Universitat Politècnica de Catalunya

MSc Thesis Dissertation

Tomlinson-Harashima Precoding for the SISO, MIMO and MISO Broadcast Channels

Author
Josep Font-Segura

Advisor
Gregori Vázquez-Grau

Barcelona, June 2008

Abstract

To achieve widespread market acceptance, future telecommunication systems must cope with the demand of high data rates in wireless transmission, which has increased significantly over the last years. Firstly, due to current bandwidth limitations, this has led to a strong interest in so-called multiple-input multiple-output (MIMO) systems which, already employed in third generation cellular systems and imminently implanted in wireless networks such as IEEE 802.11n and IEEE 802.16e Wimax standards, can significantly improve channel capacity and diversity with high spectral efficiency. Secondly, orthogonal-frequency division multiplexing (OFDM) provides a robust, scalable and low-cost solution to the distortion caused by multiple paths present in many outdoors and indoors wireless scenarios, and it has been well adopted in nowadays systems like the European digital terrestrial television broadcasting and the wireless networks IEEE 802.11a and ETSI Hiperlan II standards. And thirdly, most of the current and next-generation wireless communication systems show a highly asymmetric structure in terms of complexity —a high-complexity base station serves a number of low-complexity mobile terminals—, whose heterogeneity is due to the fact that the design of the mobile terminals, contrary to the base station, is tightly constrained in terms of cost and power consumption. The intersection of the three former realities in the third millennium information society defines an undeniable milieu where techniques based on Tomlinson-Harashima precoding (THP) might become extremely competitive. THP's leitmotiv is to rearrange the information previous to transmission, namely precoding, in order to diminish or cancel out the interference caused by the channel, together with the disposition of a non-linear modulo operation which avoids the transmit power enhancement typical in equalization systems.

The lack of depth in knowledge when designing systems based on THP married to MIMO and OFDM makes necessary the present thesis. Specifically, THP is proposed and acutely analyzed in three different scenarios in a very constructive way. The first part introduces the basic but required background to follow parts II, III and IV, which design the THP for the single-user single-input single-output (SISO), the single-user MIMO and multi-user multiple-input single-output (MISO) broadcast (BC) frequency-selective channels; the two latter with the use of OFDM. These parts, which form the body of the thesis and have very well defined objectives, are not independent neither disjunctive among them, so that the knowledge obtained from one is without any doubt applicable to the sequent. Finally, the last part provides a brief conclusion and proposes future work that might be investigated in this line. This methodology has derived to a very effective and profitable working plan, whose result is on your hands.

Resum

Per tal d'aconseguir una àmplia acceptació al mercat, els futurs sistemes de telecomunicació han de ser capaços de cobrir la demanda d'altes velocitats de transmissió sense fils, la qual ha experimentat un increment significatiu en els darrers anys. En primer lloc, a causa de les actuals limitacions d'espectre disponible, s'ha generat un fort interès en els anomenats sistemes amb múltiples entrades i múltiples sortides (MIMO) que, ja utilitzats en la tercera generació de sistemes de telefonia mòbil i amb imminent implantació en xarxes sense fils com els estàndards IEEE 802.11n i IEEE 802.16e Wimax, poden millorar notablement la capacitat del canal, així com la diversitat, amb un ús altament eficient de l'espectre. En segon lloc, la multiplexació per divisió en freqüències ortogonals (OFDM) ofereix una solució robusta, escalable i de baix cost al problema de la distorsió causada per la propagació multi camí present en molts escenaris sense fils, tant exteriors com interiors, i ha estat adoptada en sistemes actuals com l'europea televisió digital terrestre o els estàndards de xarxa sense fils IEEE 802.11a i ETSI Hiperlan II. En tercer lloc, la majoria dels sistemes de comunicació, actuals i futurs, mostren una estructura molt asimètrica en termes de complexitat —una estació base d'alta complexitat dóna servei a varis terminals mòbils de baixa complexitat—, l'heterogeneïtat de la qual es deu al fet que els dissenys dels terminals mòbils, contràriament a l'estació base, estan estretament lligats a limitacions de cost i consum energètic. La intersecció d'aquestes tres últimes realitats en l'actual societat de la informació defineix una innegable situació on tècniques basades en la precodificació de Tomlinson-Harashima (THP) poden esdevenir realment competitives. THP consisteix en reorganitzar la informació prèvia a la transmissió, procés anomenat precodificació, per tal de disminuir o cancel·lar les interferències causades pel canal, juntament amb la disposició d'una operació modular no lineal que evita els sobtats increments de potència transmesa inherents en molts sistemes d'equalització.

