrm{Tr}\left(\mathbf{H}_k^H \mathbf{R}_n^{-2}(k)\mathbf{H}_k\right) + 2\sigma_R^2 \,\mathrm{Tr}^2\left(\mathbf{H}_k^H \mathbf{R}_n^{-2}(k)\mathbf{H}_k\right)}. \tag{5.123}$$

C - Derivation of the Expression (5.55) for the HARM Technique

In order to obtain (5.55) for the HARM power allocation strategy, first, a first order approximation of the maximum SNIR degradation at the kth subcarrier (5.54) $\Delta\mathrm{SNIR}_k|_{\max} = \mathrm{SNIR}_k^{\mathrm{HARM}} - \mathrm{SNIR}_k'|_{\max}$ is calculated as follows, taking into account the first order Taylor expansion $\frac{1}{\sqrt{1\pm x}} \simeq 1 \mp \frac{1}{2}x,\ x \ll 1$, and assuming $\delta_k/\lambda(k) \ll 1$:

$$\begin{aligned}
\Delta\mathrm{SNIR}_k|_{\max} &= \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)} - \frac{P_0}{\sum_{l\neq k}\frac{1}{\sqrt{\lambda(l)-\delta_l}} + \frac{1}{\sqrt{\lambda(k)+\delta_k}}}\frac{\lambda(k)}{\sqrt{\lambda(k)+\delta_k}} && (5.124)\\
&= \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)} - \frac{P_0}{\sum_{l\neq k}\frac{1}{\sqrt{\lambda(l)}}\frac{1}{\sqrt{1-\delta_l/\lambda(l)}} + \frac{1}{\sqrt{\lambda(k)}}\frac{1}{\sqrt{1+\delta_k/\lambda(k)}}}\frac{\sqrt{\lambda(k)}}{\sqrt{1+\delta_k/\lambda(k)}}\\
&\simeq \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)}\\
&\quad - \frac{P_0}{\sum_{l\neq k}\frac{1}{\sqrt{\lambda(l)}}\left(1+\frac{1}{2}\frac{\delta_l}{\lambda(l)}\right) + \frac{1}{\sqrt{\lambda(k)}}\left(1-\frac{1}{2}\frac{\delta_k}{\lambda(k)}\right)}\left(1-\frac{1}{2}\frac{\delta_k}{\lambda(k)}\right)\sqrt{\lambda(k)} && (5.125)\\
&= \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)} - \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}} + \frac{1}{2}\left(\sum_{l\neq k}\frac{\delta_l}{\sqrt{\lambda^3(l)}} - \frac{\delta_k}{\sqrt{\lambda^3(k)}}\right)}\left(1-\frac{1}{2}\frac{\delta_k}{\lambda(k)}\right)\sqrt{\lambda(k)}\\
&= \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)}\\
&\quad - \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\frac{1}{1+\frac{1}{2}\left(\sum_{l\neq k}\frac{\delta_l}{\sqrt{\lambda^3(l)}} - \frac{\delta_k}{\sqrt{\lambda^3(k)}}\right)/\sum_l \frac{1}{\sqrt{\lambda(l)}}}\left(1-\frac{1}{2}\frac{\delta_k}{\lambda(k)}\right)\sqrt{\lambda(k)}\\
&\simeq \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)}\\
&\quad - \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\left(1-\frac{1}{2}\frac{\sum_{l\neq k}\frac{\delta_l}{\sqrt{\lambda^3(l)}} - \frac{\delta_k}{\sqrt{\lambda^3(k)}}}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\right)\left(1-\frac{1}{2}\frac{\delta_k}{\lambda(k)}\right)\sqrt{\lambda(k)} && (5.126)\\
&= \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)}\frac{1}{2}\frac{\delta_k}{\lambda(k)}\\
&\quad + \frac{P_0}{\left(\sum_l \frac{1}{\sqrt{\lambda(l)}}\right)^2}\frac{1}{2}\left(\sum_{l\neq k}\frac{\delta_l}{\sqrt{\lambda^3(l)}} - \frac{\delta_k}{\sqrt{\lambda^3(k)}}\right)\left(1-\frac{1}{2}\frac{\delta_k}{\lambda(k)}\right)\sqrt{\lambda(k)} && (5.127)\\
&\simeq \frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)}\frac{1}{2}\frac{\delta_k}{\lambda(k)} + \frac{P_0}{\left(\sum_l \frac{1}{\sqrt{\lambda(l)}}\right)^2}\sqrt{\lambda(k)}\frac{1}{2}\left(\sum_{l\neq k}\frac{\delta_l}{\sqrt{\lambda^3(l)}} - \frac{\delta_k}{\sqrt{\lambda^3(k)}}\right).
\end{aligned}$$

Using the previous approximation, the expression (5.55) can be calculated finally as

$$d(k) \triangleq \frac{\Delta\text{SNIR}_k|_{\max}}{\text{SNIR}_k^{\text{HARM}}} \simeq \frac{\frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)}\frac{1}{2}\frac{\delta_k}{\lambda(k)} + \frac{P_0}{\left(\sum_l \frac{1}{\sqrt{\lambda(l)}}\right)^2}\sqrt{\lambda(k)}\frac{1}{2}\left(\sum_{l\neq k}\frac{\delta_l}{\sqrt{\lambda^3(l)}} - \frac{\delta_k}{\sqrt{\lambda^3(k)}}\right)}{\frac{P_0}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}\sqrt{\lambda(k)}}$$

$$= \frac{1}{2}\frac{\delta_k}{\lambda(k)} + \frac{1}{2}\frac{\sum_{l\neq k}\frac{\delta_l}{\sqrt{\lambda^3(l)}} - \frac{\delta_k}{\sqrt{\lambda^3(k)}}}{\sum_l \frac{1}{\sqrt{\lambda(l)}}}. \quad (5.128)$$

5.B Appendix: Proof of the Equivalence of the Convex Problem (5.73)

The expression to be minimized is (5.72)

$$\frac{1}{p_{\widehat{\mathbf{h}}}(\widehat{\mathbf{h}})}\int \frac{\alpha_m}{N}\sum_{k=0}^{N-1} e^{-cP_k\mathbf{h}^H\mathbf{f}_k\mathbf{f}_k^H\mathbf{h}} p_{\boldsymbol{\delta}}(\widehat{\mathbf{h}}-\mathbf{h})p_{\mathbf{h}}(\mathbf{h})\mathrm{d}\mathbf{h}, \quad (5.129)$$

which is convex with respect to $\{P_k\}$ and where

$$p_{\boldsymbol{\delta}}(\widehat{\mathbf{h}}-\mathbf{h})p_{\mathbf{h}}(\mathbf{h}) = \frac{1}{\pi^L\det(\boldsymbol{\Sigma})}e^{-\left(\widehat{\mathbf{h}}-\mathbf{h}\right)^H\boldsymbol{\Sigma}^{-1}\left(\widehat{\mathbf{h}}-\mathbf{h}\right)}\frac{1}{\pi^L\det(\boldsymbol{\Lambda})}e^{-\mathbf{h}^H\boldsymbol{\Lambda}^{-1}\mathbf{h}}. \quad (5.130)$$

The expression (5.129) can be simplified by using the following equality:

$$e^{-(\widehat{\mathbf{h}}-\mathbf{h})^H\boldsymbol{\Sigma}^{-1}(\widehat{\mathbf{h}}-\mathbf{h})}e^{-\mathbf{h}^H\boldsymbol{\Lambda}^{-1}\mathbf{h}}e^{-cP_k\mathbf{h}^H\mathbf{f}_k\mathbf{f}_k^H\mathbf{h}} = e^{-(\mathbf{h}-\mathbf{m}_k)^H\mathbf{C}_k^{-1}(\mathbf{h}-\mathbf{m}_k)}e^{\mathbf{m}_k^H\mathbf{C}_k^{-1}\mathbf{m}_k}e^{-\widehat{\mathbf{h}}^H\boldsymbol{\Sigma}^{-1}\widehat{\mathbf{h}}}, \quad (5.131)$$

where

$$\mathbf{C}_k = \left(\boldsymbol{\Lambda}^{-1}+\boldsymbol{\Sigma}^{-1}+cP_k\mathbf{f}_k\mathbf{f}_k^H\right)^{-1} \in \mathcal{C}^{L\times L}, \quad (5.132)$$

$$\mathbf{m}_k = \mathbf{C}_k\boldsymbol{\Sigma}^{-1}\widehat{\mathbf{h}} \in \mathcal{C}^{L\times 1}. \quad (5.133)$$

According to this equality, the expression (5.129) can be rewritten as

$$\frac{1}{p_{\widehat{\mathbf{h}}}(\widehat{\mathbf{h}})}\frac{\alpha_m}{N}\frac{1}{\pi^L\det(\boldsymbol{\Lambda})}\frac{1}{\pi^L\det(\boldsymbol{\Sigma})}\sum_{k=0}^{N-1}\int e^{-(\mathbf{h}-\mathbf{m}_k)^H\mathbf{C}_k^{-1}(\mathbf{h}-\mathbf{m}_k)}e^{\mathbf{m}_k^H\mathbf{C}_k^{-1}\mathbf{m}_k}e^{-\widehat{\mathbf{h}}^H\boldsymbol{\Sigma}^{-1}\widehat{\mathbf{h}}}\mathrm{d}\mathbf{h}, \quad (5.134)$$

which can be simplified taking into account that

$$\frac{1}{\pi^L\det(\mathbf{C}_k)}\int e^{-(\mathbf{h}-\mathbf{m}_k)^H\mathbf{C}_k^{-1}(\mathbf{h}-\mathbf{m}_k)}\mathrm{d}\mathbf{h} = 1. \quad (5.135)$$

Collecting all these results, the original expression (5.129) can be finally written as

$$\frac{1}{p_{\widehat{\mathbf{h}}}(\widehat{\mathbf{h}})}\frac{\alpha_m}{N}\frac{1}{\pi^L\det(\boldsymbol{\Lambda})\det(\boldsymbol{\Sigma})}e^{-\widehat{\mathbf{h}}^H\boldsymbol{\Sigma}^{-1}\widehat{\mathbf{h}}}\sum_{k=0}^{N-1}\det(\mathbf{C}_k)e^{\mathbf{m}_k^H\mathbf{C}_k^{-1}\mathbf{m}_k}, \quad (5.136)$$

and, hence, the objective is to minimize the following convex function with respect to the power allocation variables $\{P_k\}$:

$$\sum_{k=0}^{N-1} \det(\mathbf{C}_k) e^{\mathbf{m}_k^H \mathbf{C}_k^{-1} \mathbf{m}_k}, \tag{5.137}$$

i.e., the following convex optimization problem has to be solved:

$$\begin{array}{lll} \underset{\{P_k\}}{\text{minimize}} & \sum_{k=0}^{N-1} \det(\mathbf{C}_k) \exp\left(\mathbf{m}_k^H \mathbf{C}_k^{-1} \mathbf{m}_k\right) & \\ \text{subject to} & P_k \geq 0, & k = 0, \ldots, N-1, \\ & \sum_{k=0}^{N-1} P_k = P_0. & \end{array} \tag{5.138}$$

In order to solve the stated optimization problem, the following equalities are used, obtained from the matrix inversion lemma [Gol96]:

$$\mathbf{C}_k = \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} - cP_k \frac{\left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{f}_k \mathbf{f}_k^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1}}{1 + cP_k \mathbf{f}_k^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{f}_k}, \tag{5.139}$$

$$\det(\mathbf{C}_k) = \frac{1}{1 + cP_k \mathbf{f}_k^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{f}_k} \frac{1}{\det\left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)}. \tag{5.140}$$

Using the previous expressions, it is easy to deduce that the problem (5.138) is equivalent to the minimization of the following convex function with respect to $\{P_k\}$ subject to the same constraints listed in (5.138):

$$\sum_{k=0}^{N-1} \frac{\exp\left(-cP_k \frac{\widehat{\mathbf{h}}^H \mathbf{\Sigma}^{-1} \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{f}_k \mathbf{f}_k^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{\Sigma}^{-1} \widehat{\mathbf{h}}}{1 + cP_k \mathbf{f}_k^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{f}_k}\right)}{1 + cP_k \mathbf{f}_k^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{f}_k}, \tag{5.141}$$

where the factors of the objective function in (5.138) that do not depend on the optimization variables $\{P_k\}$ have been eliminated. Thus, the problem can be finally rewritten as follows:

$$\begin{array}{lll} \underset{\{P_k\}}{\text{minimize}} & \sum_{k=0}^{N-1} \frac{\exp\left(-c \frac{t_k P_k}{1 + c s_k P_k}\right)}{1 + c s_k P_k} & \\ \text{subject to} & -P_k \leq 0, & k = 0, \ldots, N-1, \\ & \sum_{k=0}^{N-1} P_k - P_0 = 0, & \end{array} \tag{5.142}$$

where

$$s_k = \mathbf{f}_k^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{f}_k, \tag{5.143}$$

$$t_k = \left| \mathbf{f}_k^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{\Sigma}^{-1} \widehat{\mathbf{h}} \right|^2. \tag{5.144}$$

5.C Appendix: Algorithm to Calculate the Bayesian Robust Power Allocation

The convex problem corresponding to the Bayesian robust power allocation, which can be formulated as

$$\begin{array}{lll} \underset{\{P_i\}}{\text{minimize}} & \sum_{i=0}^{N-1} \frac{\exp\left(-c\frac{t_i P_i}{1+cs_i P_i}\right)}{1+cs_i P_i} & \\ \text{subject to} & -P_i \leq 0, & i=0,\ldots,N-1, \\ & \sum_{i=0}^{N-1} P_i - P_0 = 0, & \end{array} \tag{5.145}$$

and where

$$s_i = \mathbf{f}_i^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{f}_i, \tag{5.146}$$

$$t_i = \left|\mathbf{f}_i^H \left(\mathbf{\Lambda}^{-1} + \mathbf{\Sigma}^{-1}\right)^{-1} \mathbf{\Sigma}^{-1} \widehat{\mathbf{h}}\right|^2, \tag{5.147}$$

can be solved by using the KKT conditions. The Lagrangian function corresponding to this optimization problem can be expressed as

$$L(\{P_i\};\boldsymbol{\lambda},\nu) \triangleq \sum_{i=0}^{N-1} \frac{\exp\left(-c\frac{t_i P_i}{1+cs_i P_i}\right)}{1+cs_i P_i} - \sum_{i=0}^{N-1} \lambda_i P_i + \nu \left(\sum_{i=0}^{N-1} P_i - P_0\right), \tag{5.148}$$

and the corresponding KKT conditions are

$$P_k^\star \geq 0, \qquad k=0,\ldots,N-1, \tag{5.149}$$

$$\lambda_k^\star \geq 0, \qquad k=0,\ldots,N-1, \tag{5.150}$$

$$\sum_{k=0}^{N-1} P_k^\star = P_0, \tag{5.151}$$

$$\lambda_k^\star P_k^\star = 0, \qquad k=0,\ldots,N-1, \tag{5.152}$$

$$\frac{\partial L(\{P_i^\star\};\boldsymbol{\lambda}^\star,\nu^\star)}{\partial P_k} = -y_k(P_k^\star)\frac{x_k(P_k^\star)}{z_k(P_k^\star)} - \lambda_k^\star + \nu^\star = 0, \qquad k=0,\ldots,N-1, \tag{5.153}$$

where

$$x_k(P_k) = cs_k + \frac{ct_k}{1+cs_k P_k}, \tag{5.154}$$

$$y_k(P_k) = \exp\left(-c\frac{t_k P_k}{1+cs_k P_k}\right), \tag{5.155}$$

$$z_k(P_k) = (1+cs_k P_k)^2. \tag{5.156}$$

From the KKT conditions it is concluded that

$$x_k(P_k^\star)y_k(P_k^\star) + \lambda_k^\star z_k(P_k^\star) = \nu^\star z_k(P_k^\star). \tag{5.157}$$

The algorithm for solving the stated problem, i.e., for finding the optimum power allocation, is based on the following points:

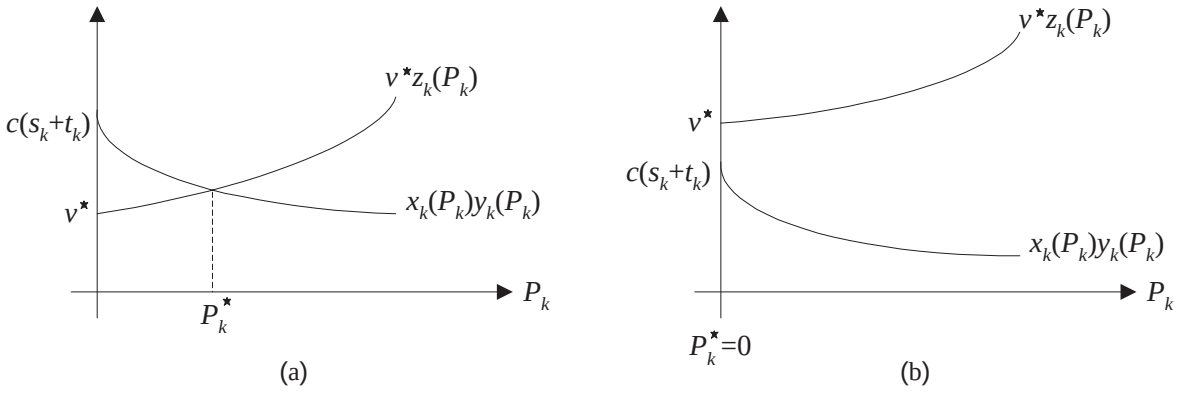


Figure 5.12: Calculation of the optimum robust power allocation. (a) Example of $\nu^\star < c(s_k + t_k)$. (b) Example of $\nu^\star \geq c(s_k + t_k)$.

- For a concrete value of $\nu^\star$, the non-linear equation (5.157) has to be solved for each subcarrier. Two different cases are possible:

  - $P_k^\star$ is positive ($P_k^\star > 0$). In this case, and based on the KKT conditions, $\lambda_k^\star = 0$. Consequently, the following non-linear equation has to be solved: $x_k(P_k^\star)y_k(P_k^\star) = \nu^\star z_k(P_k^\star)$. The solution to this problem is unique and can be found numerically in an efficient way, since $x_k(P_k)y_k(P_k)$ is a differentiable continuous decreasing function of P_k and $z_k(P_k)$ is a differentiable continuous increasing function of P_k. Note that $x_k(0)y_k(0) = c(s_k + t_k)$ and $z_k(0) = 1$, which means that this problem has a solution if and only if $\nu^\star < c(s_k + t_k)$, i.e., the optimum power allocated to the kth subcarrier is positive if and only if $\nu^\star < c(s_k + t_k)$. Note also that the optimum power $P_k^\star$ is decreasing with the value of the Lagrange multiplier $\nu^\star$, since $x_k(P_k)y_k(P_k)$ and $z_k(P_k)$ are decreasing and increasing functions, respectively. See an example of this case in Figure 5.12(a).
  - $P_k^\star$ is equal to 0 ($P_k^\star = 0$). From the KKT conditions it is concluded that the relationship $x_k(0)y_k(0) + \lambda_k^\star z_k(0) = \nu^\star z_k(0)$ has to be fulfilled for some $\lambda_k^\star \geq 0$. The previous relationship is equivalent to $c(s_k + t_k) + \lambda_k^\star = \nu^\star$, i.e., $\lambda_k^\star = \nu^\star - c(s_k + t_k)$. Note that we are assuming a different situation from the previous one and, therefore, $\nu^\star \geq c(s_k + t_k)$, which means that there exists $\lambda_k^\star = \nu^\star - c(s_k + t_k) \geq 0$ that fulfills $x_k(0)y_k(0) + \lambda_k^\star z_k(0) = \nu^\star z_k(0)$. See an example of this case in Figure 5.12(b).

- Once the optimum powers for all the subcarriers have been calculated for a concrete value of the multiplier $\nu^\star$, the total transmit power is calculated as $\sum_{k=0}^{N-1} P_k^\star$. If this power is greater than P_0, then the value of $\nu^\star$ has to be increased and the powers for all the carriers have to be calculated again (note that increasing $\nu^\star$ implies decreasing $P_k^\star$). On the contrary, if $\sum_{k=0}^{N-1} P_k^\star$ is lower than P_0, then the value of $\nu^\star$ has to be decreased

and the optimum powers for all the subcarriers have to be calculated again. This process has to be applied iteratively until a value of $\nu^\star$ is found such that $\sum_{k=0}^{N-1} P_k^\star = P_0$. An adequate iterative algorithm could be based on the *nested intervals* technique, although other algorithms could also be applied.

5.D Appendix: Closed-Form Expressions of the Matrices M and X

The matrices $\mathbf{M}$ and $\mathbf{X}$ are defined as

$$\mathbf{M} \triangleq \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\mathbf{H}\right] \in \mathcal{C}^{(L+M-1)\times n_T M}, \tag{5.158}$$

$$\mathbf{X} \triangleq \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\mathbf{H}^H\mathbf{H}\right] \in \mathcal{C}^{n_T M\times n_T M}. \tag{5.159}$$

In case the actual channel $\mathbf{h}$ and the error $\boldsymbol{\delta}$ in the channel estimate $\widehat{\mathbf{h}} = \mathbf{h}+\boldsymbol{\delta}$ are independent and circularly symmetric Gaussian distributed according to the following mean vectors and covariance matrices $\mathbf{h} \sim \mathcal{CN}(\mathbf{m}, \boldsymbol{\Lambda})$ and $\boldsymbol{\delta} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$, then the actual channel conditioned to the channel estimate $\mathbf{h}|\widehat{\mathbf{h}}$ also follows a Gaussian distribution with the following mean vector and covariance matrix (see [Kay93]):

$$\mathbf{h}|\widehat{\mathbf{h}} \sim \mathcal{CN}(\mathbf{t}, \mathbf{C}), \tag{5.160}$$

$$\mathbf{C} = \left(\boldsymbol{\Lambda}^{-1} + \boldsymbol{\Sigma}^{-1}\right)^{-1} \in \mathcal{C}^{n_T L\times n_T L}, \tag{5.161}$$

$$\mathbf{t} = \mathbf{C}\left(\boldsymbol{\Lambda}^{-1}\mathbf{m} + \boldsymbol{\Sigma}^{-1}\widehat{\mathbf{h}}\right) \in \mathcal{C}^{n_T L\times 1}. \tag{5.162}$$

Deduction of the Expression of M

Let us write the vector $\mathbf{t} = \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\mathbf{h}\right]$, whose expression is given by (5.162), as $\mathbf{t} = \left[\mathbf{t}_1^T \cdots \mathbf{t}_{n_T}^T\right]^T$, where $\mathbf{t}_p = \left[t_p(0) \cdots t_p(L-1)\right]^T \in \mathcal{C}^{L\times 1}$ and $t_p(n) = \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[h_p(n)\right]$. According to this, the matrix $\mathbf{M}$ can be written as $\mathbf{M} = \left[\mathbf{M}_1 \cdots \mathbf{M}_{n_T}\right] \in \mathcal{C}^{(L+M-1)\times n_T M}$, where $\mathbf{M}_p = \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\mathbf{H}_p\right] \in \mathcal{C}^{(L+M-1)\times M}$ is the Toeplitz convolution matrix associated to the time response $\mathbf{t}_p$, which is constructed in the same way as shown in (5.81).

Deduction of the Expression of X

The matrix $\mathbf{C} = \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[(\mathbf{h}-\mathbf{t})(\mathbf{h}-\mathbf{t})^H\right]$, whose expression is given by (5.161), can be written as

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{1,1} & \cdots & \mathbf{C}_{1,n_T} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{n_T,1} & \cdots & \mathbf{C}_{n_T,n_T} \end{bmatrix} \in \mathcal{C}^{n_T L\times n_T L}, \tag{5.163}$$

where $\mathbf{C}_{p,q} = \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[(\mathbf{h}_p - \mathbf{t}_p)(\mathbf{h}_q - \mathbf{t}_q)^H\right] \in \mathcal{C}^{L\times L}$. Based on this, the expression of the matrix $\mathbf{X}$ is given by

$$\mathbf{X} = \begin{bmatrix} \mathbf{R}_{1,1} & \cdots & \mathbf{R}_{1,n_T} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{n_T,1} & \cdots & \mathbf{R}_{n_T,n_T} \end{bmatrix} \in \mathcal{C}^{n_T M\times n_T M}, \tag{5.164}$$

where $\mathbf{R}_{p,q} = \mathbb{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left[\mathbf{H}_p^H\mathbf{H}_q\right] \in \mathcal{C}^{M\times M}$. Finally, the (m,n)th component of the matrix $\mathbf{R}_{p,q}$ can be expressed as follows:

$$\left[\mathbf{R}_{p,q}\right]_{m,n} = \begin{cases} \sum_{l=1}^{L+m-n}\left[\mathbf{C}_{q,p}\right]_{l,l+n-m} + t_p^*(l+n-m-1)t_q(l-1), & m\le n, \\ \sum_{l=1}^{L+n-m}\left[\mathbf{C}_{q,p}\right]_{l+m-n,l} + t_p^*(l-1)t_q(l+m-n-1), & m>n. \end{cases} \tag{5.165}$$

5.E Appendix: Closed-Form Expression of the Optimum MMSE Transmit Filters

From the KKT condition (5.91), the expression of the optimum receive gain factor (5.89), and the transmit power constraint (5.82), the following equations have to be fulfilled:

$$|a_R^\star|^2\mathbf{X}\mathbf{z}^\star - a_R^\star\mathbf{M}^H\mathbf{1}_l + \lambda^\star\mathbf{z}^\star = \mathbf{0}, \tag{5.166}$$

$$a_R^\star = \frac{\mathbf{1}_l^T\mathbf{M}\mathbf{z}^\star}{\mathbf{z}^{\star H}\mathbf{X}\mathbf{z}^\star + \sigma_n^2}, \tag{5.167}$$

$$\mathbf{z}^{\star H}\mathbf{z}^\star = P_0. \tag{5.168}$$

From (5.166), the following relationship is obtained by left-multiplying by $\mathbf{z}^{\star H}$, and taking into account the expression of the optimum receive gain factor (5.167) and the power constraint (5.168):

$$\frac{\mathbf{z}^{\star H}\mathbf{M}^H\mathbf{1}_l\mathbf{1}_l^T\mathbf{M}\mathbf{z}^\star}{\left(\mathbf{z}^{\star H}\mathbf{X}\mathbf{z}^\star + \sigma_n^2\right)^2}\mathbf{z}^{\star H}\mathbf{X}\mathbf{z}^\star - \frac{\mathbf{z}^{\star H}\mathbf{M}^H\mathbf{1}_l\mathbf{1}_l^T\mathbf{M}\mathbf{z}^\star}{\mathbf{z}^{\star H}\mathbf{X}\mathbf{z}^\star + \sigma_n^2} + \lambda^\star P_0 = 0. \tag{5.169}$$

By means of a correct manipulation of the previous expression, the optimum value of the Lagrange multiplier $\lambda^\star$ can be expressed as a function of the optimum transmit filters as

$$\lambda^\star = \frac{\sigma_n^2}{P_0}\frac{\mathbf{z}^{\star H}\mathbf{M}^H\mathbf{1}_l\mathbf{1}_l^T\mathbf{M}\mathbf{z}^\star}{\left(\mathbf{z}^{\star H}\mathbf{X}\mathbf{z}^\star + \sigma_n^2\right)^2} \ge 0. \tag{5.170}$$

Using the previous result in (5.166), the following expression is obtained after some simple mathematical manipulations:

$$\frac{\mathbf{z}^{\star H}\mathbf{M}^H\mathbf{1}_l\mathbf{1}_l^T\mathbf{M}\mathbf{z}^\star}{\mathbf{z}^{\star H}\mathbf{X}\mathbf{z}^\star + \sigma_n^2}\left(\mathbf{X} + \frac{\sigma_n^2}{P_0}\mathbf{I}\right)\mathbf{z}^\star = \mathbf{1}_l^T\mathbf{M}\mathbf{z}^\star\mathbf{M}^H\mathbf{1}_l, \tag{5.171}$$

from which it is deduced the final expression of the optimum transmit filters:

$$\mathbf{z}^\star = \mu\left(\mathbf{X} + \frac{\sigma_n^2}{P_0}\mathbf{I}\right)^{-1}\mathbf{M}^H\mathbf{1}_l, \tag{5.172}$$

where the constant μ, defined as

$$\mu = \mathbf{1}_l^T \mathbf{M}\mathbf{z}^\star \frac{\mathbf{z}^{\star H}\mathbf{X}\mathbf{z}^\star + \sigma_n^2}{\mathbf{z}^{\star H}\mathbf{M}^H \mathbf{1}_l \mathbf{1}_l^T \mathbf{M}\mathbf{z}^\star}, \tag{5.173}$$

guarantees that the transmit power constraint $\mathbf{z}^{\star H}\mathbf{z}^\star = P_0$ is fulfilled.

Chapter 6

Robust Maximin Design of MIMO Single-User Communications

6.1 Introduction

As explained in Chapter 5, the performance of a system can be severely degraded if a non-robust design is taken, i.e., if the imperfections in the available CSI during the design stage are not taken into account explicitly. In this chapter, a robust design is proposed for a MIMO communication system under the maximin approach. This strategy, as explained in §5.4.1, consists in modeling the error in the CSI as belonging to a predefined uncertainty region and then looking for the design that optimizes the worst system performance for any error in this region.

This chapter is structured in the following way. Section 6.2 is devoted to the system model and the problem formulation. The solution to the maximin robust design is obtained in Section 6.3 using the theory of convex optimization. The different sources of errors in the CSI jointly with the description of the corresponding uncertainty regions are provided in Section 6.4, whereas a closed-form solution to the case of spherical uncertainty regions is deduced in Section 6.5. In Section 6.6, some applications and extensions of the robust design are given for the case of multicarrier systems and adaptive modulation schemes. The benefits of the proposed solution are shown in Section 6.7 by means of simulation results and, finally, a summary and some conclusions are obtained in Section 6.8.

The publications regarding the results shown in this chapter are [Pal05, PI04b, PI04d, PI04c].

6.2 System Model and Problem Formulation

As in Chapter 3, a single-user communication system is assumed, where the transmitter and the receiver have n_T and n_R antennas, respectively. First, a single-carrier flat fading MIMO channel is

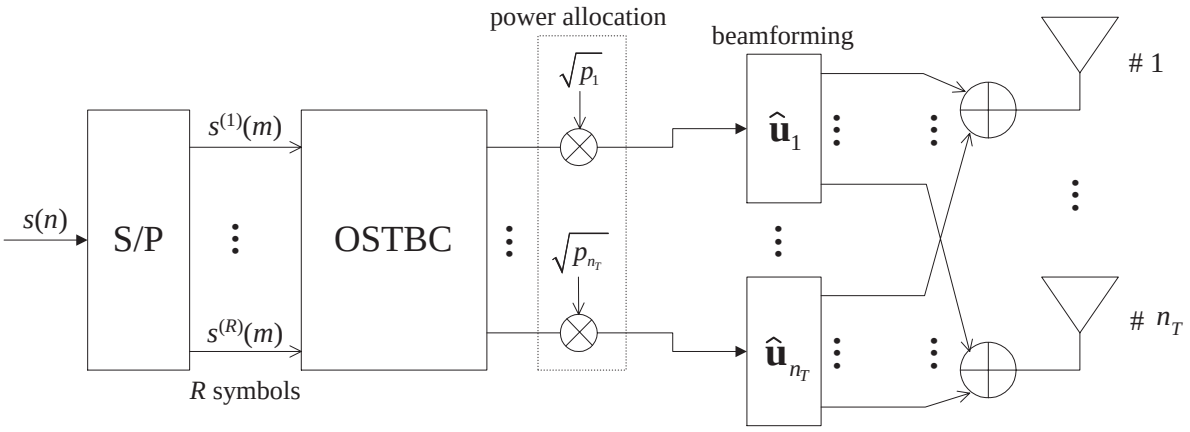


Figure 6.1: General architecture of the robust transmitter based on the concatenation of an OSTBC block, a power allocation, and multiple beamforming.

studied, although afterwards, the obtained results are extended to the case of frequency selective channels and the use of the OFDM modulation in §6.6.4. At the receiver, only the presence of AWGN is considered with a power equal to σ_n^2, leading to the following correlation matrix: $\mathbf{R}_n = \sigma_n^2\mathbf{I}$. The extension of the results to scenarios in which there are also interferences and, consequently, the noise plus interferences correlation matrix is not proportional to the identity matrix, is direct and is presented in §6.6.3. Under these considerations, the matrix $\mathbf{H} \in \mathcal{C}^{n_R \times n_T}$ represents the response of the MIMO channel, as explained in §3.2.2.

As commented previously, the final objective is to obtain a robust maximin design of the communication system according to an imperfect channel estimate $\widehat{\mathbf{H}}$ at the transmitter, which is modeled as

$$\mathbf{H} \triangleq \widehat{\mathbf{H}} + \mathbf{\Delta}, \tag{6.1}$$

where $\mathbf{\Delta} \in \mathcal{C}^{n_R \times n_T}$ is the error in the estimate.[1] The CSI at the receiver is assumed to be perfect.

Consider, for illustrative purposes, that one symbol has to be transmitted at one time instant. In case that a perfect CSI is available, it has been shown in Chapter 3 that the optimum beamforming solution consists in the transmission through the eigenvector of $\mathbf{H}^H\mathbf{H}$ associated to the maximum eigenvalue (see §3.2.3 and (3.22) for more details, taking into account that $\mathbf{R}_n = \sigma_n^2\mathbf{I}$). In case that the channel knowledge is imperfect, transmitting through the maximum eigenmode of $\widehat{\mathbf{H}}^H\widehat{\mathbf{H}}$ constitutes the non-robust or naive solution, which may be quite sensitive to the errors. Therefore, a robust design is expected to use more eigenmodes than the maximum one.

[1] Note that the notation $\mathbf{H} = \widehat{\mathbf{H}} + \mathbf{\Delta}$ has been taken instead of $\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{\Delta}$, as used in Chapter 5, in order to facilitate the formulation and the problem statement corresponding to the robust maximin design.

The proposed design of the robust transmitter is based on a linear processing scheme, whereas at the receiver, an optimum ML detector is considered assuming a perfect channel knowledge. The robust transmitter architecture is composed of an OSTBC block, whose outputs are transmitted through all the eigenmodes of $\widehat{\mathbf{H}}^H\widehat{\mathbf{H}}$ using an adequate power distribution among them, as opposed to the non-robust design, which uses only the maximum eigenmode. This architecture is shown in Figure 6.1, where the OSTBC, the power allocation, and the beamforming stages are explicitly shown (similar transmitter architectures have been proposed in other works, such as [Jön02, Xia04, Zho02, Zho03]).

Consider that an OSTBC block is used, such that R independent complex symbols $(\{s^{(l)}(m)\}_{l=1}^{R})$ are transmitted simultaneously over T periods of time, i.e., the code rate is R/T. According to this, the transmitted signal model can be formulated as follows, similarly to linear dispersion codes [Has02] and OSTBC [Tar99a, Gan01]:

$$\mathbf{X}_T(m) = \widehat{\mathbf{U}}\,\mathrm{diag}\left(\{\sqrt{p_i}\}\right)\sum_{l=1}^{R}\left(\mathbf{T}_l^{(r)}\,\mathrm{Re}\{s^{(l)}(m)\} + j\mathbf{T}_l^{(i)}\,\mathrm{Im}\{s^{(l)}(m)\}\right) \in \mathbb{C}^{n_T\times T}, \tag{6.2}$$

where the complex symbols are assumed to have a normalized energy, i.e., $\mathbb{E}\left[|s^{(l)}(m)|^2\right] = 1$, $\mathbf{T}_l^{(r)}$ and $\mathbf{T}_l^{(i)} \in \mathbb{C}^{n_T\times T}$ are the matrices that multiply the real and imaginary parts of the symbol $s^{(l)}(m)$ before being transmitted, respectively, and each of the n_T rows of $\mathbf{X}_T(m)$ corresponds to the T signal samples that are transmitted through each antenna corresponding to the mth block of R symbols $(\{s^{(l)}(m)\}_{l=1}^{R})$. The matrices $\mathbf{T}_l^{(r)}$ and $\mathbf{T}_l^{(i)}$ represent the OSTBC and belong to the Hurwitz-Radon family of matrices (see [Tar99a] and [Gan01]). The unitary matrix $\widehat{\mathbf{U}} = \left[\widehat{\mathbf{u}}_1\cdots\widehat{\mathbf{u}}_{n_T}\right] \in \mathbb{C}^{n_T\times n_T}$ contains the n_T eigenvectors of $\widehat{\mathbf{H}}^H\widehat{\mathbf{H}}$ with eigenvalues $\{\widehat{\lambda}_i\}$ sorted in deacreasing order, p_i is the power allocated to the transmission through the ith estimated eigenmode, and $\mathrm{diag}\left(\{\sqrt{p_i}\}\right)$ is a diagonal matrix whose elements are $\{\sqrt{p_i}\}$. According to this signal model, the ML detector reduces to a bank of linear filters, as in the case of OSTBC, applied to the received signal samples collected in the matrix

$$\mathbf{X}_R(m) = \mathbf{H}\mathbf{X}_T(m) + \mathbf{N}(m) \in \mathbb{C}^{n_R\times T}, \tag{6.3}$$

where $\mathbf{N}(m)$ represents the AWGN at the receiver.

The design objective is to calculate the optimum power allocation strategy $\{p_i\}$ subject to a transmit power constraint under an adequate performance criterion. If the transmit power budget is P_0, the power constraint can be expressed in terms of the factors $\{p_i\}$ as

$$\frac{1}{R}\mathbb{E}\left[\|\mathbf{X}_T(m)\|_F^2\right] = \sum_{i=1}^{n_T} p_i \le P_0, \quad p_i \ge 0. \tag{6.4}$$

Note that the set of feasible power distributions is convex in $\{p_i\}$, since the constraints detailed in (6.4) are linear. Note also that, according to this notation, the non-robust design corresponds to $p_1 = P_0,\ \ p_i = 0,\ \ i = 2,\ldots,n_T$, i.e., only the maximum eigenmode is used for transmission.

For the considered system (6.2) using OSTBC with ML detection, the performance can be measured by the SNR expressed as (see [Tar99a] and [Gan01])

$$\mathrm{SNR} = \frac{1}{\sigma_n^2} \mathrm{Tr}\left(\widehat{\mathbf{U}}^H \mathbf{H}^H \mathbf{H} \widehat{\mathbf{U}} \,\mathrm{diag}(\mathbf{p})\right), \tag{6.5}$$

where $\mathbf{p} = \left[p_1 \cdots p_{n_T}\right]^T \in I\!R^{n_T \times 1}$ and $\mathrm{diag}(\mathbf{p})$ is a diagonal matrix with elements $\{p_i\}$. Based on this, the performance function f in this system can be defined as

$$f\left(\mathbf{p}, \boldsymbol{\Delta}\right) \triangleq \mathrm{Tr}\left(\widehat{\mathbf{U}}^H \mathbf{H}^H \mathbf{H} \widehat{\mathbf{U}} \,\mathrm{diag}(\mathbf{p})\right) = \mathrm{Tr}\left(\widehat{\mathbf{U}}^H \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta}\right)^H \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta}\right) \widehat{\mathbf{U}} \,\mathrm{diag}(\mathbf{p})\right), \tag{6.6}$$

whose maximization with respect to $\mathbf{p}$ is the design objective and where the error model (6.1) has been introduced. Note that f is linear and, therefore, concave in $\mathbf{p}$; and convex-quadratic in $\boldsymbol{\Delta}$.

As commented previously, the maximin approach has been chosen to include robustness in the design of the power allocation. According to it, an uncertainty region $\mathcal{R}$ for the error in the estimate $\boldsymbol{\Delta}$ has to be known, which, in the following, will be assumed to be a convex set and to have a non-empty interior. This region models the imprecise knowledge of the channel and, therefore, the size of the region should be larger as the quality of the CSI decreases and its shape should be related to the physical phenomenon producing the error, as will be shown in §6.4. The objective of the maximin design is to look for the power allocation $\mathbf{p}$ that optimizes the worst performance for any error in the uncertainty region, expressed in this case as $\inf_{\boldsymbol{\Delta} \in \mathcal{R}} f\left(\mathbf{p}, \boldsymbol{\Delta}\right)$. Therefore, the robust approach can be formulated as

$$\begin{array}{ll} \underset{\mathbf{p}}{\text{maximize}} & \inf\limits_{\boldsymbol{\Delta} \in \mathcal{R}} f\left(\mathbf{p}, \boldsymbol{\Delta}\right) \\ \text{subject to} & \mathbf{1}^T \mathbf{p} \leq P_0, \\ & p_i \geq 0, \qquad \forall i, \end{array} \tag{6.7}$$

where $\mathbf{1} = \left[1 \cdots 1\right]^T \in I\!R^{n_T \times 1}$ is the all-one vector.

6.3 Solution to the Maximin Problem

In the following, different ways of finding the solution to the stated robust maximin design problem (6.7) are presented. First, an exhaustive numerical search method is proposed and, afterwards, a reformulation of the problem is given, obtaining an equivalent and much simpler convex optimization problem.

6.3.1 Direct Solution to the Original Problem

The direct way to solve the maximin problem is to obtain the minimization of f analytically and then solve the outer maximization, either numerically or analytically. Such an approach, however, is difficult because it is not clear what is the minimizing $\boldsymbol{\Delta}$ for a given $\mathbf{p}$ in closed-form.

One can also consider the inner minimization numerically for a given $\mathbf{p}$

$$\widetilde{f}(\mathbf{p}) \triangleq \inf_{\mathbf{\Delta}\in\mathcal{R}} f\left(\mathbf{p},\mathbf{\Delta}\right), \tag{6.8}$$

and then solving again the outer maximization $\sup_{\mathbf{p}} \widetilde{f}(\mathbf{p})$ numerically. Note that the inner minimization is a convex problem, since f is convex in $\mathbf{\Delta}$ and the constraint set $\mathcal{R}$ is also convex. The outer maximization is a concave optimization problem (see §2.3.3), since the constraint set for $\mathbf{p}$ is convex (the constraints in (6.4) are linear) and $\widetilde{f}$ is concave. This procedure allows to find the robust power allocation $\mathbf{p}^\star$, although it is computationally very costly. This is because each iteration for the outer maximization requires an evaluation of $\widetilde{f}(\mathbf{p})$ (and possibly also of its gradient), which in turns requires solving the inner minimization numerically with as many iterations as needed to converge.