La falta de profunditat de coneixement al dissenyar sistemes basats en THP lligat a MIMO i OFDM fa necessària aquesta tesi. Específicament, THP és proposat i minuciosament analitzat en tres escenaris diferents en un procediment molt constructiu. La primera part introdueix els fonaments bàsics però necessaris per desenvolupar les següents parts —II, III i IV—, que dissenyen THP per als canals selectius en freqüència d'un sol usuari amb una entrada i una sortida (SISO), un sol usuari amb MIMO i múltiples usuaris de múltiples entrades i una sortida (MISO), també anomenat canal de difusió (BC); els dos últims casos amb l'ús addicional d'OFDM. Aquestes parts, que formen el cos de la tesi i tenen uns objectius molt ben definits, no són independents ni disjuntives entre elles, de tal manera que les conjectures i deduccions obtingudes en una són intuïtivament aplicables a la següent. Finalment, unes breus conclusions, així com propostes de línies d'investigació, són tractades en l'última part. Aquesta metodologia ha derivat a un pla de treball efectiu i fructífer, el resultat del qual teniu a les vostres mans.

Contents

List of Figures

List of Tables

Acronyms

The set of acronyms described in the following lines is widely used along the thesis. Though most of them are expanded in their first appearance, the reader may use this list as reference.

AWGN Additive white Gaussian noise.

BC Broadcast channel.

CDM Code domain multiplexing.

CO Central office.

DFE Decision feedback equalization.

DPC Dirty paper coding.

DSL Digital subscriber line.

EDGE Enhanced data rates for GSM evolution.

ETSI European Telecommunications Standards Institute.

EVD Eigenvalue decomposition.

FIR Finite length response.

FDM Frequency domain multiplexing.

FFT Fast Fourier transform.

GSM Global system for mobile communications.

ICI Inter-channel interference.

IEEE Institute of Electrical and Electronics Engineers.

ISI Inter-symbol interference.

KKT Karush-Kuhn-Tucker.

LE Linear equalization.

MAC Multiple-access channel.

MAI Multiple-access interference.

MIMO Multiple-input multiple-output.

MISO Multiple-input single-output.

MMSE Minimum mean-square error.

MSE Mean-square error.

MUI Multi-user interference.

NMSE Normalized mean-square error.

OFDM Orthogonal frequency division multiplexing.

OFDMA Orthogonal frequency division multiple access.

ONU Optical network unit.

PAM Pulse amplitude modulation.

PAR Peak-to-average power ratio.

PDF Probability density function.

PER Package error rate.

PLC Powerline communication.

QoS Quality of service.

SDM Spatial domain multiplexing.

SER Symbol error rate.

SIMO Single-input multiple-output.

SINR Signal-to-interference plus noise ratio.

SISO Single-input single-output.

SNR Signal-to-noise ratio.

SVD Singular value decomposition.

TDM Time domain multiplexing.

THP Tomlinson-Harashima precoding.

UMTS Universal mobile telecommunications system.

WLAN Wireless local area network.

ZF Zero-forcing.

Other contextual abbreviations that are extensively employed are:

i.e. "Id Est", from the Latin, meaning that is.

e.g. "Exempli Gratia", from the Latin, meaning for example.

cf. "Confer", from the Latin, meaning refer to, compare with.

i.i.d. Independent and identically distributed.

Notation

In the sequel, matrices are indicated by uppercase boldface letters, vectors are indicated by lowercase boldface letters, and scalars are indicated by italics letters. Other specific notation has been introduced as follows:

$=$	Equal to.
$\neq$	Not equal to.
$\triangleq$	Defined as.
$\approx$	Approximately equal to.
$\mapsto$	Maps to.
$\Leftrightarrow$	If and only if.
$\hat{x}$	Estimate of x.
$x < y$	x is smaller than y.
$x > y$	x is greater than y.
$x \ll y$	x is much smaller than y.
$x \gg y$	x is much greater than y.
$x \leq y$	x is smaller or equal to y.
$x \geq y$	x is greater or equal to y.
$\min(x, y)$	The smallest of x and y.
$\max(x, y)$	The largest of x and y.
$\log(x)$	Natural logarithm of x, which evaluates the information in nats.
$\log_2(x)$	Logarithm in base 2 of x, which evaluates the information in bits.
$\mathbb{R}^{N \times M}$	Set of $N \times M$ matrices with real valued entries.
$\mathbb{C}^{N \times M}$	Set of $N \times M$ matrices with complex valued entries.
$\mathrm{Re}(z)$	Real part of $z \in \mathbb{C}$.
$\mathrm{Im}(z)$	Imaginary part of $z \in \mathbb{C}$.