The proof of the concavity of $\widetilde{f}(\mathbf{p})$ is shown below:

$$\begin{aligned}
\widetilde{f}(\theta\mathbf{p}_1+(1-\theta)\mathbf{p}_2) &= \inf_{\mathbf{\Delta}\in\mathcal{R}} f(\theta\mathbf{p}_1+(1-\theta)\mathbf{p}_2,\mathbf{\Delta}) \\
&= \inf_{\mathbf{\Delta}\in\mathcal{R}} [\theta f(\mathbf{p}_1,\mathbf{\Delta})+(1-\theta)f(\mathbf{p}_2,\mathbf{\Delta})] \\
&\geq \theta \inf_{\mathbf{\Delta}\in\mathcal{R}} f(\mathbf{p}_1,\mathbf{\Delta}) + (1-\theta)\inf_{\mathbf{\Delta}\in\mathcal{R}} f(\mathbf{p}_2,\mathbf{\Delta}) \\
&= \theta\widetilde{f}(\mathbf{p}_1)+(1-\theta)\widetilde{f}(\mathbf{p}_2), \qquad \forall\mathbf{p}_1,\mathbf{p}_2,\ \ \forall\theta\in[0,1],
\end{aligned} \tag{6.9}$$

where the linearity of f in $\mathbf{p}$ has been used in the second equality.

Other numerical methods could also be used, such as the algorithm proposed in [Roc71b] to find saddle-points of maximin problems based on a modified steepest descent over $\mathbf{p}$ and $\mathbf{\Delta}$ simultaneously. In [Ž00], an alternative algorithm for the same problem is derived based on the interior point method.

In the following, a more efficient and elegant way to solve the problem is shown, based on a transformation of the original maximin problem (6.7) into a simple convex optimization problem.

6.3.2 Reformulation of the Problem as a Simplified Convex Problem

Before presenting the technique that allows to transform the original problem into a simplified convex optimization problem, some mathematical preliminaries are given concerning concave-convex functions and saddle-points.

Some Mathematical Preliminaries: Concave-Convex Functions and Saddle-Points

In this subsection, the concept of saddle-point is defined and two basic results are then given, since they will be useful in the following.

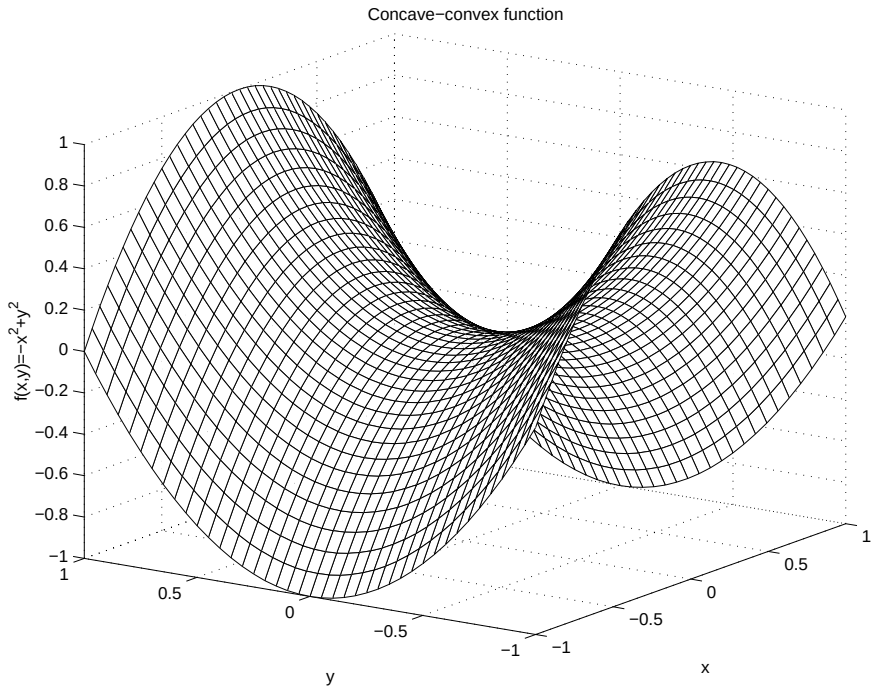


Figure 6.2: 3D representation of the concave-convex function $f(x, y) = -x^2 + y^2$.

Definition 2 *A point* $(\mathbf{x}^\star, \mathbf{y}^\star) \in \mathcal{X} \times \mathcal{Y}$ *is a saddle-point of the function* $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ *with respect to maximizing over* $\mathcal{X}$ *and minimizing over* $\mathcal{Y}$ *if*

$$f(\mathbf{x}, \mathbf{y}^\star) \leq f(\mathbf{x}^\star, \mathbf{y}^\star) \leq f(\mathbf{x}^\star, \mathbf{y}), \quad \forall \mathbf{x} \in \mathcal{X},\ \forall \mathbf{y} \in \mathcal{Y}. \tag{6.10}$$

Given a saddle-point $(\mathbf{x}^\star, \mathbf{y}^\star)$, $f^\star \triangleq f(\mathbf{x}^\star, \mathbf{y}^\star)$ *is defined as the saddle-value.*

Lemma 1 *[Roc71a, Cor. 37.6.2] Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be non-empty closed bounded convex sets and let* f *be a real continuous finite concave-convex function on* $\mathcal{X} \times \mathcal{Y}$.[2] *Then,* f *has a saddle-point with respect to* $\mathcal{X} \times \mathcal{Y}$.

In Figure 6.2, a 3D representation of the concave-convex function $f(x, y) = -x^2 + y^2$ is given, where the point $(0, 0)$ is the saddle-point.

Lemma 2 *[Roc71a, Lem. 36.2] Let* f *be any function from a non-empty product set* $\mathcal{X} \times \mathcal{Y}$ *to* $[-\infty, +\infty]$. *If a saddle-point* $(\mathbf{x}^\star, \mathbf{y}^\star) \in \mathcal{X} \times \mathcal{Y}$ *of* f *exists (with respect to maximizing over* $\mathcal{X}$ *and minimizing over* $\mathcal{Y}$*), then*

$$f(\mathbf{x}^\star, \mathbf{y}^\star) = \sup_{\mathbf{x} \in \mathcal{X}} \inf_{\mathbf{y} \in \mathcal{Y}} f(\mathbf{x}, \mathbf{y}) = \inf_{\mathbf{y} \in \mathcal{Y}} \sup_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}, \mathbf{y}), \tag{6.11}$$

[2] The function $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ is concave-convex if $f(\mathbf{x}, \mathbf{y})$ is concave with respect to $\mathbf{x}$ for any $\mathbf{y} \in \mathcal{Y}$, and is convex with respect to $\mathbf{y}$ for any $\mathbf{x} \in \mathcal{X}$.

and the outer optimizations in $\sup_{\mathbf{x}\in\mathcal{X}} \inf_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{x},\mathbf{y})$ *and* $\inf_{\mathbf{y}\in\mathcal{Y}} \sup_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x},\mathbf{y})$ *are attained at* $\mathbf{x}^\star$ *and* $\mathbf{y}^\star$*, respectively. In other words, if a saddle-point exists, the order of the supremum and infimum operators can be interchanged.*

Robust Maximin Design

In the following, the original problem (6.7) is transformed into a simplified convex problem consisting of a single optimization stage, instead of the two stages defined originally. Thanks to this transformation, the powerful numerical algorithms available in software packages for solving convex problems can be applied to find the optimum solution to the design problem, requiring much less computational effort than the algorithms mentioned previously.

The proof of the following proposition describing the problem transformation is based on the existence of a saddle-point of the function f defined in (6.6). This saddle-point can be interpreted as a Nash equilibrium in a pure strategic game, as defined in *game theory*, in which the transmitter and the error in the channel estimate are players with opposed objectives, i.e., players that try to maximize and minimize the SNR, respectively. A more complete description of pure strategic games can be found in [Osb94] and an application of them to capacity maximization in MIMO channels is shown in [Pal03b].

Proposition 1 *The original maximin problem (6.7) can be transformed into the simplified convex optimization problem*

$$\begin{array}{ll} \underset{t,\boldsymbol{\Delta}}{\text{minimize}} & t \\ \text{subject to} & t \geq P_0 \widehat{\mathbf{u}}_i^H \left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H \left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{u}}_i, \qquad \forall i, \\ & \boldsymbol{\Delta}\in\mathcal{R}. \end{array} \tag{6.12}$$

The optimum robust power allocation $\mathbf{p}^\star = \left[p_1^\star \cdots p_{n_T}^\star\right]^T$ *is equal to the optimum dual variables* $\{\gamma_i^\star\}$ *associated to the inequalities* $t \geq P_0 \widehat{\mathbf{u}}_i^H \left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H \left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{u}}_i$ *in (6.12) multiplied by the power budget* P_0*, i.e.,* $p_i^\star = P_0\gamma_i^\star$*. Besides, the optimum primal variable* $\boldsymbol{\Delta}^\star$ *of (6.12) minimizes* $f(\mathbf{p}^\star,\boldsymbol{\Delta})$*. Hence, the primal and dual variables of (6.12) give the saddle-point of* f*.*

Proof. The function $f(\mathbf{p},\boldsymbol{\Delta})$, which is concave-convex, and the optimization sets satisfy the conditions required by Lemma 1. Consequently, there exists a saddle-point (see Definition 2) of the maximin problem (6.7), i.e., there exist $\mathbf{p}^\star$ and $\boldsymbol{\Delta}^\star$ fulfilling the constraints and satisfying

$$f(\mathbf{p},\boldsymbol{\Delta}^\star) \leq f(\mathbf{p}^\star,\boldsymbol{\Delta}^\star) \leq f(\mathbf{p}^\star,\boldsymbol{\Delta}) \tag{6.13}$$

for any feasible $\mathbf{p}$ and $\boldsymbol{\Delta}$. The solution to the original problem (6.7) is $\mathbf{p}^\star$ and the saddle-value $f^\star \triangleq f(\mathbf{p}^\star,\boldsymbol{\Delta}^\star)$ is $\widetilde{f}(\mathbf{p}^\star)$ (see Lemma 2). The existence of the saddle-point permits to

interchange the outer and inner optimizations in the original maximin problem, obtaining the minimax problem

$$\begin{array}{ll} \underset{\boldsymbol{\Delta}}{\text{minimize}} & \sup\limits_{\mathbf{1}^T\mathbf{p}\leq P_0, p_i\geq 0} f\left(\mathbf{p},\boldsymbol{\Delta}\right) \\ \text{subject to} & \boldsymbol{\Delta}\in\mathcal{R}, \end{array} \tag{6.14}$$

with the advantage that the inner maximization is a linear program with linear constraints:

$$\begin{array}{ll} \underset{\mathbf{p}}{\text{maximize}} & \sum_{i=1}^{n_T} p_i \left[\widehat{\mathbf{U}}^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{U}}\right]_{ii} \\ \text{subject to} & \mathbf{1}^T\mathbf{p}\leq P_0, \\ & p_i\geq 0, \qquad \forall i. \end{array} \tag{6.15}$$

It can be shown that the optimum value of this maximization is the maximum element of the diagonal of the matrix $\widehat{\mathbf{U}}^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{U}}$ multiplied by the power budget P_0, i.e.:

$$\begin{aligned} \sup_{\mathbf{1}^T\mathbf{p}\leq P_0, p_i\geq 0} f(\mathbf{p},\boldsymbol{\Delta}) &= P_0 \max_i \left[\widehat{\mathbf{U}}^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{U}}\right]_{ii} \\ &= P_0 \max_i \left\{\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{u}}_i\right\}, \end{aligned} \tag{6.16}$$

where the power allocation $\mathbf{p}$ achieving this optimum value is not unique if the maximum value is attained by more than one element of the diagonal of the matrix $\widehat{\mathbf{U}}^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{U}}$. As a consequence of this result, the original problem (6.7) can be written as the convex problem (6.12) shown in Proposition 1, where the dummy variable t has been introduced (note that minimizing the maximum of a set of numbers is equivalent to minimizing an upper-bound of all the numbers in the set).

Solving the convex problem (6.12) gives the saddle-value $t^\star = f^\star = f(\mathbf{p}^\star,\boldsymbol{\Delta}^\star)$ and the worst-case error $\boldsymbol{\Delta}^\star$ of the saddle-point of the problem (see Lemma 2); however, the optimal robust power distribution $\mathbf{p}^\star$ is still unknown. It turns out that the optimum Lagrange multipliers $\gamma_i^\star$ associated to the inequality constraints $t\geq P_0\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{u}}_i$ in problem (6.12) provide the optimum normalized robust power distribution, i.e., $p_i^\star = P_0\gamma_i^\star$, as proved below.

The problem (6.12) can be solved by formulating the necessary and sufficient Karush-Kuhn-Tucker (KKT) conditions (Slater's condition described in §2.3.4 holds since $\mathcal{R}$ has a non-empty interior) [Boy04], which, according to Lemma 2, are satisfied by the the worst-case error $\boldsymbol{\Delta}^\star$ along with the optimum dual variables. On the other hand, it is clear that $\boldsymbol{\Delta}^\star$ is also the solution to the convex problem $\min_{\boldsymbol{\Delta}} f(\mathbf{p}^\star,\boldsymbol{\Delta})$ (from the second inequality in (6.13)), where $\mathbf{p}^\star$ is the robust power distribution, and, therefore, the worst-case error $\boldsymbol{\Delta}^\star$ must satisfy the KKT conditions for the problem $\min_{\boldsymbol{\Delta}} f(\mathbf{p}^\star,\boldsymbol{\Delta})$ as well. By a simple comparison between both sets of KKT conditions, it can be seen that, for $p_i^\star = P_0\gamma_i^\star$, the worst-case error $\boldsymbol{\Delta}^\star$ satisfies both sets of conditions and, hence, that is an optimal power allocation.

The Lagrangian of the problem (6.12) (characterizing for convenience and w.l.o.g. the uncertainty convex region $\mathcal{R}$ as the intersection of a set of convex constraints of the form $f_i(\boldsymbol{\Delta})\leq 0$)

is

$$L_1(t,\boldsymbol{\Delta};\boldsymbol{\gamma},\boldsymbol{\mu}) \triangleq t+\sum_{i=1}^{n_T}\gamma_i\left(P_0\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{u}}_i-t\right)+\sum_i\mu_i f_i(\boldsymbol{\Delta}) \tag{6.17}$$

$$= t\left(1-\sum_{i=1}^{n_T}\gamma_i\right)+P_0\,\mathrm{Tr}\left(\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{U}}\,\mathrm{diag}(\{\gamma_i\})\widehat{\mathbf{U}}^H\right) +\sum_i\mu_i f_i(\boldsymbol{\Delta}), \tag{6.18}$$

where the equality $\sum_{i=1}^{n_T}\gamma_i\widehat{\mathbf{u}}_i\widehat{\mathbf{u}}_i^H=\widehat{\mathbf{U}}\,\mathrm{diag}(\{\gamma_i\})\widehat{\mathbf{U}}^H$ has been used. Therefore, the KKT conditions for this problem are:

$$f_i(\boldsymbol{\Delta}^\star)\le 0,\quad t^\star\ge P_0\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}^\star\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}^\star\right)\widehat{\mathbf{u}}_i, \tag{6.19}$$

$$\mu_i^\star\ge 0,\quad \gamma_i^\star\ge 0, \tag{6.20}$$

$$\sum_{i=1}^{n_T}\gamma_i^\star=1,\quad P_0\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}^\star\right)\widehat{\mathbf{U}}\,\mathrm{diag}(\{\gamma_i^\star\})\widehat{\mathbf{U}}^H+\sum_i\mu_i^\star\nabla f_i(\boldsymbol{\Delta}^\star)=\mathbf{0}, \tag{6.21}$$

$$\mu_i^\star f_i(\boldsymbol{\Delta}^\star)=0,\quad \gamma_i^\star\left(P_0\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}^\star\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}^\star\right)\widehat{\mathbf{u}}_i-t^\star\right)=0. \tag{6.22}$$

Now the Lagrangian for the problem $\min_{\boldsymbol{\Delta}} f(\mathbf{p}^\star,\boldsymbol{\Delta})$ is

$$L_2(\boldsymbol{\Delta};\boldsymbol{\alpha})\triangleq\mathrm{Tr}\left(\widehat{\mathbf{U}}^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{U}}\,\mathrm{diag}(\mathbf{p}^\star)\right)+\sum_i\alpha_i f_i(\boldsymbol{\Delta}) \tag{6.23}$$

and the KKT conditions for the optimal error and multipliers are:

$$f_i(\boldsymbol{\Delta}^\star)\le 0, \tag{6.24}$$

$$\alpha_i^\star\ge 0, \tag{6.25}$$

$$\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}^\star\right)\widehat{\mathbf{U}}\,\mathrm{diag}(\mathbf{p}^\star)\widehat{\mathbf{U}}^H+\sum_i\alpha_i^\star\nabla f_i(\boldsymbol{\Delta}^\star)=\mathbf{0}, \tag{6.26}$$

$$\alpha_i^\star f_i(\boldsymbol{\Delta}^\star)=0. \tag{6.27}$$

From the comparison of both sets of KKT conditions (6.19)-(6.22) and (6.24)-(6.27), it is clear that they are satisfied by the same worst-case error $\boldsymbol{\Delta}^\star$ taking $\alpha_i^\star=\mu_i^\star$ and $p_i^\star=P_0\gamma_i^\star$. Besides, from (6.22) it is concluded that the saddle-value is $f^\star=t^\star=\sum_{i=1}^{n_T}P_0\gamma_i^\star\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}^\star\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}^\star\right)\widehat{\mathbf{u}}_i=f(\mathbf{p}^\star,\boldsymbol{\Delta}^\star)$. In other words, given a solution to (6.19)-(6.22), a solution to (6.24)-(6.27) is automatically obtained, which means that $\boldsymbol{\Delta}^\star$ is the worst-case error for $\mathbf{p}^\star$; moreover, the pair $(\mathbf{p}^\star,\boldsymbol{\Delta}^\star)$ is a saddle-point of the problem and, consequently, $\mathbf{p}^\star$ is a robust power allocation. Note that the transmit power constraint is fulfilled with equality, since the optimum dual variables $\{\gamma_i^\star\}$ are required to satisfy $\gamma_i^\star\ge 0$ (see (6.20)) and $\sum_{i=1}^{n_T}\gamma_i^\star=1$ (see (6.21)). ■

Summarizing, the original maximin power allocation problem (6.7) can be solved by considering the simplified convex problem (6.12). The values of the optimum Lagrange multipliers

for this problem provide the normalized power distribution to be applied among the estimated eigenmodes. Currently, there exist many software packages implementing very efficient numerical algorithms, such as the primal-dual interior point methods, that are able to solve convex optimization problems and give, not only the value of the optimum primal variables, i.e., $t^\star$ and $\mathbf{\Delta}^\star$ in problem (6.12), but also the optimum value of the dual variables, i.e., the Lagrange multipliers $\{\gamma_i^\star\}$. Consequently, by using these algorithms, the saddle-value, the worst-case error, and also the optimum robust power allocation can be calculated efficiently. In addition, it can be shown that, for some concrete uncertainty regions $\mathcal{R}$, problem (6.12) simplifies to a quadratic problem and, even in some cases, a closed-form solution exists.

6.3.3 Duality Interpretation of the Lagrange Multipliers

In this subsection, a completely different interpretation of the problem is given to derive an alternative proof of Proposition 1, in which the optimum Lagrange multipliers are stated to provide the normalized robust power distribution.

Consider the original maximin problem

$$\begin{array}{ll} \underset{\mathbf{p}}{\text{maximize}} & \inf\limits_{\mathbf{\Delta}\in\mathcal{R}} \ \mathrm{Tr}\left(\widehat{\mathbf{U}}^H\left(\widehat{\mathbf{H}}+\mathbf{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\mathbf{\Delta}\right)\widehat{\mathbf{U}}\,\mathrm{diag}(\mathbf{p})\right) \\ \text{subject to} & \mathbf{1}^T\mathbf{p}\leq P_0, \\ & p_i\geq 0, \qquad \forall i, \end{array} \tag{6.28}$$

where, obviously, the optimum solution is attained when the transmit power constraint inequality is fulfilled with equality, i.e., $\mathbf{1}^T\mathbf{p}=P_0$. This problem can be rewritten in terms of the variables $\overline{p}_i=p_i/P_0$, so that the constraints are formulated as $\mathbf{1}^T\overline{\mathbf{p}}=1$ ($\overline{\mathbf{p}}=\left[\overline{p}_1\cdots\overline{p}_{n_T}\right]^T$) and $\overline{p}_i\geq 0,\ \forall i$. The problem is then

$$\begin{array}{ll} \underset{\overline{\mathbf{p}}}{\text{maximize}} & \inf\limits_{\mathbf{\Delta}\in\mathcal{R}} \ P_0\sum_{i=1}^{n_T}\overline{p}_i\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\mathbf{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\mathbf{\Delta}\right)\widehat{\mathbf{u}}_i \\ \text{subject to} & \mathbf{1}^T\overline{\mathbf{p}}=1, \\ & \overline{p}_i\geq 0, \qquad \forall i. \end{array} \tag{6.29}$$

Let us include a dummy variable t, obtaining:

$$\begin{array}{ll} \underset{\overline{\mathbf{p}}}{\text{maximize}} & \inf\limits_{t,\mathbf{\Delta}\in\mathcal{R}} \ t+\sum_{i=1}^{n_T}\overline{p}_i\left(P_0\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\mathbf{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\mathbf{\Delta}\right)\widehat{\mathbf{u}}_i-t\right) \\ \text{subject to} & \mathbf{1}^T\overline{\mathbf{p}}=1, \\ & \overline{p}_i\geq 0, \qquad \forall i. \end{array} \tag{6.30}$$

It turns out that the constraint $\mathbf{1}^T\overline{\mathbf{p}}=1$ can be removed since, if it is not satisfied, the minimization with respect to t would be unbounded below (simply by looking at the term $t(1-\sum_{i=1}^{n_T}\overline{p}_i)$). Hence, the original problem can be rewritten as

$$\begin{array}{ll} \underset{\overline{\mathbf{p}}}{\text{maximize}} & \inf\limits_{t,\mathbf{\Delta}\in\mathcal{R}} \ t+\sum_{i=1}^{n_T}\overline{p}_i\left(P_0\widehat{\mathbf{u}}_i^H\left(\widehat{\mathbf{H}}+\mathbf{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\mathbf{\Delta}\right)\widehat{\mathbf{u}}_i-t\right) \\ \text{subject to} & \overline{p}_i\geq 0, \qquad \forall i, \end{array} \tag{6.31}$$

which can be recognized as the maximization of the dual function (which in turn is defined as the minimization of the Lagrangian [Boy04]) associated to the problem (see §2.3.4)

$$\begin{array}{ll} \underset{t,\boldsymbol{\Delta}\in\mathcal{R}}{\text{minimize}} & t \\ \text{subject to} & t \geq P_0 \widehat{\mathbf{u}}_i^H \left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H \left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right) \widehat{\mathbf{u}}_i, \qquad \forall i, \end{array} \tag{6.32}$$

which is the same problem as (6.12) and, therefore, the dual variables or Lagrange multipliers γ_i associated to the constraints $t \geq P_0 \widehat{\mathbf{u}}_i^H \left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H \left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right) \widehat{\mathbf{u}}_i$ coincide with $\overline{p}_i$, i.e., $p_i = P_0\gamma_i$, proving Proposition 1 in an alternative way. Note that the constraint $\boldsymbol{\Delta} \in \mathcal{R}$ is implicitly included in both problems (6.31) and (6.32) by defining the domain of the functions in the variable $\boldsymbol{\Delta}$ as $\mathcal{R}$.

From this interpretation, it can be seen that the fundamental reason why the relationship $p_i^\star = P_0\gamma_i^\star$ holds is that the original function f is linear in the power distribution variables $\{p_i\}$ and, hence, they can be interpreted as the Lagrange multipliers.

6.4 Convex Uncertainty Regions

The definition of the uncertainty region $\mathcal{R}$ may impact importantly on the system performance. The size and the shape of this region should take into account the quality of the channel estimate and the imperfections that generate the error, already listed in §5.2, linking the optimization problem and the physical phenomenon producing the error.

In the following, two sources of errors are identified and three different uncertainty regions, jointly with their sizes, are described. In all the cases, the proposed uncertainty regions are convex, as required to solve the optimization problem. Afterwards, a more general list of possible uncertainty regions is given.

6.4.1 Estimation Gaussian Noise

A usual error in the channel estimate comes from the Gaussian noise, especially in TDD systems, where the transmitter can estimate the channel using the signals received in the reverse link, and use it as an estimate in the forward link due to the channel reciprocity principle.

In this subsection, the objective is to derive the expression of an uncertainty region according to an unbiased estimate of the channel and taking into account that the error is Gaussian distributed. Let the unbiased channel estimate be formulated as $\widetilde{\mathbf{H}} \triangleq \mathbf{H}+\mathbf{E}$, where $\mathbf{E}$ is the zero-mean estimation noise, independent from the actual channel realization. Note that a different notation is used for the unbiased channel estimate $\widetilde{\mathbf{H}}$ and the estimation error $\mathbf{E}$ when compared to $\widehat{\mathbf{H}}$ and $\boldsymbol{\Delta}$, as used in the previous sections. In the following, the relationship between $\widehat{\mathbf{H}}$, $\widetilde{\mathbf{H}}$, $\boldsymbol{\Delta}$, and $\mathbf{E}$ is shown, and the corresponding uncertainty region for $\boldsymbol{\Delta}$ is deduced.

Let us define $\mathbf{h} \triangleq \text{vec}(\mathbf{H})$ and $\mathbf{e} \triangleq \text{vec}(\mathbf{E})$, where $\mathbf{h}$ and $\mathbf{e}$ are column vectors resulting from stacking the columns of $\mathbf{H}$ and $\mathbf{E}$, respectively. A usual assumption is to consider that $\mathbf{h}$ and $\mathbf{e}$ are jointly Gaussian distributed with mean values and covariance matrices $\mathbf{m}_h$ and $\mathbf{C}_h$ for $\mathbf{h}$, and $\mathbf{0}$ and $\mathbf{C}_e$ for $\mathbf{e}$. According to this, the distribution of the actual channel conditioned to the unbiased channel estimate follows also a Gaussian distribution [Kay93]:

$$\mathbf{h}|\widetilde{\mathbf{h}} \sim \mathcal{CN}\left(\mathbf{m}_{h|\widetilde{h}}, \mathbf{C}_{h|\widetilde{h}}\right), \quad p_{h|\widetilde{h}}\left(\mathbf{h}|\widetilde{\mathbf{h}}\right) = \frac{1}{\pi^{n_T n_R}\left|\mathbf{C}_{h|\widetilde{h}}\right|} e^{-\left(\mathbf{h}-\mathbf{m}_{h|\widetilde{h}}\right)^H \mathbf{C}_{h|\widetilde{h}}^{-1}\left(\mathbf{h}-\mathbf{m}_{h|\widetilde{h}}\right)}, \tag{6.33}$$

where

$$\mathbf{m}_{h|\widetilde{h}} = \mathbf{m}_h + \mathbf{C}_h\left(\mathbf{C}_h + \mathbf{C}_e\right)^{-1}\left(\widetilde{\mathbf{h}} - \mathbf{m}_{\widetilde{h}}\right) \in \mathcal{C}^{n_T n_R \times 1}, \tag{6.34}$$

$$\mathbf{C}_{h|\widetilde{h}} = \left(\mathbf{C}_h^{-1} + \mathbf{C}_e^{-1}\right)^{-1} \in \mathcal{C}^{n_T n_R \times n_T n_R}. \tag{6.35}$$

Consequently, from (6.33) it is concluded that the actual channel $\mathbf{h}$ can be assumed to be in a region centered at $\mathbf{m}_{h|\widetilde{h}}$, i.e., at the conditional mean of the actual channel, also known as the MMSE Bayesian channel estimate [Kay93]. Based on this, $\widehat{\mathbf{h}} \triangleq \text{vec}(\widehat{\mathbf{H}})$ is defined as $\mathbf{m}_{h|\widetilde{h}}$ and, therefore, the error $\boldsymbol{\delta} \triangleq \text{vec}(\boldsymbol{\Delta})$ is equal to $\mathbf{h} - \mathbf{m}_{h|\widetilde{h}}$. According to these results, the uncertainty region for the error can be defined as an ellipsoid and, consequently, problem (6.12) is quadratic:

$$\mathcal{R} = \left\{\boldsymbol{\Delta}:\ \boldsymbol{\delta} = \text{vec}(\boldsymbol{\Delta}),\ \Pr\left(\boldsymbol{\delta}^H \mathbf{C}_{h|\widetilde{h}}^{-1}\boldsymbol{\delta} \leq r^2\right) = P_{\text{in}}\right\}. \tag{6.36}$$

Obviously, as the error $\boldsymbol{\delta}$ is Gaussian distributed, it will be inside the uncertainty region $\mathcal{R}$ (i.e., $\boldsymbol{\delta}^H \mathbf{C}_{h|\widetilde{h}}^{-1}\boldsymbol{\delta} \leq r^2$) with a certain probability P_{in} lower than 1. This probability will be equal to the probability of providing the required QoS to the user (i.e., the probability of having a SNR higher than the target SNR_0 or, equivalently, a BER lower than a maximum target BER_0). The mathematical relationship between the size of the uncertainty region, measured by r^2, and P_{in} is given by $r^2 = \phi^{-1}\left(P_{\text{in}}\right)$, where ϕ is the cdf of the chi-square distribution with $2n_R n_T$ degrees of freedom and normalized variance 1/2 (this result can be easily obtained taking into account the statistical distribution (6.33) and that the vector $\mathbf{C}_{h|\widetilde{h}}^{-1/2}\boldsymbol{\delta}$ is complex Gaussian distributed with zero-mean and covariance matrix $\mathbf{I}$).

For the concrete case where both the channel $\mathbf{H}$ and the error $\mathbf{E}$ matrices have i.i.d. components with zero-mean and variances σ_h^2 and σ_e^2, respectively, the uncertainty region for the channel reduces to a sphere of radius $\sqrt{\epsilon}$ centered at the Bayesian channel estimate, obtaining

$$\mathcal{R} = \left\{\boldsymbol{\Delta}:\ \|\boldsymbol{\Delta}\|_F^2 \leq \epsilon,\ \epsilon = r^2 \frac{\sigma_h^2}{1 + \text{SNR}_{\text{est}}}\right\}, \tag{6.37}$$

where $\text{SNR}_{\text{est}} \triangleq \sigma_h^2/\sigma_e^2$ is the received SNR during the transmission of the training sequence and r^2 is calculated as explained previously.

6.4.2 Quantization Errors

In FDD systems, the channel estimate at the transmitter has to be obtained through a feedback channel from the receiver to the transmitter. Since this feedback is expected to be discrete, the channel response has to be quantized introducing an error in the CSI available at the transmitter. Assuming that the receiver has a perfect knowledge of the channel response $\mathbf{H}$, it can quantize uniformly the real and imaginary parts of all the components of $\mathbf{H}$ using a quantization step equal to Δ_q and obtaining $\widehat{\mathbf{H}}$ as a result. Taking this parameter, the quantization SNR is defined as $\text{SNR}_q \triangleq 6\sigma_h^2/\Delta_q^2$, where σ_h^2 is the variance of each component of $\mathbf{H}$ assuming i.i.d. Gaussian components. Consequently, the uncertainty region for the channel can be defined as a hypercube centered at $\widehat{\mathbf{H}}$ and, therefore, $\mathcal{R}$ is defined as shown below, leading to a convex quadratic problem:

$$\mathcal{R} = \left\{ \boldsymbol{\Delta} : \ \left|\text{Re}\{[\boldsymbol{\Delta}]_{ij}\}\right| \leq \frac{\Delta_q}{2}, \ \left|\text{Im}\{[\boldsymbol{\Delta}]_{ij}\}\right| \leq \frac{\Delta_q}{2} \right\}. \tag{6.38}$$

Usually, the assumed dynamic range for the quantization is equal to 6 times the standard deviation. If this approach is taken, the relationship between SNR_q and the total number of bits N_b to be fed back is $N_b = n_T n_R \log_2(3\text{SNR}_q)$. Obviously, as the capacity of the feedback channel increases, more bits can be used in the quantization and, therefore, the size of the uncertainty region can be reduced.

6.4.3 Combined Estimation and Quantization Errors

In a realistic scenario, the two effects considered previously, i.e., the Gaussian noise from the estimation process and the quantization errors, are expected to be combined. This can be modeled mathematically by defining an appropriate uncertainty region for the error, which can be expressed as

$$\mathcal{R} = \left\{ \begin{array}{l} \boldsymbol{\Delta} = \boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2 : \|\boldsymbol{\Delta}_1\|_F^2 \leq \epsilon, \\ \qquad\qquad \left|\text{Re}\{[\boldsymbol{\Delta}_2]_{ij}\}\right| \leq \frac{\Delta_q}{2}, \ \left|\text{Im}\{[\boldsymbol{\Delta}_2]_{ij}\}\right| \leq \frac{\Delta_q}{2} \end{array} \right\} \tag{6.39}$$

and is convex. Note that, in the expression above, white Gaussian noise and uncorrelated MIMO channels have been considered, although the extension to ellipsoidal regions combined with quantization is direct. According to this region, the optimization problem (6.12) can be rewritten as the following quadratic problem:

$$\begin{array}{ll} \underset{t,\boldsymbol{\Delta}_1,\boldsymbol{\Delta}_2}{\text{minimize}} & t \\ \text{subject to} & t \geq P_0 \widehat{\mathbf{u}}_i^H \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2\right)^H \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2\right) \widehat{\mathbf{u}}_i, \qquad \forall i, \\ & \text{Tr}\left(\boldsymbol{\Delta}_1^H \boldsymbol{\Delta}_1\right) \leq \epsilon, \\ & \left|\text{Re}\{[\boldsymbol{\Delta}_2]_{ij}\}\right| \leq \frac{\Delta_q}{2}, \ \left|\text{Im}\{[\boldsymbol{\Delta}_2]_{ij}\}\right| \leq \frac{\Delta_q}{2}, \end{array} \tag{6.40}$$

which comprises the previous uncertainty regions and the corresponding optimization problems as particular cases.

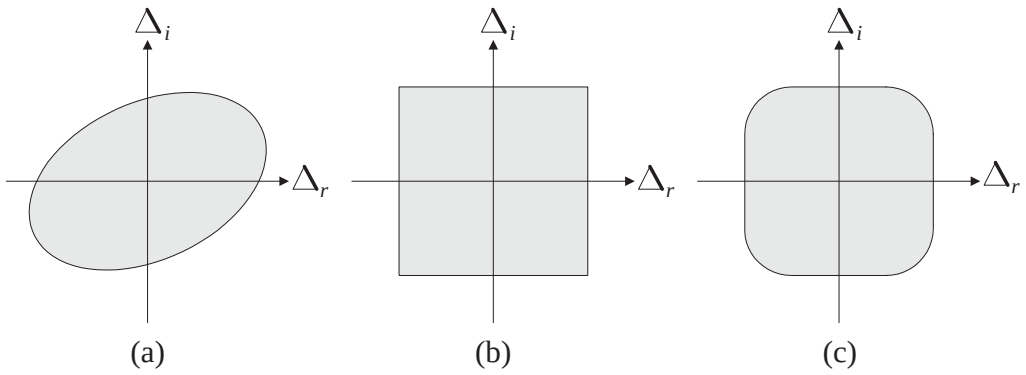


Figure 6.3: Different uncertainty regions for the case of a scalar error Δ, where $\Delta_r = \mathrm{Re}\{\Delta\}$ and $\Delta_i = \mathrm{Im}\{\Delta\}$. (a) estimation Gaussian noise, (b) quantization errors, and (c) combined estimation and quantization errors.

Figure 6.3 illustrates the shape of the three considered uncertainty regions for the concrete case of a scalar error Δ, where $\Delta_r = \mathrm{Re}\{\Delta\}$ and $\Delta_i = \mathrm{Im}\{\Delta\}$.

6.4.4 Other Uncertainty Regions

In addition to the previous uncertainty regions, there are many other possibilities and feedback strategies, whose error and imperfection models also lead to convex uncertainty regions. In the following, some of these examples are given, also including the previous ones:

1. **Spherical/ellipsoidal uncertainty regions:** represented by $\|\mathbf{\Delta}\|_F^2 \leq \epsilon$ and $\boldsymbol{\delta}^H \mathbf{C}_{h|\tilde{h}}^{-1} \boldsymbol{\delta} \leq r^2$, respectively. These regions correspond to the Gaussian noise from the estimation process.

2. **Quantization errors:** represented by $\left|\mathrm{Re}\{[\mathbf{\Delta}]_{ij}\}\right| \leq \frac{\Delta_q}{2}$, $\left|\mathrm{Im}\{[\mathbf{\Delta}]_{ij}\}\right| \leq \frac{\Delta_q}{2}$.

3. **Combination of Gaussian noise and quantization errors:** represented by $\mathbf{\Delta} = \mathbf{\Delta}_1 + \mathbf{\Delta}_2$, where $\|\mathbf{\Delta}_1\|_F^2 \leq \epsilon$, $\left|\mathrm{Re}\{[\mathbf{\Delta}_2]_{ij}\}\right| \leq \frac{\Delta_q}{2}$, $\left|\mathrm{Im}\{[\mathbf{\Delta}_2]_{ij}\}\right| \leq \frac{\Delta_q}{2}$.

All the previous regions make sense when the whole channel matrix $\mathbf{H}$ is estimated or each element of $\mathbf{H}$ is independently quantized, i.e., a scalar quantization is carried out. In the case of SIMO channels, a possible approach is to represent the channel in terms of the modes of the channel correlation matrix, as in the KL transform (see [Jai89] and references therein). Consider the channel $\mathbf{h} \in \mathcal{C}^{n_R \times 1}$, where this column vector represents the response of the SIMO channel with n_R receive antennas and fixed covariance matrix $\mathbb{E}[\mathbf{h}\mathbf{h}^H] = \mathbf{U}\mathbf{D}\mathbf{U}^H \in \mathcal{C}^{n_R \times n_R}$, where $\mathbf{U} \in \mathcal{C}^{n_R \times n_R}$ is the unitary matrix $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_{n_R}]$. Given this decomposition, the channel response can be coded taking the eigenvectors as the basis vectors for the representation,

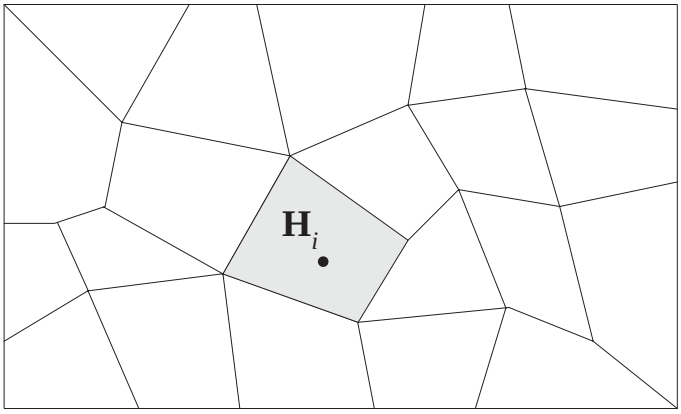


Figure 6.4: Uncertainty regions resulting from the intersection of halfspaces in vector quantization. In this example, the channel is within the indicated region and the index i is fed back to the transmitter.

obtaining: $\alpha_i = \mathbf{u}_i^H \mathbf{h}$, so that $\mathbf{h} = \sum_i \alpha_i \mathbf{u}_i$ and $\widehat{\mathbf{h}} = \sum_i \widehat{\alpha}_i \mathbf{u}_i$ (the eigenvectors are assumed to be known both at the transmitter and the receiver, and the only parameters that have to be fed back are the coefficients α_i).

In a general setup for a MIMO channel, this strategy could also be used and, therefore, the representation of the MIMO channel in terms of its coordinates $\{\alpha_i\}$ could be expressed as

$$\mathbf{H} = \sum_i \alpha_i \mathbf{H}_i, \tag{6.41}$$

where the matrices $\{\mathbf{H}_i\}$, which are fixed and known at both sides of the system, form the basis for the expression of the channel estimate. The estimated/quantized channel available at the transmitter is

$$\widehat{\mathbf{H}} = \sum_i \widehat{\alpha}_i \mathbf{H}_i, \tag{6.42}$$

where the error in the coefficients is represented by $\delta_i = \alpha_i - \widehat{\alpha}_i$. If the set defined for the error vector $\boldsymbol{\delta} = \begin{bmatrix}\delta_1 \cdots \delta_N\end{bmatrix}^T$ is convex, then the uncertainty region for the error in the channel estimate $\boldsymbol{\Delta} = \mathbf{H} - \widehat{\mathbf{H}} = \sum_i \delta_i \mathbf{H}_i$ is also convex, since the channel is expressed as a linear combination of the matrices of the basis multiplied by the coefficients $\{\alpha_i\}$ [Boy04].

The imperfections and uncertainty regions described previously can also be adopted for the coordinates $\{\alpha_i\}$ as:

4. **Spherical/ellipsoidal regions for the coordinates:** represented by $\sum_i w_i |\delta_i|^2 \leq r^2$.
5. **Quantization regions for the coordinates:** represented by $|\mathrm{Re}\{\delta_i\}| \leq \frac{\Delta_q}{2}$, $|\mathrm{Im}\{\delta_i\}| \leq \frac{\Delta_q}{2}$.
6. **Combination of 4 and 5.**

Previously, the quantization has been assumed to be a scalar quantization; however, a vector quantization is usually preferred. Consider a space with N points $\{\mathbf{H}_i\}$, each one representing the region given by $\mathbf{H}_i + \mathcal{R}_i$, i.e., if $\mathbf{H} \in \mathbf{H}_i + \mathcal{R}_i$, the ith index corresponding to $\mathbf{H}_i$ is sent (the number of bits for the feedback is equal to $\log_2(N)$). Each region $\mathcal{R}_i$ could be defined as a polyhedron obtained from the intersection of a finite number of halfspaces (see Figure 6.4 for an example). The region is then:

7. **Vector quantization:** $\widehat{\mathbf{H}} = \mathbf{H}_i$ and $\mathbf{\Delta} \in \mathcal{R}_i$, where i is the received index.

An example of vector quantization is shown in [Nar98], in which the Lloyd algorithm [Ger92] is applied to find a suboptimal region partition of the channel space maximizing the SNR in a MISO system. See also [Lov04], and references therein, for a general discussion on vector quantization and adequate criteria to define an appropriate feedback strategy in MIMO channels.

6.5 A Closed-Form Solution for Spherical Uncertainty Regions

The general convex problem presented in (6.12) can be extremely simplified for the case of spherical uncertainty regions. In this section, a closed-form solution is given for this case.

Proposition 2 *Consider the maximin problem (6.7) and the uncertainty region* $\mathcal{R} = \left\{\|\mathbf{\Delta}\|_F^2 \leq \epsilon\right\}$. *If* $\|\widehat{\mathbf{H}}\|_F^2 \leq \epsilon$, *then the saddle-value is 0 and no SNR can be guaranteed, i.e., no robust power allocation exists. Otherwise, the optimum robust power allocation is given by*

$$p_i^\star = \begin{cases} \mu\left(\sqrt{\widehat{\lambda}_i} - \alpha\right), & 1 \leq i \leq i_{\max}, \\ 0, & i > i_{\max}, \end{cases} \tag{6.43}$$

where μ *is a normalization factor such that* $\sum_{i=1}^{n_T} p_i^\star = P_0$, $i_{\max}$ *is the maximum index such that the following inequality is fulfilled:*

$$\sum_{i=1}^{i_{\max}} \left(\sqrt{\widehat{\lambda}_i} - \sqrt{\widehat{\lambda}_{i_{\max}}}\right)^2 < \epsilon, \tag{6.44}$$

and α *is the minimum solution to the following second degree equation:*

$$i_{\max}\alpha^2 - \left(2\sum_{i=1}^{i_{\max}} \sqrt{\widehat{\lambda}_i}\right)\alpha + \sum_{i=1}^{i_{\max}} \widehat{\lambda}_i - \epsilon = 0. \tag{6.45}$$

Proof. The concave-convex function f can be rewritten in terms of the matrices $\overline{\mathbf{H}}$ and $\overline{\mathbf{\Delta}}$, which are obtained by performing a linear transformation of the original channel estimate and error matrices using the unitary matrix of estimated eigenvectors:

$$\overline{\mathbf{H}} \triangleq \widehat{\mathbf{H}}\widehat{\mathbf{U}}, \quad \overline{\mathbf{\Delta}} \triangleq \mathbf{\Delta}\widehat{\mathbf{U}}, \tag{6.46}$$

and, therefore,

$$f\left(\mathbf{p},\overline{\boldsymbol{\Delta}}\right)=\mathrm{Tr}\left(\widehat{\mathbf{U}}^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)^H\left(\widehat{\mathbf{H}}+\boldsymbol{\Delta}\right)\widehat{\mathbf{U}}\,\mathrm{diag}\left(\mathbf{p}\right)\right)=\mathrm{Tr}\left(\left(\overline{\mathbf{H}}+\overline{\boldsymbol{\Delta}}\right)^H\left(\overline{\mathbf{H}}+\overline{\boldsymbol{\Delta}}\right)\mathrm{diag}\left(\mathbf{p}\right)\right). \tag{6.47}$$

Note that the uncertainty region defined for $\boldsymbol{\Delta}$ as $\|\boldsymbol{\Delta}\|_F^2\leq\epsilon$ can be equivalently written in terms of the transformed error $\overline{\boldsymbol{\Delta}}$ as $\overline{\mathcal{R}}=\left\{\overline{\boldsymbol{\Delta}}:\|\overline{\boldsymbol{\Delta}}\|_F^2\leq\epsilon\right\}$, since the multiplication by the unitary matrix $\widehat{\mathbf{U}}$ does not modify the value of the norm. Besides, the norms of the columns of the transformed channel matrix $\overline{\mathbf{H}}$ are related to the estimated eigenvalues by $\|\overline{\mathbf{h}}_i\|=\sqrt{\widehat{\lambda}_i}$.

The original maximin problem (6.7) can be rewritten as a minimax problem (the order of the inner and outer optimizations can be interchanged according to Lemma 2 and as used in the proof of Proposition 1, since a saddle-point of the problem exists) that can be formulated as

$$\begin{array}{ll}\underset{\overline{\boldsymbol{\Delta}}}{\text{minimize}} & \max\limits_i\ P_0\left[\left(\overline{\mathbf{H}}+\overline{\boldsymbol{\Delta}}\right)^H\left(\overline{\mathbf{H}}+\overline{\boldsymbol{\Delta}}\right)\right]_{ii}\\ \text{subject to} & \|\overline{\boldsymbol{\Delta}}\|_F^2\leq\epsilon.\end{array} \tag{6.48}$$

The elements of the diagonal can be written as $\left[\left(\overline{\mathbf{H}}+\overline{\boldsymbol{\Delta}}\right)^H\left(\overline{\mathbf{H}}+\overline{\boldsymbol{\Delta}}\right)\right]_{ii}=\left\|\overline{\mathbf{h}}_i+\overline{\boldsymbol{\delta}}_i\right\|^2$. Consider now the minimization with respect to each $\overline{\boldsymbol{\delta}}_i$, i.e., to each column of the matrix $\overline{\boldsymbol{\Delta}}$. The vector $\overline{\boldsymbol{\delta}}_i$ with norm $\|\overline{\boldsymbol{\delta}}_i\|=c_i$ that minimizes $\left\|\overline{\mathbf{h}}_i+\overline{\boldsymbol{\delta}}_i\right\|^2$ is $\overline{\boldsymbol{\delta}}_i^\star=-c_i\overline{\mathbf{h}}_i/\|\overline{\mathbf{h}}_i\|$. Using this result, the minimized ith component of the diagonal can be written as

$$\left[\left(\overline{\mathbf{H}}+\overline{\boldsymbol{\Delta}}^\star\right)^H\left(\overline{\mathbf{H}}+\overline{\boldsymbol{\Delta}}^\star\right)\right]_{ii}=\left\|\overline{\mathbf{h}}_i\left(1-\frac{1}{\|\overline{\mathbf{h}}_i\|}c_i\right)\right\|^2=\|\overline{\mathbf{h}}_i\|^2\left(1-\frac{1}{\|\overline{\mathbf{h}}_i\|}c_i\right)^2=\left(\|\overline{\mathbf{h}}_i\|-c_i\right)^2. \tag{6.49}$$

According to this, the problem (6.48) can be equivalently expressed as

$$\begin{array}{lll}\underset{t,\mathbf{c}}{\text{minimize}} & t & \\ \text{subject to} & t\geq P_0\left(\|\overline{\mathbf{h}}_i\|-c_i\right)^2, & \forall i,\\ & \sum_{i=1}^{n_T}c_i^2=\|\mathbf{c}\|^2\leq\epsilon, & \end{array} \tag{6.50}$$

where $\mathbf{c}=\left[c_1\cdots c_{n_T}\right]^T\in\mathbb{R}^{n_T\times 1}$. Note that the constraint $\|\overline{\boldsymbol{\Delta}}\|_F^2\leq\epsilon$ can be equivalently written in terms of the scalars $\{c_i\}$ as $\sum_{i=1}^{n_T}c_i^2\leq\epsilon$.

It is now clear how to find the optimum solution to this problem according to the points below, taking into account that the eigenvalues $\{\widehat{\lambda}_i\}$ are sorted in decreasing order and that $\|\overline{\mathbf{h}}_i\|=\sqrt{\widehat{\lambda}_i}$:

1. *Detection of saddle-value equal to 0:* in case that $\sum_{i=1}^{n_T}\|\overline{\mathbf{h}}_i\|^2=\|\overline{\mathbf{H}}\|_F^2=\|\widehat{\mathbf{H}}\|_F^2\leq\epsilon$, then the worst-error is attained when $c_i^\star=\|\overline{\mathbf{h}}_i\|$, which corresponds to $\boldsymbol{\Delta}^\star=-\widehat{\mathbf{H}}$, and the saddle-value $t^\star$ is 0, which means that no SNR can be guaranteed for any power allocation.

2. *Detection of the number of active eigenmodes:* when the saddle-value is different from 0, the optimum solution corresponds to the case in which some of the inequality constraints

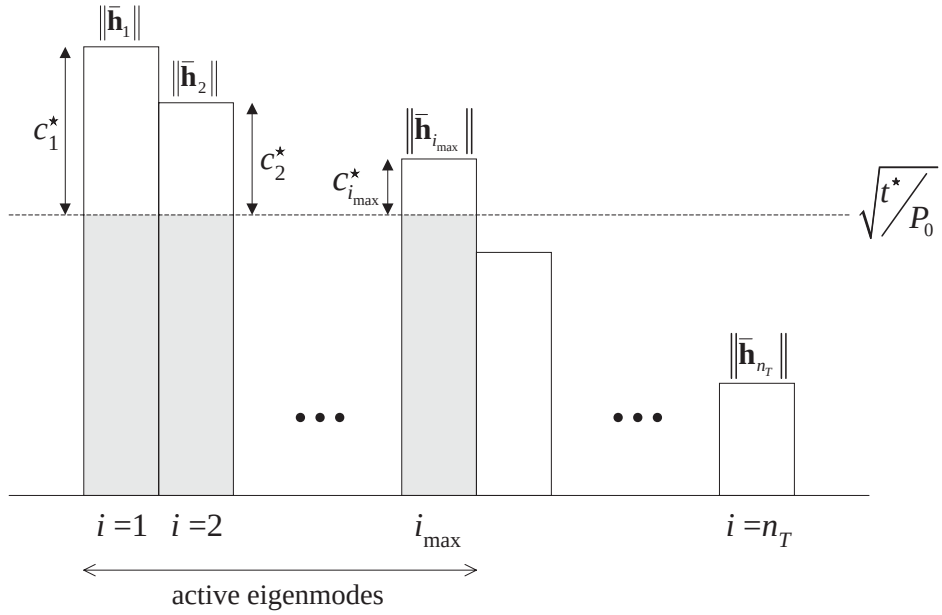


Figure 6.5: Representation of the optimal solution for the case of a spherical uncertainty region.

in (6.50) are fulfilled with equality, whereas the value of $c_i^\star$ for the other constraints is 0. This is the optimum solution, since, in case that there exists an index j such that $c_j > 0$ but the corresponding constraint is not fulfilled with equality (i.e., $t > P_0\left(\|\overline{\mathbf{h}}_j\| - c_j\right)^2$), then the value of t can be reduced by decreasing the value of c_j and increasing c_i, $\forall i \neq j$, while still fulfilling $\|\mathbf{c}\|^2 \leq \epsilon$. Consequently, the optimum solution is attained when $\|\overline{\mathbf{h}}_i\| - c_i^\star$ is constant for the active constraints and the value of t cannot be further reduced since $\|\mathbf{c}^\star\|^2 = \epsilon$ (see Figure 6.5 for an example of an optimal solution). Taking all this into account, the number of active constraints is the maximum index $i_{\max}$ such that the following inequality is fulfilled (this is obtained by taking into account that the constraints are activated in increasing order in Figure 6.5 and by looking for the maximum number of active constraints using $\|\mathbf{c}^\star\|^2 = \epsilon$):

$$\sum_{i=1}^{i_{\max}} \left(\|\overline{\mathbf{h}}_i\| - \|\overline{\mathbf{h}}_{i_{\max}}\|\right)^2 < \epsilon. \tag{6.51}$$

Once this index has been calculated by a simple finite iteration, the optimum values of the coefficients $\{c_i\}$ for the active constraints can be expressed as a function of the constant for the last active constraint $c_{i_{\max}}$ (note that for the other constraints, the optimum constants are equal to 0, i.e., $c_i^\star = 0,\ i > i_{\max}$):

$$\|\overline{\mathbf{h}}_i\| - c_i^\star = \|\overline{\mathbf{h}}_{i_{\max}}\| - c_{i_{\max}}^\star \quad \Rightarrow \quad c_i^\star = \|\overline{\mathbf{h}}_i\| - \|\overline{\mathbf{h}}_{i_{\max}}\| + c_{i_{\max}}^\star, \qquad i = 1, \ldots, i_{\max}. \tag{6.52}$$

Using this result, $c_{i_{\max}}^\star$ can be easily calculated by taking the positive solution to the

following second degree equation resulting from the constraint $\sum_{i=1}^{n_T} c_i^2 = \epsilon$:

$$\sum_{i=1}^{n_T} c_i^{\star 2} = \sum_{i=1}^{i_{\max}} \left(\|\overline{\mathbf{h}}_i\| - \|\overline{\mathbf{h}}_{i_{\max}}\| + c_{i_{\max}}^{\star} \right)^2 = \epsilon \quad \Rightarrow$$

$$i_{\max} \left(c_{i_{\max}}^{\star} \right)^2 + 2 \sum_{i=1}^{i_{\max}} \left(\|\overline{\mathbf{h}}_i\| - \|\overline{\mathbf{h}}_{i_{\max}}\| \right) c_{i_{\max}}^{\star} + \sum_{i=1}^{i_{\max}} \left(\|\overline{\mathbf{h}}_i\| - \|\overline{\mathbf{h}}_{i_{\max}}\| \right)^2 - \epsilon = 0. \quad (6.53)$$

Collecting all these results, the worst-case error $\boldsymbol{\Delta}^{\star}$ can be finally calculated as

$$\boldsymbol{\Delta}^{\star} = \left[\begin{array}{ccc} -c_1^{\star} \frac{\overline{\mathbf{h}}_1}{\|\overline{\mathbf{h}}_1\|} & \cdots & -c_{n_T}^{\star} \frac{\overline{\mathbf{h}}_{n_T}}{\|\overline{\mathbf{h}}_{n_T}\|} \end{array} \right] \widehat{\mathbf{U}}^H. \quad (6.54)$$

The optimal robust power allocation can be obtained as the power $\mathbf{p}^{\star}$ such that the worst-case error is a solution to $\min_{\boldsymbol{\Delta}} f\left(\mathbf{p}^{\star}, \boldsymbol{\Delta}\right)$, i.e., $\boldsymbol{\Delta}^{\star}$ satisfies the corresponding KKT conditions. The Lagrangian associated to the problem is

$$L(\boldsymbol{\Delta}; \mu) \triangleq \mathrm{Tr}\left(\widehat{\mathbf{U}}^H \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta} \right)^H \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta} \right) \widehat{\mathbf{U}} \, \mathrm{diag}(\mathbf{p}^{\star}) \right) + \mu \left(\mathrm{Tr}\left(\boldsymbol{\Delta}^H \boldsymbol{\Delta} \right) - \epsilon \right) \quad (6.55)$$

and, therefore, one of the KKT conditions is

$$\left(\widehat{\mathbf{H}} + \boldsymbol{\Delta}^{\star} \right) \widehat{\mathbf{U}} \, \mathrm{diag}(\mathbf{p}^{\star}) \widehat{\mathbf{U}}^H + \mu^{\star} \boldsymbol{\Delta}^{\star} = \mathbf{0} \quad \Rightarrow \quad \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta}^{\star} \right) \widehat{\mathbf{U}} \, \mathrm{diag}(\mathbf{p}^{\star}) + \mu^{\star} \boldsymbol{\Delta}^{\star} \widehat{\mathbf{U}} = \mathbf{0}, \quad (6.56)$$

$$\left(\overline{\mathbf{H}} + \overline{\boldsymbol{\Delta}^{\star}} \right) \mathrm{diag}(\mathbf{p}^{\star}) + \mu^{\star} \overline{\boldsymbol{\Delta}^{\star}} = \mathbf{0}, \quad (6.57)$$

which has to be satisfied at the worst-case error $\boldsymbol{\Delta}^{\star}$ and for the robust power allocation $\{p_i^{\star}\}$. From this equation, the power to be allocated to the ith estimated eigenmode can be calculated as

$$\left(\overline{\mathbf{h}}_i + \overline{\boldsymbol{\delta}}_i^{\star} \right) p_i^{\star} + \mu^{\star} \overline{\boldsymbol{\delta}}_i^{\star} = 0 \Rightarrow \overline{\mathbf{h}}_i \left(1 - \frac{c_i^{\star}}{\|\overline{\mathbf{h}}_i\|} \right) p_i^{\star} - \mu^{\star} \overline{\mathbf{h}}_i \frac{c_i^{\star}}{\|\overline{\mathbf{h}}_i\|} = 0$$

$$\Rightarrow p_i^{\star} = \mu^{\star} \frac{\frac{c_i^{\star}}{\|\overline{\mathbf{h}}_i\|}}{1 - \frac{c_i^{\star}}{\|\overline{\mathbf{h}}_i\|}} = \mu^{\star} \frac{c_i^{\star}}{\sqrt{\widehat{\lambda}_i} - c_i^{\star}} = \begin{cases} \mu^{\star} \frac{\sqrt{\widehat{\lambda}_i} - \sqrt{\widehat{\lambda}_{i_{\max}}} + c_{i_{\max}}^{\star}}{\sqrt{\widehat{\lambda}_{i_{\max}}} - c_{i_{\max}}^{\star}}, & 1 \leq i \leq i_{\max}, \\ 0, & i > i_{\max}, \end{cases}$$

where $\mu^{\star}$ is a normalization factor such that $\sum_{i=1}^{n_T} p_i^{\star} = P_0$. By defining the constant $\alpha = \sqrt{\widehat{\lambda}_{i_{\max}}} - c_{i_{\max}}^{\star}$, absorbing the factor $1/\alpha$ in $\mu^{\star}$, and introducing the change of variable $c_{i_{\max}}^{\star} = \sqrt{\widehat{\lambda}_{i_{\max}}} - \alpha$ in (6.53), the results in Proposition 2 are directly obtained, including the equation (6.45). ■

6.6 Applications and Extensions

In this section, some applications and extensions of the already presented design strategy are presented.

6.6.1 Minimum Transmit Power with an Instantaneous Performance Constraint

From the KKT conditions (6.19)-(6.22) for the reformulated convex problem (6.12), it can be shown that the optimal dual variables $\{\gamma_i^\star\}$ and the worst-case error $\mathbf{\Delta}^\star$ do not depend on the power budget P_0. Consequently, the optimum robust power allocation scales linearly with P_0, and also the saddle-value, which is given by

$$f^\star \triangleq f\left(\mathbf{p}^\star, \mathbf{\Delta}^\star\right) = P_0 \operatorname{Tr}\left(\widehat{\mathbf{U}}^H \left(\widehat{\mathbf{H}} + \mathbf{\Delta}^\star\right)^H \left(\widehat{\mathbf{H}} + \mathbf{\Delta}^\star\right) \widehat{\mathbf{U}} \operatorname{diag}\left(\{\gamma_i^\star\}\right)\right). \tag{6.58}$$

This result can be used to calculate the solution to the problem consisting in minimizing the transmit power, while still guaranteeing that the instantaneous performance, in terms of the SNR, is better than a minimum target SNR_0 for any error in the uncertainty region. This problem is the complementary to that solved previously, in which the performance was optimized subject to a power constraint, although both problems are essentially equivalent. The solution to this new problem is also attained by taking the robust power allocation given by $\{\gamma_i^\star\}$, where the required transmit power is calculated as

$$P_0^\star = \mathrm{SNR}_0 \frac{\sigma_n^2}{\operatorname{Tr}\left(\widehat{\mathbf{U}}^H \left(\widehat{\mathbf{H}} + \mathbf{\Delta}^\star\right)^H \left(\widehat{\mathbf{H}} + \mathbf{\Delta}^\star\right) \widehat{\mathbf{U}} \operatorname{diag}\left(\{\gamma_i^\star\}\right)\right)}. \tag{6.59}$$

6.6.2 Application to Adaptive Modulation with Maximum BER Constraints

The previous robust design can be combined with AM strategies [Cav72, Web95, Gol97, Chu01] to maximize the throughput subject to BER constraints, i.e., the objective is to maximize the transmission rate by employing high level modulations while still guaranteeing a minimum quality in terms of a maximum BER for any possible error in the uncertainty region.

Consider that the transmit power is bounded by $P_0^{\max}$ and let $\mathrm{BER}_L(\mathrm{SNR})$ be the function that relates the SNR with the BER for a L-QAM modulation. Obviously, given a certain SNR, the BER increases as the number of levels L in the modulation also grows. Taking this into account, the proposed robust AM is based on the following steps, in which all the symbols are assumed to be taken from the same signal constellation:

1. Define the desired QoS in terms of a maximum allowed BER_0.

2. Given the channel estimate $\widehat{\mathbf{H}}$ and the uncertainty region $\mathcal{R}$, calculate the robust normalized power allocation given by the optimum Lagrange multipliers $\{\gamma_i^\star\}$ and the worst-case error given by the optimum primal variables $\mathbf{\Delta}^\star$.

3. Calculate the maximum achievable SNR for any error in the uncertainty region similarly as in (6.59):

$$\mathrm{SNR}^{\max} = \frac{P_0^{\max}}{\sigma_n^2} \mathrm{Tr}\left(\widehat{\mathbf{U}}^H \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta}^\star\right)^H \left(\widehat{\mathbf{H}} + \boldsymbol{\Delta}^\star\right) \widehat{\mathbf{U}} \,\mathrm{diag}\left(\{\gamma_i^\star\}\right)\right). \tag{6.60}$$

4. Calculate the maximum number of levels $L^\star$ fulfilling $\mathrm{BER}_L(\mathrm{SNR}^{\max}) \leq \mathrm{BER}_0$. This can be trivially done by using a look-up table in which the values of the SNR required for each BER and signal constellation size are saved. If the previous constraint cannot be fulfilled for any value of L, set $L^\star = 0$. In this case, no signal is transmitted since the QoS requirement cannot be satisfied for all the possible errors in the uncertainty region while still fulfilling the maximum transmit power constraint.

5. For the selected value $L^\star$, calculate the necessary instantaneous transmit power as in (6.59). Note that the transmit power required to fulfill the BER constraint with equality may be lower than the maximum available transmit power $P_0^{\max}$ since the number of modulation levels L is discrete.

Summarizing, this algorithm proposes a robust AM technique, in which the throughput is maximized while a certain QoS can be guaranteed to the user given a channel estimate and a transmit power constraint.

6.6.3 Extension to Scenarios with Interferences

Previously, no interferences have been assumed at the receiver. In case that interferences have to be taken into account in the system, the previous design strategy can also be applied, as explained in the following.

Let $\mathbf{R}_n$ be the noise plus interferences correlation matrix, as defined in §3.2.3 (see (3.15)). The optimum receiver maximizing the SNIR is based on a matched filter, leading to

$$\mathrm{SNIR} = \mathrm{Tr}\left(\underline{\widehat{\mathbf{U}}}^H \mathbf{H}^H \mathbf{R}_n^{-1} \mathbf{H} \underline{\widehat{\mathbf{U}}} \,\mathrm{diag}(\mathbf{p})\right) = \mathrm{Tr}\left(\underline{\widehat{\mathbf{U}}}^H \underline{\mathbf{H}}^H \underline{\mathbf{H}} \underline{\widehat{\mathbf{U}}} \,\mathrm{diag}(\mathbf{p})\right), \tag{6.61}$$

where $\underline{\mathbf{H}}$ is the whitened channel matrix defined as $\underline{\mathbf{H}} = \mathbf{R}_n^{-1/2}\mathbf{H}$, and $\underline{\widehat{\mathbf{U}}}$ is the unitary matrix containing the estimates of the eigenvectors of $\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H} = \underline{\mathbf{H}}^H\underline{\mathbf{H}}$, i.e., the eigenvectors of $\widehat{\mathbf{H}}^H\widehat{\mathbf{R}}_n^{-1}\widehat{\mathbf{H}} = \underline{\widehat{\mathbf{H}}}^H\underline{\widehat{\mathbf{H}}}$, where $\widehat{\mathbf{H}}$ and $\widehat{\mathbf{R}}_n$ are the estimates of the channel and correlation matrices, respectively. Note that the previous expression of the SNIR is equal to the one corresponding to a system with no interferences (6.5), but including the matrix $\mathbf{R}_n^{-1}$. Based on this, the following performance function to be maximized can be defined:

$$f\left(\mathbf{p}, \underline{\boldsymbol{\Delta}}\right) = \mathrm{Tr}\left(\underline{\widehat{\mathbf{U}}}^H \left(\underline{\widehat{\mathbf{H}}} + \underline{\boldsymbol{\Delta}}\right)^H \left(\underline{\widehat{\mathbf{H}}} + \underline{\boldsymbol{\Delta}}\right) \underline{\widehat{\mathbf{U}}} \,\mathrm{diag}(\mathbf{p})\right), \tag{6.62}$$

where $\underline{\mathbf{\Delta}}$ is the error in the estimate of $\underline{\mathbf{H}}$, i.e.,

$$\underline{\mathbf{H}} = \underline{\widehat{\mathbf{H}}} + \underline{\mathbf{\Delta}}. \tag{6.63}$$

According to these definitions, the design strategy presented previously can be applied directly assuming that the error $\underline{\mathbf{\Delta}}$ belongs to a convex uncertainty region $\underline{\mathcal{R}}$.

In case that the estimate of the correlation matrix $\widehat{\mathbf{R}}_n$ is perfect, i.e., $\widehat{\mathbf{R}}_n = \mathbf{R}_n$, then the uncertainty region $\underline{\mathcal{R}}$ for the error in the estimate of the whitened channel matrix $\underline{\mathbf{H}}$ can be written in terms of the uncertainty region $\mathcal{R}$ for the error in the estimate of the channel matrix $\mathbf{H}$ as

$$\underline{\mathcal{R}} = \left\{\underline{\mathbf{\Delta}} = \mathbf{R}_n^{-1/2}\mathbf{\Delta} : \mathbf{\Delta} \in \mathcal{R}\right\}, \tag{6.64}$$

which is also convex if $\mathcal{R}$ is convex, since the transformed error $\underline{\mathbf{\Delta}}$ is obtained by a linear transformation of $\mathbf{\Delta}$.

6.6.4 Extension to OFDM Modulations

In the previous sections in this chapter, a single-carrier MIMO flat fading channel has been assumed. In many communication systems, however, the delay spread of the channel can be high when compared with the symbol duration, leading to frequency selective channels. Consequently, the previous results should be extended to these scenarios.

A common approach when the channel is frequency selective is to apply the OFDM modulation, as commented in previous chapters [Bin90, Nee98, Wan00]. The objective in this section is to extend the results presented previously to the case of using a multicarrier modulation in a frequency selective MIMO channel. Let us consider a modulation consisting of N carriers, where $k = 0, \ldots, N-1$ is the carrier index. For each subcarrier, a channel estimate is available at the transmitter $\widehat{\mathbf{H}}_k$, where the error is $\mathbf{\Delta}_k = \mathbf{H}_k - \widehat{\mathbf{H}}_k$. The errors at each carrier will be assumed to be independent and an uncertainty region $\mathcal{R}_k$ will be defined for each of them ($\mathbf{\Delta}_k \in \mathcal{R}_k$).

Let P_0 be the total transmit power to be distributed among the N carriers of the OFDM modulation. The power allocated to the kth carrier is represented by P_k. Obviously, for a concrete power distribution among the different carriers, the maximin robust design presented previously can be applied to each carrier as shown in Figure 6.6, obtaining the robust power distribution $\left\{p_{k,i}^{\star}\right\}$. In the figure, $\widehat{\mathbf{u}}_{k,i}$ is the ith estimated eigenvector at the kth carrier, i.e., the unitary matrix $\widehat{\mathbf{U}}_k = \left[\widehat{\mathbf{u}}_{k,1} \cdots \widehat{\mathbf{u}}_{k,n_T}\right] \in \mathcal{C}^{n_T \times n_T}$ contains the n_T eigenvectors of $\widehat{\mathbf{H}}_k^H \widehat{\mathbf{H}}_k$ sorted in decreasing order, and $p_{k,i}$ represents the power allocated to $\widehat{\mathbf{u}}_{k,i}$. Since P_k is the total power allocated to the kth carrier, the following constraint holds: $\sum_{i=1}^{n_T} p_{k,i} \leq P_k$. According to this, the saddle-value for the kth carrier, i.e., the minimum SNR that can be guaranteed for any error in the uncertainty region $\mathcal{R}_k$ is

$$P_k \beta_k, \tag{6.65}$$

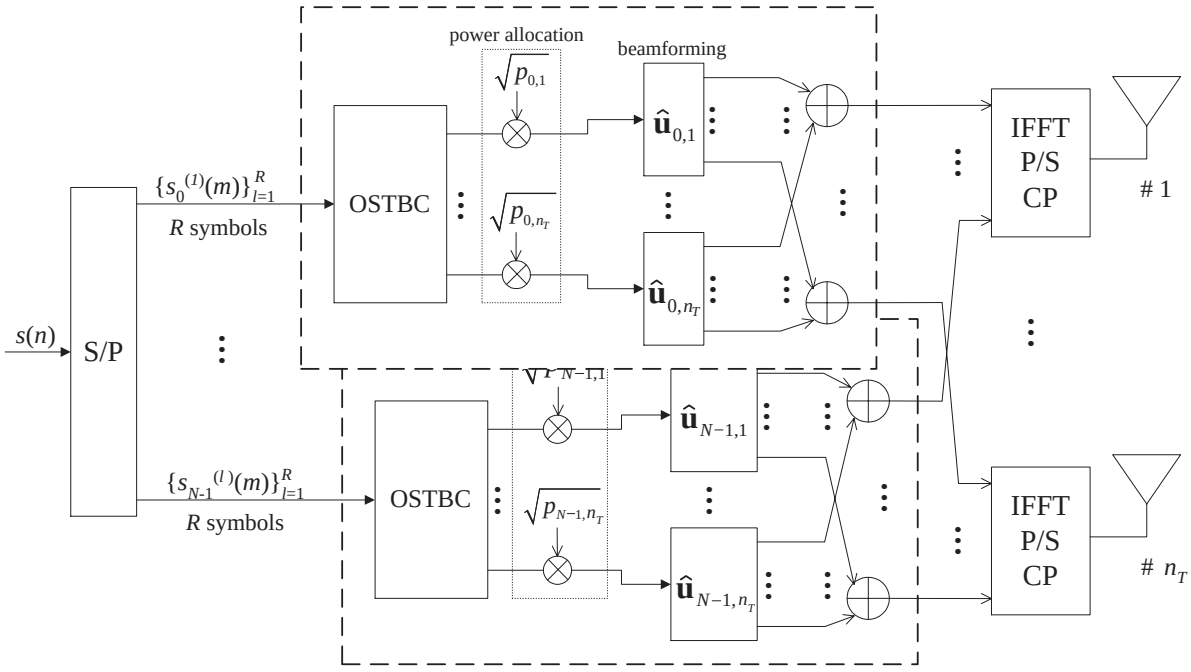


Figure 6.6: Transmitter architecture corresponding to a robust maximin design combined with the use of the OFDM modulation.

where

$$\beta_k = \frac{1}{\sigma_n^2} \mathrm{Tr}\left(\widehat{\mathbf{U}}_k^H \left(\widehat{\mathbf{H}}_k + \boldsymbol{\Delta}_k^\star\right)^H \left(\widehat{\mathbf{H}}_k + \boldsymbol{\Delta}_k^\star\right) \widehat{\mathbf{U}}_k \,\mathrm{diag}\left(\{\gamma_{k,i}^\star\}\right)\right). \tag{6.66}$$

Note that β_k does not depend on the power P_k allocated to the carrier.

A design strategy has still to be defined for the power distribution among the carriers $\{P_k\}$ subject to the total transmit power constraint $\sum_{k=0}^{N-1} P_k = P_0$. One possibility consists in maximizing the worst saddle-value for all the carriers, i.e., the MAXMIN strategy, presented in §3.3.3 could also be applied over the performance of all the carriers. According to this strategy, the power distribution is found as

$$P_k = \frac{P_0}{\sum_{j=0}^{N-1} \frac{1}{\beta_j}} \frac{1}{\beta_k}, \tag{6.67}$$

which leads to a solution in which all the carriers have the same saddle-value, i.e., the minimum guaranteed SNR is the same for all the carriers.

6.7 Simulation Results

In this section, several simulation results are presented to show the robustness capabilities of the already presented technique and compare its performance with other classical solutions, such as the non-robust approach and the pure OSTBC approach.

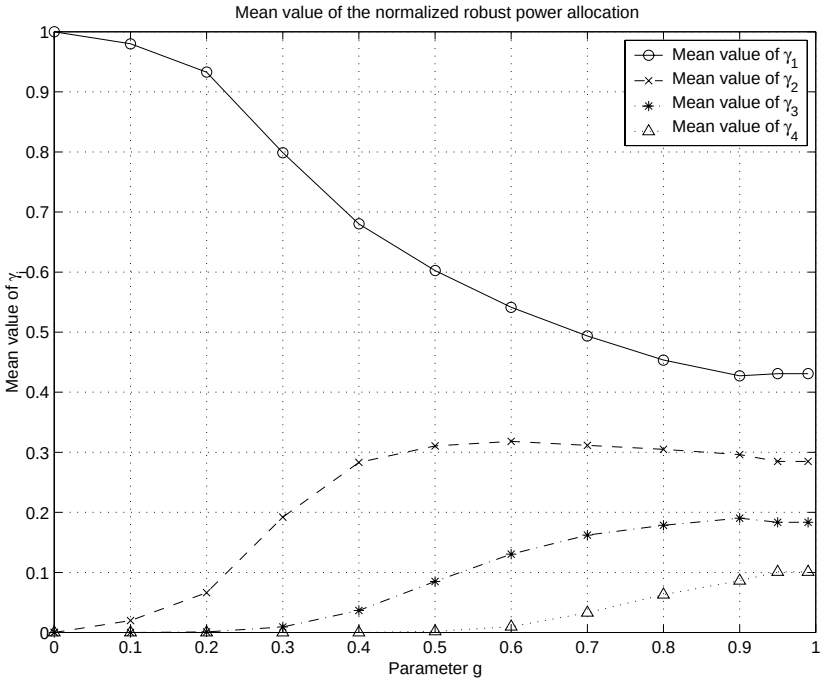


Figure 6.7: Mean value of the robust power distribution for different sizes of the uncertainty region.

As presented in §6.3, the robust maximin technique distributes the available power among the estimated eigenmodes taking into account the errors in the channel estimate. Obviously, if the channel estimate is perfect, the robust solution should be equal to the non-robust beamforming, i.e., to the power allocation formulated as $p_1 = P_0, \;\; p_i = 0, \;\; i = 2, \ldots, n_T$. When the uncertainty in the actual channel increases, the robust design tends to distribute the power in a more uniform way.

In the first simulations, a system with 4 transmit and 6 receive antennas is analyzed. The noise in the channel estimate is assumed to be Gaussian and spherical uncertainty regions with a radius equal to $\sqrt{\epsilon} = g\|\widehat{\mathbf{H}}\|_F, \;\; 0 \leq g < 1$ are considered. Note that for these uncertainty regions, $\mathbf{H} = \widehat{\mathbf{H}} + \mathbf{\Delta} \neq \mathbf{0}, \; \forall \mathbf{\Delta} \in \mathcal{R}$. This condition has to hold since, otherwise, the saddle-value would be equal to 0.

Since $n_T = 4$, the total transmit power has to be distributed among the 4 estimated eigenvectors. In Figure 6.7, the mean value of the normalized robust power allocation $\{\gamma_i^\star\}$ is shown as a function of g. As can be seen, for $g = 0$ the power distribution corresponds to the non-robust approach, as expected. As g increases, the power allocation profile changes and tends to distribute the power in a more uniform way. Note that the pure OSTBC approach is equivalent to a uniform power allocation $p_i = 1/4 = 0.25, \; 1 \leq i \leq 4$. As can be shown in Figure 6.7, this uniform distribution is not attained by the robust approach, even when $g \to 1$. In [Gan01], it

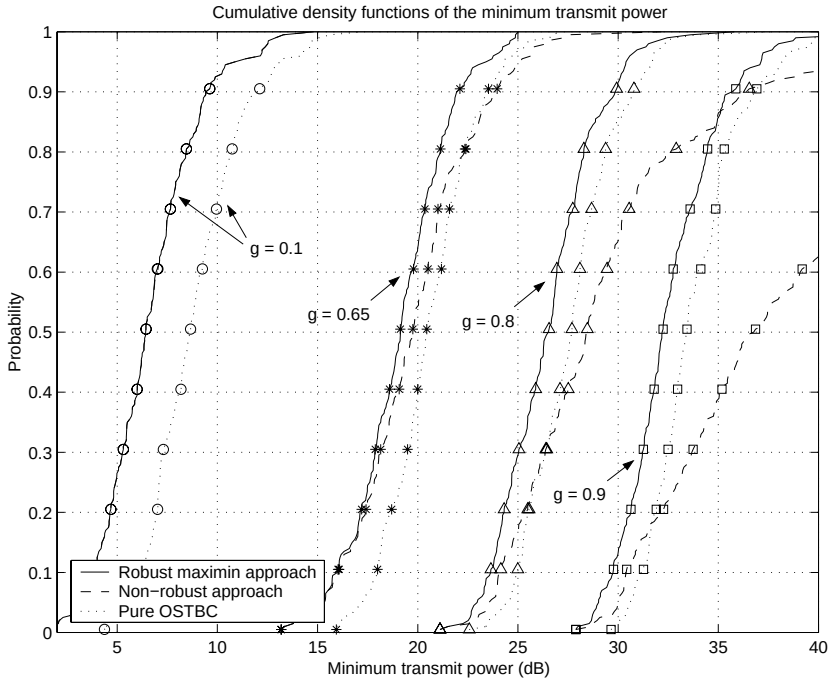


Figure 6.8: Cumulative density functions of the minimum required transmit power assuming different sizes for the spherical uncertainty regions, and according to a target SNR equal to $\text{SNR}_0 = 10$ dB.

is deduced that OSTBC is optimum in a robust maximin sense when the channel is totally unknown. This can be expressed mathematically as the following uncertainty region for the actual channel: $\widetilde{\mathcal{R}} = \left\{\mathbf{H} : \|\mathbf{H}\|_F^2 \geq \rho\right\}$, where ρ is a positive real value that avoids the channel $\mathbf{H} = \mathbf{0}$ to belong to the uncertainty region. Note that, when $g = 1$, the spherical uncertainty region is different from that for which OSTBC is the robust maximin solution and, consequently, the robust power allocation policy for $g = 1$ does not have to be necessarily uniform.

As stated in §6.6.1, the maximin design can be used to guarantee a minimum target SNR_0 with the minimum required transmit power for any error in the uncertainty region. In Figure 6.8, the cdf of the minimum required transmit power is shown for SNR_0=10 dB. This cdf is represented for $n_T = 2$ and $n_R = 2$ and for three different transmission techniques: the robust approach, the non-robust classical beamforming, and a pure OSTBC strategy. The uncertainty regions that are considered are spherical taking four different values for the parameter g: 0.1, 0.65, 0.8, and 0.9. As can be seen, for small uncertainty regions, both the robust and the non-robust approaches have a similar performance and need less transmit power than OSTBC, as expected. When the size of the uncertainty region increases, the non-robust approach increases the necessary transmit power to fulfill the QoS requirements. Note that, for an extreme case corresponding to big uncertainty regions, the non-robust technique may need even more power than OSTBC. This means that in case that the CSI may have high errors, it is more convenient

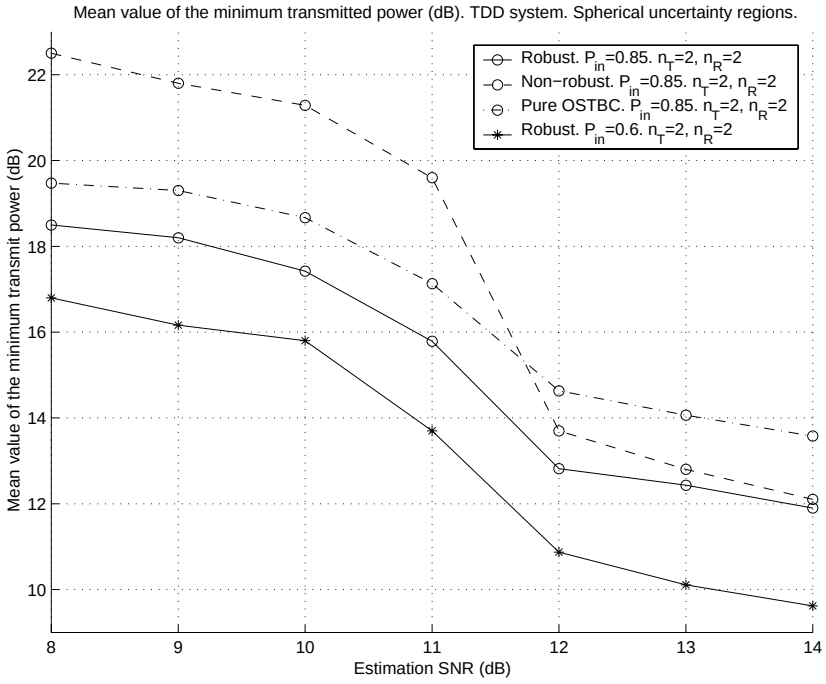


Figure 6.9: Mean value of the minimum required transmit power in a TDD system (Gaussian noise, spherical uncertainty regions) according to a target SNR equal to $\mathrm{SNR}_0 = 10$ dB.

to use an OSTBC approach than to assume that the channel estimate is perfect, despite not being true. Note also that, in all cases, the robust solution is the technique requiring the least transmit power.

Figures 6.9 and 6.10 show some results on the mean value of the minimum required transmit power to attain a SNR_0=10 dB when considering spherical (TDD systems) and cubic (FDD systems) uncertainty regions. The sizes of the regions are directly related to the estimation and quantization SNR, as deduced in §6.4. For the case of spherical uncertainty regions, two different QoS probabilities (as defined in subsection 6.4.1) have been used: $P_{\mathrm{in}} = 0.85$ and $P_{\mathrm{in}} = 0.6$. The same conclusions can be obtained from the observation of both figures. If the estimation or quantization SNR is high, OSTBC needs more power than the non-robust and the robust designs, since it does not exploit the channel knowledge available at the transmitter. As the estimation or quantization SNR decreases, all the techniques need more power to fulfill the instantaneous SNR requirements, since the size of the uncertainty region increases. Note that for all the cases, the technique requiring the least transmit power is the robust approach. Also, as commented previously, if the estimation or quantization SNR is low enough, the non-robust solution needs more power than OSTBC, concluding that in case of having a very low quality channel estimate, it is not convenient to use it without taking into account explicitly the errors in the estimate, i.e., in a non-robust way. In the case of spherical regions, increasing the probability

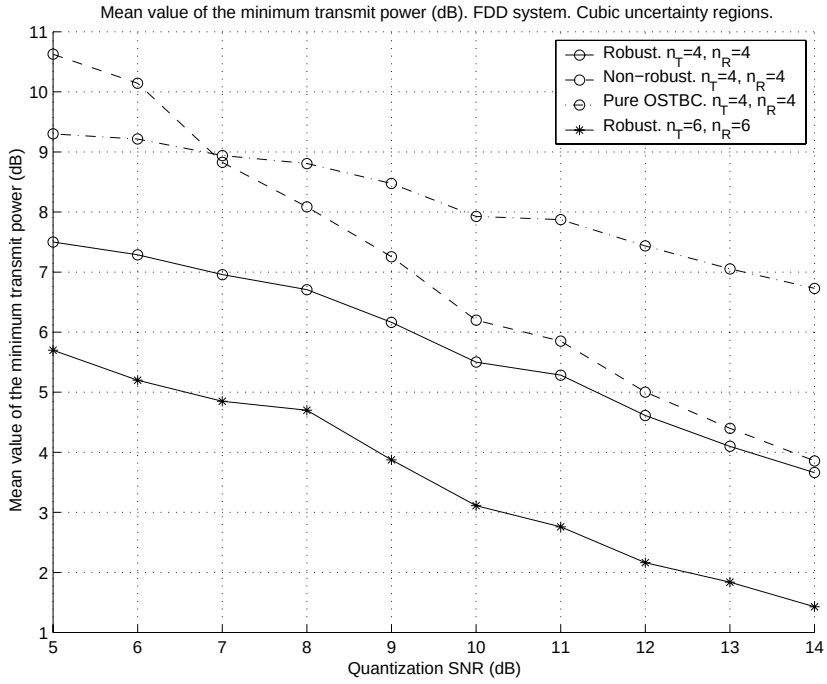


Figure 6.10: Mean value of the minimum required transmit power in a FDD system (quantization errors, cubic uncertainty regions) according to a target SNR equal to $\text{SNR}_0 = 10$ dB.

P_{in} of providing a QoS implies an increase of the minimum required transmit power, since the size of the uncertainty region also increases. From the figures, it is also concluded that very important savings in terms of transmit power can be obtained when using the robust approach instead of the non-robust beamforming. For example, for $n_T = 2$ and $n_R = 2$ and $\text{SNR}_{\text{est}} = 11$ dB, a saving of almost 4 dB can be obtained, whereas for $n_T = 4$ and $n_R = 4$ and $\text{SNR}_q = 5$ dB, a saving of 3 dB is achieved.

As explained at the beginning of this section, an instantaneous SNR can be guaranteed to the user only in case that $\mathbf{H} = \mathbf{0}$ does not belong to the uncertainty region for the actual channel. In Figures 6.11 and 6.12 the *service provision probability* is shown, i.e., $\Pr\left(\widehat{\mathbf{H}} + \boldsymbol{\Delta} \neq \mathbf{0}, \forall \boldsymbol{\Delta} \in \mathcal{R}\right)$, as a function of the QoS probability P_{in} required by the user in the case of TDD, and the quantization SNR in the case of FDD. These results have been obtained for different number of transmit and receive antennas. As a general conclusion, it can be observed that increasing the number of antennas and the quantization SNR implies an increase of the service provision probability, as expected. On the other hand, if the user demands a higher QoS probability, the service provision probability decreases, since an increase of the required QoS implies an increase of the uncertainty region $\mathcal{R}$ and, therefore, it is not always possible to guarantee that QoS since the actual channel $\mathbf{H} = \mathbf{0}$ may be possible.

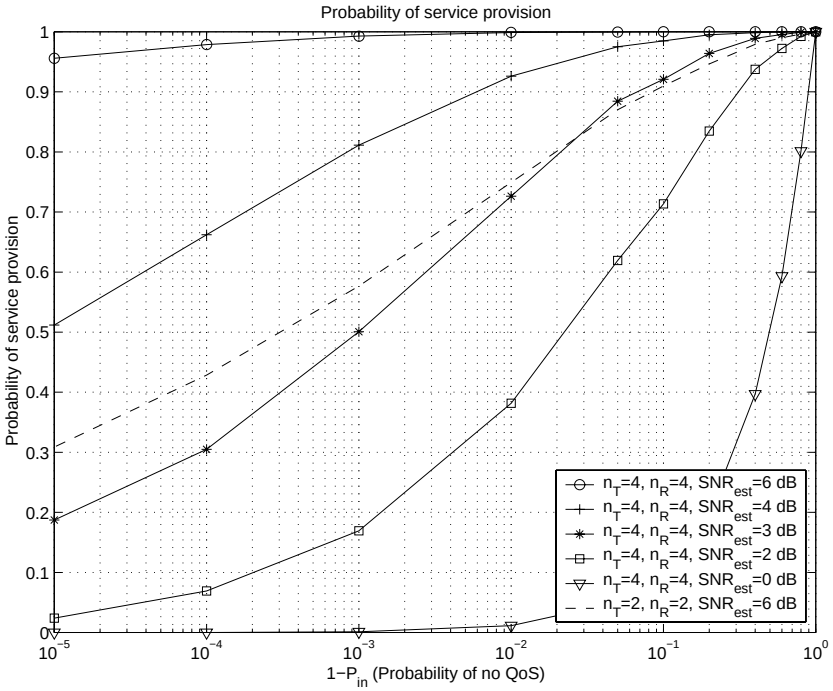


Figure 6.11: Probability of service provision vs. $1 - P_{\text{in}}$, where P_{in} is the probability of attaining the desired QoS assuming Gaussian noise and spherical uncertainty regions.

As explained in §6.6.2, the throughput can be maximized while guaranteeing a maximum BER for any possible error in the channel estimate by using the robust power allocation. Figure 6.13 shows the mean throughput that can be achieved in order to guarantee a maximum BER equal to 10^{-3} for any possible error in the uncertainty region, which is considered to be spherical with a radius equal to $\sqrt{\epsilon} = g\|\widehat{\mathbf{H}}\|_F$. The techniques that are compared are the robust approach and the non-robust classical beamforming solution. For both techniques, the mean throughput is shown as a function of the maximum available power at the transmitter. Besides, the plots regarding the application of fixed modulation formats corresponding to QPSK and 16-QAM are also given. From the figure it is concluded that, thanks to the use of the robust maximin design, very important savings in terms of transmit power can be obtained when compared to the non-robust solution, specially when the size of the uncertainty region is high, as expected. In the same figure, the improvement of the system can also be shown when the number of antennas increases.

6.8 Chapter Summary and Conclusions

In this chapter, a design strategy has been presented for a MIMO system in which the available CSI is imperfect. The design of the transmitter has been done according to a channel estimate

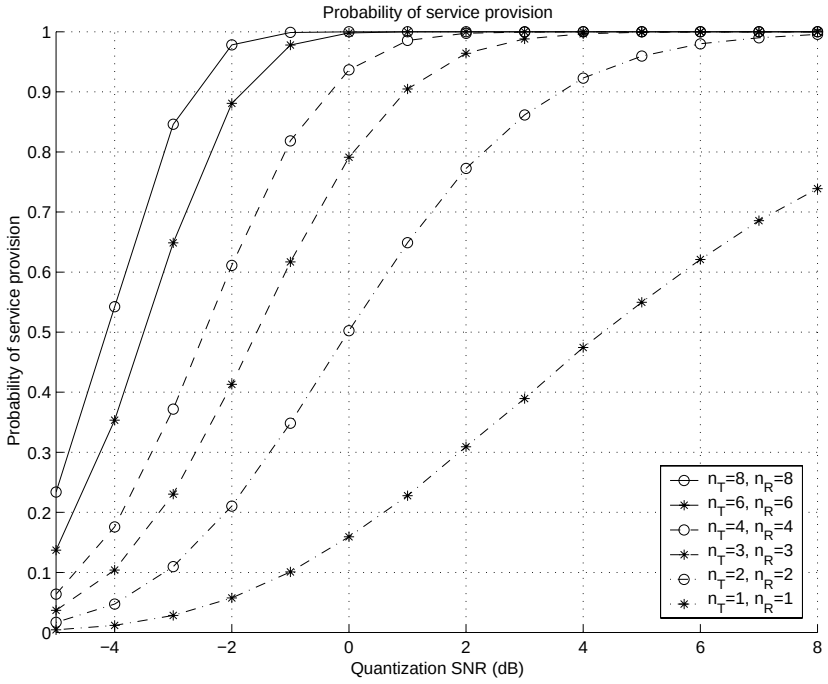


Figure 6.12: Probability of service provision vs. quantization SNR, assuming quantization errors and cubic uncertainty regions.

that may have errors from different origins, such as the Gaussian noise from the estimation process, or the errors from the quantization, among others. These errors have been taken into account explicitly in the design, obtaining a robust solution less sensitive to them. The robustness in the design has been obtained under the maximin philosophy. This strategy is characterized by attaining the best worst performance, in terms of SNR, for any possible error modeled by an uncertainty region. The shape and the size of the uncertainty region have to be chosen according to the source and the level of the imperfections in the CSI.

The transmitter architecture that has been proposed is based on the concatenation of an OSTBC block, a power allocation, and a set of beamformers, each one connected to one of the outputs of the OSTBC and corresponding to an eigenmode of the MIMO channel estimate. The robustness has been included according to an adequate power distribution of the total transmit power among the estimated eigenmodes. Thanks to this optimum robust power allocation, the necessary transmit power is minimized while guaranteeing a minimum instantaneous SNR for any possible estimation error in the uncertainty region.

The mathematical optimization problem corresponding to the maximin robust power allocation has been transformed into a simple convex optimization problem, where the values of the optimum Lagrange multipliers provide the normalized robust power distribution. Note that this

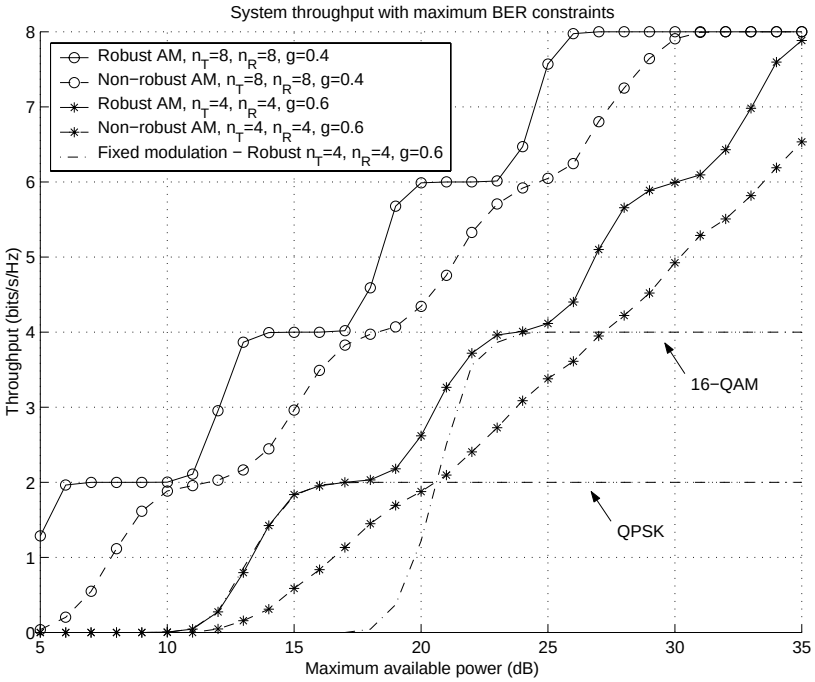


Figure 6.13: Mean throughput for the robust and the non-robust approaches combined with AM taking a maximum allowed $\mathrm{BER}_0 = 10^{-3}$.

convex problem can be solved efficiently with existing software packages. For many uncertainty regions, the convex problem has been shown to be quadratic, and even for the case of a spherical uncertainty region, a closed-form solution exists.

Finally, this robust solution has been compared with a pure OSTBC strategy and also with the classical non-robust beamforming design, corresponding to the use of only the maximum estimated channel eigenvector. From the simulation results, it has been observed that the gains in terms of saving in transmit power are quite important when compared to the non-robust and the pure OSTBC techniques, especially when the estimation and quantization SNR's are low.

Chapter 7

Conclusions and Future Work

This dissertation has considered the design of multi-antenna communication systems, where the most general case corresponds to the communication through MIMO channels with multiple antennas at both the transmitter and the receiver. More specifically, the impact of the CSI on the design has been studied. First, a perfect and complete CSI has been assumed, obtaining an optimum solution for the case of single-user and multi-user communications. The impact of an imperfect CSI has been studied in the case of a single-user communication, when the design is carried out assuming that the available channel estimate is perfect, despite not being true. Different robustness strategies have been proposed taking into account explicitly the errors in the CSI, obtaining designs less sensitive to these errors. Two examples of Bayesian robust designs have been deduced and, finally, a maximin robust solution for a MIMO single-user channel has been obtained.

7.1 Conclusions

A motivation of the Ph.D. dissertation has been presented in Chapter 1, jointly with an outline of the work and the research contributions in terms of publications.

Chapter 2 has presented an overview of the different design strategies in MIMO channels, providing a general description of the state of the art and focusing the attention specifically on the impact of the quality and the quantity of the CSI on the design and the final system performance. Additionally, a brief description of the most useful tools derived from convex optimization theory have been presented, since they have been used throughout the dissertation.

Chapter 3 has been devoted to the analysis and design of a single-user MIMO system. The assumed modulation at the physical layer is OFDM, i.e., a multicarrier approach has been taken. First, a brief review of the OFDM modulation has been done. Afterwards, the signal model corresponding to the transmission through a MIMO channel combined with the OFDM

modulation has been detailed. In this case, both the transmitter and the receiver architectures have been based on a beamforming per carrier approach, i.e., a different beamvector is applied at each subcarrier at the transmitter and the receiver. The design of the beamvectors has been carried out assuming a perfect CSI available at both sides of the communication system. This CSI encompasses both the channel response and the noise plus interferences correlation matrices at the receiver. The design of the beamvectors maximizing the SNIR leads to a solution composed of two stages. First, at each subcarrier, the beamvectors selecting the strongest eigenmode of the channel are calculated and, afterwards, a power distribution of the total available transmit power among the subcarriers is performed. Three different and simple power allocation strategies have been deduced based on different norms of the SNIR at all the subcarriers and using convex optimization theory. Finally, these power allocation strategies have been compared with other classical approaches, showing that the new solutions have a lower computational load with even better performances.

Chapter 4 has extended the results and the formulation in the previous chapter to the case of a multi-user scenario, in which all the terminals are allowed to have multiple antennas and to transmit simultaneously using the same frequency band, leading to several and parallel MIMO transmissions. As previously, the OFDM modulation and a beamforming per carrier approach have been considered, where the CSI has been assumed to be perfect at all the transmitters and the receivers. The main problem of the design is that a coupling effect appears, since the transmitter design for one user modifies the noise plus interferences correlation matrices for all the other links. The design problem has been formulated as the minimization of the total transmit power subject to QoS constraints in terms of a maximum BER for each link, and to optional individual maximum transmit power constraints. The problem has been shown to be non-convex due to the QoS constraints and the coupling effect mentioned previously. In order to find the optimum solution despite the non-convexity, the SA algorithm has been proposed. The SA, an heuristic and stochastic optimization tool, has been first briefly described and, afterwards, its application to the communication system design has been detailed. The performance of this algorithm has been compared to other classical approaches, such as AM and GS, showing that SA is able to find a solution fulfilling the constraints and requiring less transmit power.

Chapter 5 has been devoted to the analysis of the imperfect CSI. First, different sources of errors in the channel estimate have been identified and briefly analyzed. Afterwards, the impact of the errors in the CSI on the system performance has been studied. The system considered in this analysis has been that presented in Chapter 3, i.e., a MIMO-OFDM single-user communication. The design has been performed assuming that the CSI is perfect, despite not being true, and the performance degradation has been analyzed in terms of the mean value of an upper-bound on the relative SNIR reduction. Two different ways of introducing robustness in the design have been then summarized: the Bayesian and the maximin approaches. Both of

them take into account explicitly the errors in the channel estimate and, therefore, the obtained designs are less sensitive to these errors. The difference between both approaches consists in the way the error is modeled. Finally, two concrete examples of Bayesian designs have been given: a power allocation technique that minimizes the mean Chernoff upper-bound on the error probability in a SISO-OFDM transmission, and the design of a bank of FIR transmit filters in a MISO channel, where the objective is either the minimization of the mean MSE or the maximization of the mean SNR. In both cases, the performance corresponding to the robust design is much better than the one attained by the non-robust counterpart.

Finally, in Chapter 6, a robust maximin design of a single-user flat fading MIMO communication has been derived. The objective of the design has been the maximization of the SNR, where different error models have been taken according to different sources of imperfections or errors in the channel estimate. The transmitter architecture has been based on the concatenation of an OSTBC, a power allocation among the outputs of the OSTBC, and a set of beamformers corresponding to the estimated eigenmodes of the MIMO chanel. The problem formulation, corresponding to the design of the power allocation, leads to a two stage optimization problem that has shown to be transformed into a single stage convex optimization problem under some technical conditions regarding the convexity of the error region. This transformation permits to find the optimum solution either analytically or numerically. Afterwards, some extensions and applications of this robust design have been proposed, such as the extension to OFDM modulation and frequency selective channels, and its combination with AM techniques to maximize the throughput subject to maximum BER constraints and with minimum transmit power requirements.

7.2 Future Work

There exist several lines for future research that can be taken as an extension of the work carried out in this dissertation.

Concerning the design of multi-antenna systems assuming a perfect CSI, the following points and issues are still open:

- Evaluation of the different joint beamforming strategies when they are combined with space-time or space-frequency coding strategies.
- Introduction of new degrees of freedom in the design of the joint beamforming approach by considering rate adaptive modulations.
- Use of extensions and variants of the SA algorithm to increase the convergence speed while still guaranteeing that a global optimum is reached.

- Use of other heuristic stochastic approaches to solve the multi-user system design problem, such as genetic algorithms and taboo search.
- Development of suboptimal designs with a lower computational complexity than SA and evaluation of the degradation when compared to the optimal solution.

Finally, and as far as robust designs is concerned, the following research lines are proposed:

- Development of Bayesian designs optimizing the performance for an outage probability instead of the mean performance.
- Proposal of transmitter architectures different from the one used in Chapter 6 (OSTBC, power allocation, and beamforming), and evaluation of the robustness capabilities.
- Design of feedback strategies that improve the system performance while maintaining the same feedback rate.
- Design of convex uncertainty regions in vector quantization taking as the distortion criterion not the error in the channel estimate, but directly the cost function describing the system performance, such as the SNR, the BER, etc.
- Design of feedback strategies where the information to be fed back is the transmitter design instead of the channel estimate.
- Development of robust design strategies where the CSI is incomplete instead of imperfect, which means that, for example, only the gains of the channel are known but not their phases.
- Development of robust designs for MIMO multi-user systems, extending the results obtained in the single-user case.

Bibliography

[Aar89] E. Aarts, and J. Korst, *Simulated Annealing and Boltzmann Machines*, John Wiley & Sons, 1989.

[AH02] C. Anton-Haro, and M. A. Lagunas Hernández, "Array Processing Techniques for Wireless: a Cross-Layer Perspective", *Proc. International Forum on Future Mobile Telecommunications and China-EU Post Conference on Beyond 3G*, November 2002.

[Akh03] J. Akhtar, and D. Gesbert, "Partial Feedback Based Orthogonal Block Coding", *Proc. IEEE Vehicular Technology Conference Spring (VTC'03)*, April 2003.

[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, October 1998.

[Bar02] D. Bartolomé, A. I. Pérez Neira, and A. Pascual Iserte, "Blind and Semiblind Spatio-Temporal Diversity for OFDM Systems", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*, vol. 3, pp. 2769–2772, May 2002.

[Bar03a] D. Bartolomé, A. Pascual Iserte, and A. I. Pérez Neira, "Spatial Scheduling Algorithms for Wireless Systems", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, vol. 4, pp. 185–188, April 2003.

[Bar03b] D. Bartolomé, A. Pascual Iserte, A. I. Pérez Neira, and P. Rosson, "From a Theoretical Framework to a Feasible Hardware Implementation of Antenna Array Algorithms for WLAN", *Proc. IST Mobile & Wireless Communications Summit (IST'03)*, vol. 1, pp. 1–5, June 2003.

[Ben99] M. Bengtsson, and B. Ottersten, "Optimal Downlink Beamforming Using Semidefinite Optimization", *Proc. Allerton Conference on Communications, Control, and Computing (ALLERTON'99)*, pp. 987–996, September 1999.

[Ben01] M. Bengtsson, and B. Ottersten, "Optimal and Suboptimal Transmit Beamforming", L. C. Godara (ed.), *Handbook of Antennas in Wireless Communications*, CRC Press, August 2001.

[Ben02] M. Bengtsson, "A Pragmatic Approach to Multi-User Spatial Multiplexing", *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM'02)*, pp. 130–134, August 2002.

[Ber82] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Computer Science and Applied Mathematics, Academic Press, 1982.

[Ber99] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, 2nd ed., 1999.

[Bha02] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback Gain in Multiple Antenna Systems", *IEEE Trans. on Communications*, vol. 50, no. 5, pp. 785–798, May 2002.

[Big03] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman, "Robust Power Adjustment for Transmit Beamforming in Cellular Communications Systems", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, vol. 5, pp. 105–108, April 2003.

[Bin90] J. A. C. Bingham, "Multicarrier Modulation for Data Transmission: An Idea Whose Time Has Come", *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, May 1990.

[Blu01] Bluetooth SIG, *Specification of the Bluetooth System. v.1.1*, February 2001.

[Boc02] H. Boche, and M. Schubert, "A General Duality Theory for Uplink and Downlink Beamforming", *Proc. IEEE Vehicular Technology Conference Fall (VTC'02)*, vol. 1, pp. 87–91, September 2002.

[Boy00] S. Boyd, and L. Vandenberghe, *Introduction to Convex Optimization with Engineering Applications*, Course Notes (available at http://www.stanford.edu/class/ee364). Stanford University, 2000.

[Boy04] S. Boyd, and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[Car86] A. B. Carlson, *Communication Systems*, McGraw-Hill, 3rd ed., 1986.

[Cav72] J. Cavers, "Variable-Rate Transmission for Rayleigh Fading Channels", *IEEE Trans. on Communications*, vol. 20, no. 1, pp. 15–22, February 1972.

[Cha02] J-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint Transmitter Receiver Diversity for Efficient Space Division Multiaccess", *IEEE Trans. on Wireless Communications*, vol. 1, no. 1, pp. 16–27, January 2002.

[Cho02a] J. Choi, "A Semiblind Method for Transmit Antenna Arrays in CDMA Systems", *IEEE Trans. on Vehicular Technology*, vol. 51, no. 4, pp. 624–635, 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, January 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, and A. J. Goldsmith, "Degrees of Freedom in Adaptive Modulation: A Unified View", *IEEE Trans. on Communications*, vol. 49, no. 9, pp. 1561–1571, September 2001.

[Cov91] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.

[Del01] J-P. Delmas, "Asymptotic Normality of Sample Covariance Matrix for Mixed Spectra Time Series: Application to Sinusoidal Frequencies Estimation", *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1681–1687, May 2001.

[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.

[ETS98] ETSI, *ETSI EP BRAN 3ERI085B: Channel Models for HIPERLAN/2 in Different Indoor Scenarios*, March 1998.

[ETS00] ETSI, *ETSI TS 101 475 v1.1.1: Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Physical (PHY) layer*, April 2000.

[Fen00] X. Feng, and C. Leung, "Performance Sensitivity Comparison of Two Diversity Schemes", *Electronics Letters*, vol. 36, no. 9, pp. 838–839, April 2000.

[Fos96] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multi-Element Antennas", *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[Gan01] G. Ganesan, and P. Stoica, "Space-Time Block Codes: A Maximum SNR Approach", *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1650–1656, May 2001.

[Gan02a] G. Ganesan, *Designing Space-Time Codes Using Orthogonal Designs*, PhD Thesis, Uppsala University, 2002.

[Gan02b] G. Ganesan, and P. Stoica, "Differential Modulation Using Space-Time Block Codes", *IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 57–60, February 2002.

[Gar94] W. A. Gardner (ed.), *Cyclostationarity in Communications and Signal Processing*, IEEE Press, 1994.

[Ger92] A. Gersho, and R. M. Gray (eds.), *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, 1992.

[Gia89] G. B. Giannakis, and J. M. Mendel, "Identification of Nonminimum Phase Systems Using Higher Order Statistics", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 37, no. 3, pp. 360–377, March 1989.

[Glo89] F. Glover, "Taboo Search - Part I", *ORSA Journal on Computing*, vol. 1, no. 3, pp. 190–206, 1989.

[Gof99] J-L. Goffin, and J-P. Vial, "Convex Nondifferentiable Optimization: A Survey Focussed on the Analytic Center Cutting Plane Method", Tech. rep., Department of Management Studies, University of Geneva, Switzerland, February 1999.

[Gol88] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1988.

[Gol96] G. Golub, and C. Van Loan, *Matrix Computations*, Johns Hopkins, Baltimore, 1996.

[Gol97] A. J. Goldsmith, and S-G. Chua, "Variable-Rate Variable-Power MQAM for Fading Channels", *IEEE Trans. on Communications*, vol. 45, no. 10, pp. 1218–1230, October 1997.

[Has99] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite Quadratic Estimation and Control: A Unified Approach to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Theories*, SIAM, Philadelphia, PA, 1999.

[Has02] B. Hassibi, and B. M. Hochwald, "Cayley Differential Unitary Space-Time Codes", *IEEE Trans. on Information Theory*, vol. 48, no. 6, pp. 1485–1503, June 2002.

[Hen04] C. Hennebert, P. Rosson, D. Bartolomé, A. Pascual Iserte, and A. I. Pérez Neira, "Practical Implementation of Space-Diversity Receivers in OFDM Systems: Structure, Performance, and Complexity", *Proc. IST Mobile & Wireless Communications Summit (IST'04)*, June 2004.

[Her00] L. Herault, "Rescaled Simulated Annealing - Accelerating Convergence of Simulated Annealing by Rescaling the States Energies", *Journal of Heuristics*, vol. 6, no. 2, pp. 215–252, June 2000.

[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, March 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, September 2000.

[Hoc00c] B. M. Hochwald, and W Sweldens, "Differential Unitary Space-Time Modulation", *IEEE Trans. on Communications*, vol. 48, no. 12, pp. 2041–2052, December 2000.

[Hol02] H. Holma, and A. Toskala, *WCDMA for UMTS - Radio Access for Third Generation Mobile Communications*, John Wiley & Sons, 2nd ed., 2002.

[Hon92] M. L. Honig, P. Crespo, and K. Steiglitz, "Suppression of Near- and Far-End Crosstalk by Linear Pre- and Post-Filtering", *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 614–629, April 1992.

[Hug00] B. L. Hughes, "Differential Space-Time Modulation", *IEEE Trans. on Information Theory*, vol. 46, no. 7, pp. 2567–2578, November 2000.

[IEE99] IEEE, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY), IEEE Std. 802.11a*, December 1999.

[Jaf01] H. Jafharkani, "A Quasi-Orthogonal Space-Time Block Code", *IEEE Trans. on Communications*, vol. 49, no. 1, pp. 1–4, January 2001.

[Jai89] A. K. Jain, *Fundamentals of Digital Signal Processing*, Prentice Hall International, Englewood Cliffs, NJ, 1989.

[Jan98] W. M. Jang, R. Vojčić, and R. L. Pickholtz, "Joint Transmitter-Receiver Optimization in Synchronous Multiuser Communications over Multipath Channels", *IEEE Trans. on Communications*, vol. 46, no. 2, pp. 269–278, February 1998.

[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, March 2002.

[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, March 1985.

[Kay93] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall International, 1993.

[Kha86] A. Khachaturyan, "Statistical Mechanics Approach in Minimizing a Multivariable Function", *Journal of Mathematical Physics*, vol. 27, no. 7, pp. 1834–1838, July 1986.

[Laa87] P. J. M. van Laarhoven, and E. H. L. Aarts, *Simulated Annealing: Theory and Applications*, Kluwer Academic Publishers, 1987.

[Lar02] E. G. Larsson, G. Ganesan, P. Stoica, and W-H. Wong, "On the Performance of Orthogonal Space-Time Block Coding With Quantized Feedback", *IEEE Communications Letters*, vol. 6, no. 11, pp. 487–489, November 2002.

[LH00] M. A. Lagunas Hernández, J. Vidal, and A. I. Pérez Neira, "Joint Array Combining and MLSE for Single-User Receivers in Multipath Gaussian Multiuser Channels", *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2252–2259, November 2000.

[Li02] Y. G. Li, "Simplified Channel Estimation for OFDM Systems With Multiple Transmit Antennas", *IEEE Trans. on Wireless Communications*, vol. 1, no. 1, pp. 67–75, January 2002.

[Li03] J. Li, P. Stoica, and Z. Wang, "On Robust Capon Beamforming and Diagonal Loading", *IEEE Trans. on Signal Processing*, vol. 51, no. 7, pp. 1702–1715, July 2003.

[Liu03] J. Liu, A. Pascual Iserte, and M. A. Lagunas Hernández, "Blind Separation of OSTBC Signals Using ICA Neural Networks", *Proc. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'04)*, December 2003.

[Lo99] T. K. Y. Lo, "Maximum Ratio Transmission", *IEEE Trans. on Communications*, vol. 47, no. 10, pp. 1458–1461, October 1999.

[Lok00] T. M. Lok, and T. F. Wong, "Transmitter and Receiver Optimization in Multicarrier CDMA Systems", *IEEE Trans. on Communications*, vol. 48, no. 7, pp. 1197–1207, July 2000.

[Lor04] R. Lorenz, and S. P. Boyd, "Robust Minimum Variance Beamforming", *IEEE Trans. on Signal Processing*, to appear 2004.

[Lov04] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What Is the Value of Limited Feedback for MIMO Channels?", *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, October 2004.

[Loz02] 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, January 2002.

[Mag99] J. R. Magnus, and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons, 1999.

[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, January 1999.

[Mon80] R. A. Monzingo, and T. W. Miller, *Introduction to Adaptive Arrays*, John Wiley & Sons, New York, 1980.

[Mor04] A. Morell, A. Pascual Iserte, and A. I. Pérez Neira, "Fuzzy Inference Robust Beamforming", *submitted to EURASIP Signal Processing*, 2004.

[Mou92] M. Mouly, and M-B. Pautet, *The GSM System for Mobile Communications*, Telecom Pub, 2nd ed., 1992.

[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, November 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, October 1998.

[Nee98] R. van Nee, and R. Prasad, *OFDM for Mobile Multimedia Communications*, Artech House, Boston/London, 1st ed., 1998.

[Nes94] Y. Nesterov, and A. Nemirovski, "Interior-point Polynomial Methods in Convex Programming", *SIAM Studies in Applied Mathematics*, vol. 13, 1994.

[Ong03] E. N. Onggosanusi, A. M. Sayeed, and B. D. Van Veen, "Efficient Signaling Schemes for Wideband Space-Time Wireless Channels Using Channel State Information", *IEEE Trans. on Vehicular Technology*, vol. 52, no. 1, pp. 1–13, January 2003.

[Osb94] M. Osborne, and A. Rubinstein, *A Course in Game Theory*, MIT Press, Cambridge, 1994.

[Pal01] D. P. Palomar, M. A. Lagunas Hernández, A. Pascual Iserte, and A. I. Pérez Neira, "Practical Implementation of Jointly Designed Transmit-Receive Space-Time IIR Filters", *Proc. IEEE International Symposium on Signal Processing and its Applications (ISSPA'01)*, pp. 521–524, August 2001.

[Pal03a] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas Hernández, "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, September 2003.

[Pal03b] 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.

[Pal05] D. P. Palomar, A. Pascual Iserte, J. M. Cioffi, and M. A. Lagunas Hernández, "Convex Optimization Theory Applied to Joint Transmitter-Receiver Design in MIMO Channels", A. B. Gershman, N. D. Sidiropoulos (eds.), *Space-Time Processing for MIMO Communications (to appear)*, John Wiley & Sons, 2005.

[Pap91] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 3rd ed., 1991.

[Pau97] A. J. Paulraj, and C. B. Papadias, "Space-Time Processing for Wireless Communications", *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, November 1997.

[PI01a] A. Pascual Iserte, M. A. Lagunas Hernández, and A. I. Pérez Neira, "Space-Time Diversity Applied to Single-User Environments and MIMO Transmission Channels", *Proc. IEEE International Conference on Electronics, Circuits, and Systems (ICECS'01)*, vol. 3, pp. 1179–1182, September 2001.

[PI01b] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Iterative Algorithm for the Estimation of Distributed Sources Localization Parameters", *Proc. IEEE Workshop on Statistical Signal Processing (SSP'01)*, pp. 528–531, August 2001.

[PI01c] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Pre- and Post-Beamforming in MIMO Channels Applied to HIPERLAN/2 and OFDM", *Proc. IST Mobile & Wireless Communications Summit (IST'01)*, pp. 3–8, September 2001.

[PI02a] A. Pascual Iserte, N. Ahmed Awad, and A. I. Pérez Neira, "Array Antennas for Packet Transmission Networks", *Proc. ETSI Workshop on Broadband Wireless Ad-Hoc Networks and Services*, September 2002.

[PI02b] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Joint Beamforming Strategies in OFDM-MIMO Systems", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*, vol. 3, pp. 2845–2848, May 2002.

[PI02c] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Joint Transceiver Optimization in Wireless Multiuser MIMO-OFDM Channels Based on Simulated Annealing", *Proc. European Signal Processing Conference (EUSIPCO'02)*, vol. 2, pp. 421–424, September 2002.

[PI02d] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Simulated Annealing Techniques for Joint Transmitter-Receiver Design in a Multiple User Access MIMO-OFDM Channel", *Proc. IST Mobile & Wireless Communications Summit (IST'02)*, pp. 325–329, June 2002.

[PI02e] A. Pascual Iserte, A. I. Pérez Neira, D. P. Palomar, and M. A. Lagunas Hernández, "Power Allocation Techniques for Joint Beamforming in OFDM-MIMO Channels", *Proc. European Signal Processing Conference (EUSIPCO'02)*, vol. 1, pp. 383–386, September 2002.

[PI03a] A. Pascual Iserte, M. A. Lagunas Hernández, and A. I. Pérez Neira, "Diseño Cooperativo de Sistemas de Comunicaciones con Información Parcial del Canal", *Proc. Simposium Nacional de la Unión Científica Internacional de Radio (URSI'03)*, September 2003.

[PI03b] A. Pascual Iserte, M. A. Lagunas Hernández, and A. I. Pérez Neira, "Robust Power Allocation for Minimum BER in a SISO-OFDM System", *Proc. IEEE International Conference on Communications (ICC'03)*, vol. 2, pp. 1263–1267, May 2003.

[PI03c] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Exploiting Transmission Spatial Diversity in Frequency Selective Systems with Feedback Channel", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, vol. 4, pp. 85–88, April 2003.

[PI03d] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Performance Degradation of an OFDM-MIMO System with Imperfect Channel State Information at the Transmitter", *Proc. IST Mobile & Wireless Communications Summit (IST'03)*, vol. 2, pp. 396–400, June 2003.

[PI04a] A. Pascual Iserte, M. A. Lagunas Hernández, and A. I. Pérez Neira, "Robustness Criteria for Transmit Spatial Diversity Systems in Frequency Selective Channels", *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM'04)*, July 2004.

[PI04b] A. Pascual Iserte, D. P. Palomar, A. I. Pérez Neira, and M. A. Lagunas Hernández, "A Robust Maximin Approach for MIMO Communications with Partial Channel State Infor-

mation Based on Convex Optimization", *submitted to IEEE Trans. on Signal Processing*, September 2004.

[PI04c] A. Pascual Iserte, M. Payaró, A. I. Pérez Neira, and M. A. Lagunas Hernández, "Robust Adaptive Modulation for Throughput Maximization in MIMO Systems Combining OSTBC and Beamforming", *Proc. IST Mobile & Wireless Communications Summit (IST'04)*, June 2004.

[PI04d] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "A Maximin Approach for Robust MIMO Design: Combining OSTBC and Beamforming with Minimum Transmission Power Requirements", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04)*, vol. 2, pp. 1–4, May 2004.

[PI04e] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "An Approach to Optimum Joint Beamforming Design in a MIMO-OFDM Multiuser System", *accepted for publication in EURASIP Journal on Wireless Communications and Networking*, 2004.

[PI04f] A. Pascual Iserte, A. I. Pérez Neira, and M. A. Lagunas Hernández, "On Power Allocation Strategies for Maximum Signal to Noise and Interference Ratio in an OFDM-MIMO System", *IEEE Trans. on Wireless Communications*, vol. 3, no. 3, pp. 808–820, May 2004.

[Pro95] J. G. Proakis, *Digital Communications*, McGraw-Hill, 3$^{\underline{\text{rd}}}$ 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, March 1998.

[Rey02a] F. Rey, M. Lamarca, and G. Vázquez, "A Joint Transmitter-Receiver Design in MIMO Systems Robust to Channel Uncertainty for W-LAN Applications", *Proc. IST Mobile & Wireless Communications Summit (IST'02)*, June 2002.

[Rey02b] F. Rey, M. Lamarca, and G. Vázquez, "Optimal Power Allocation with Partial Channel Knowledge for MIMO Multicarrier Systems", *Proc. IEEE Vehicular Technology Conference Fall (VTC'02)*, vol. 4, pp. 2121–2125, September 2002.

[Rey03] F. Rey, M. Lamarca, and G. Vázquez, "Transmit Filter Optimization based on Partial CSI Knowledge for Wireless Applications", *Proc. IEEE International Conference on Communications (ICC'03)*, vol. 4, pp. 2567–2571, May 2003.

[RF98a] F. Rashid-Farrokhi, K. J. Ray Liu, and L. Tassiulas, "Transmit Beamforming and Power Control for Cellular Wireless Systems", *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, October 1998.

[RF98b] F. Rashid-Farrokhi, L. Tassiulas, and K. J. Ray Liu, "Joint Optimal Power Control and Beamforming in Wireless Networks Using Antenna Arrays", *IEEE Trans. on Communications*, vol. 46, no. 10, pp. 1313–1324, October 1998.

[Rhe04] W. Rhee, W. Yu, and J. M. Cioffi, "The Optimality of Beamforming in Uplink Multiuser Wireless Systems", *IEEE Trans. on Wireless Communications*, vol. 3, no. 1, pp. 86–96, January 2004.

[Roc71a] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 2nd ed., 1971.

[Roc71b] R. T. Rockafellar, "Saddle-Points and Convex Analysis", H. W. Kuhn, G. P. Szego (eds.), *Differential Games and Related Topics*, pp. 109–127, North-Holland Publ. Co., 1971.

[Rog80] G. S. Rogers, *Matrix Derivatives, Lecture Notes in Statistics, vol. 2*, Marcel Dekker, 1980.

[Sal85] J. Salz, "Digital Transmission over Cross-coupled Linear Channels", *AT&T Technical J.*, vol. 64, no. 6, pp. 1147–1159, July 1985.

[Sam01] H. Sampath, P. Stoica, and A. Paulraj, "Generalized Precoder and Decoder Design for MIMO Channels Using the Weighted MMSE Criterion", *IEEE Trans. on Communications*, vol. 49, no. 12, pp. 2198–2206, December 2001.

[Sca99a] 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, April 1999.

[Sca99b] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers Part I: Unification and Optimal Designs", *IEEE Trans. on Signal Processing*, vol. 47, no. 7, pp. 1988–2006, July 1999.

[Sca02] A. Scaglione, P. Stoica, S. Barbarossa, and G. B. Giannakis, "Optimal Designs for Space-Time Linear Precoders and Decoders", *IEEE Trans. on Signal Processing*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[Sch91] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, ans Time Series Analysis*, Addison-Wesley, 1991.

[Ser04] S. Serbetli, and A. Yener, "Transceiver Optimization for Multiuser MIMO Systems", *IEEE Trans. on Signal Processing*, vol. 52, no. 1, pp. 214–226, January 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 (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.

[Sha03] S. Shahbazpanahi, A. B. Gershman, Z-Q. Luo, and K. M. Wong, "Robust Adaptive Beamforming for General-Rank Signal Models", *IEEE Trans. on Signal Processing*, vol. 51, no. 9, pp. 2257–2269, September 2003.

[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*, October 2002.

[Ste99] R. Steele, and L. Hanzo (eds.), *Mobile Radio Communications*, John Wiley & Sons, 2nd ed., 1999.

[Ste01] M. Stege, M. Bronzel, and G. Fettweis, "On the Performance of Space-Time Block Codes", *Proc. IEEE Vehicular Technology Conference Spring (VTC'01)*, vol. 3, pp. 2282–2286, May 2001.

[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, December 2002.

[Sto03] P. Stoica, Z. Wang, and J. Li, "Robust Capon Beamforming", *IEEE Signal Processing Letters*, vol. 10, no. 6, pp. 172–175, June 2003.

[Szu87] H. Szu, and R. Hartley, "Fast Simulated Annealing", *Physics Letters A*, vol. 122, no. 3, pp. 157–162, June 1987.

[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, no. 2, pp. 744–765, March 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, "Combined Array Processing and Space-Time Coding", *IEEE Trans. on Information Theory*, vol. 45, no. 4, pp. 1121–1128, May 1999.

[Tar99c] 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, February 1999.

[Tar00] V. Tarokh, and H. Jafarkhani, "A Differential Detection Scheme for Transmit Diversity", *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 7, pp. 1169–1174, July 2000.

[Van84] D. Vanderbilt, and S. G. Louie, "A Monte Carlo Simulated Annealing Approach to Optimization over Continuous Variables", *Journal of Computational Physics*, vol. 56, no. 2, pp. 259–271, November 1984.

[Č84] V. Černy, "Minimization of Continuous Funcstions by Simulated Annealing", Tech. Rep. Preprint No. HU-TFT-84-51, Research Insitute for Theoretical Physics, University of Helsinki, 1984.

[Ver84] S. Verdú, and V. Poor, "On Minimax Robustness: A General Approach and Applications", *IEEE Trans. on Information Theory*, vol. 30, no. 2, pp. 328–340, March 1984.

[Vis99] E. Visotsky, and U. Madhow, "Optimum Beamforming Using Transmit Antenna Arrays", *Proc. IEEE Vehicular Technology Conference Spring (VTC'99)*, vol. 1, pp. 851–856, May 1999.

[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.

[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, February 2003.

[Ž00] S. Žaković, and C. Pantelides, "An Interior Point Method Algorithm for Computing Saddle Points of Constrained Continuous Minimax", *Annals of Operations Research*, vol. 99, pp. 59–77, December 2000.

[Wan00] Z. Wang, and G. B. Giannakis, "Wireless Multicarrier Communications", *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, May 2000.

[Web95] W. T. Webb, and R. Steele, "Variable Rate QAM for Mobile Radio", *IEEE Trans. on Communications*, vol. 43, no. 7, pp. 2223–2230, July 1995.

[Wil85] L. T. Wille, and J. Vennik, "Electrostatic Energy Minimization by Simulated Anneling", *Journal of Physics A: Mathematical and General*, vol. 18, no. 17, pp. L1113–1117, December 1985.

[Wit95] A. Wittneben, "Optimal predictive TX combining diversity in correlated fading for microcellular mobile radio applications", *Proc. IEEE Global Telecommunications Conference (GLOBECOM'95)*, pp. 13–17, November 1995.

[Won01] K. K. Wong, R. S. K. Cheng, K. B. Letaief, and R. D. Murch, "Adaptive Antennas at the Mobile and Base Stations in an OFDM/TDMA System", *IEEE Trans. on Communications*, vol. 49, no. 1, pp. 195–206, January 2001.

[Wor97] G. W. Wornell, and M. D. Trott, "Efficient Signal Processing Techniques for Exploiting Transmit Antenna Diversity on Fading Channels", *IEEE Trans. on Signal Processing*, vol. 45, no. 1, pp. 191–205, January 1997.

[Xia04] P. Xia, S. Zhou, and G. B. Giannakis, "Adaptive MIMO-OFDM Based on Partial Channel State Information", *IEEE Trans. on Signal Processing*, vol. 52, no. 1, pp. 202–213, January 2004.

[Yan94a] J. Yang, and S. Roy, "Joint Transmitter-Receiver Optimization for Multi-Input Multi-Output Systems with Decision Feedback", *IEEE Trans. on Information Theory*, vol. 40, no. 5, pp. 1334–137, September 1994.

[Yan94b] J. Yang, and S. Roy, "On Joint Transmitter and Receiver Optimization for Multiple-Input-Multiple-Output (MIMO) Transmission Systems", *IEEE Trans. on Communications*, vol. 42, no. 12, pp. 3221–3231, December 1994.

[Zho96] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice Hall International, Upper Saddle River, NJ, 1996.

[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, October 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.