$\boldsymbol{X}^*$, $\boldsymbol{X}^T$, $\boldsymbol{X}^H$	Complex conjugate, transpose and transpose conjugate or Hermitian of matrix $\boldsymbol{X}$, respectively.
$\boldsymbol{X}_{ij}$	The (i,j)-th element of matrix $\boldsymbol{X}$.
$\boldsymbol{I}_N$	The $N \times N$ identity matrix.
$\boldsymbol{0}_{N \times M}$	The $N \times M$ all zeros matrix.
$\lvert x \rvert$	Absolute value of x.
$\lvert \boldsymbol{x} \rvert^2$	Euclidean norm, i.e. $\boldsymbol{x}^H \boldsymbol{x}$.
$\lvert \boldsymbol{X} \rvert = \det(\boldsymbol{X})$	Determinant of matrix $\boldsymbol{X}$.
$\lvert\lvert \boldsymbol{X} \rvert\rvert^2$	Squared Frobenius norm, i.e. the $\text{tr}(\boldsymbol{X}\boldsymbol{X}^H)$.
$\text{tr}(\boldsymbol{X})$	The trace operation over the squared matrix $\boldsymbol{X}$.
$E[\bullet]$	Expected value operator.
$\text{diag}(\boldsymbol{x})$	Returns a diagonal matrix whose main diagonal is the vector $\boldsymbol{x}$.
$\text{diag}(x_1 \dots x_N)$	Returns a diagonal matrix whose main diagonal is formed by the set of scalars $\{x_1 \dots x_N\}$.
$\text{diag}(\boldsymbol{X})$	Returns a diagonal matrix whose main diagonal is the main diagonal of matrix $\boldsymbol{X}$.
$\text{diag}(\boldsymbol{X}_1 \dots \boldsymbol{X}_N)$	Returns a block diagonal matrix whose diagonal is formed by the set of matrices $\{\boldsymbol{X}_1 \dots \boldsymbol{X}_N\}$.
$\boldsymbol{X} \odot \boldsymbol{Y}$	Hadamard product, that is the element-wise product between matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, $(\boldsymbol{X} \odot \boldsymbol{Y})_{ij} = \boldsymbol{X}_{ij} \cdot \boldsymbol{Y}_{ij}$.
$(x)^+$	Sets the value to zero if $x \leq 0$. In other words, $(x)^+ \triangleq \max(0, x)$.
$\nabla_{\boldsymbol{X}} f(\boldsymbol{X})$	Gradient of function f with respect to matrix $\boldsymbol{X}$.

Part I

Introduction

Chapter 1

Motivation to THP

Tomlinson-Harashima precoding (THP) has recently received much attention as a very competing signal processing modus operandi for wireless communications, though it was originally developed by Tomlinson [Tom71] and Harashima [HM72] during the early seventies as an attractive technique to reduce the inter-symbol interference (ISI) present in classical communication channels. The initial axiom of Tomlinson and Harashima was to prepare the user information at signal level by means of a feedback loop and a non-linear modulo operation in order to compensate, prior to transmission, the effect of the communication channel. This has led to a very well-known structure known as THP. However, during the last years, THP has been regarded as a practical and low-complex implementation of the so-called dirty paper coding (DPC), motivated by Costa [Cos83] in the early eighties, which promises high data rates over any channel whose interferences are known to the transmitter, if appropriate precoding is used. Hence, it has been at the present age, when the frequency spectrum has become a very limited and expensive resource, that efficient techniques such as THP have an essential role in current and next-generation multi-user systems.

The efficient utilization of a communication channel is translated to simultaneously dwarf any type of interference and independent random noise —we understand as interference any noise or undesired signal that can be known or estimated. Equalization methods are in charge of mitigating these effects and many techniques have been described in the literature. Though linear equalizers have relatively good performance (e.g. [JUN05]), they become limited in some scenarios and other structures, such as decision feedback equalization (DFE), must be considered. DFEs are motivated by their ability to reduce the noise enhancement in linear equalization (LE) (e.g. [ADC95] or [RG07]), but the concept of error propagation turns out to be of a great concern in practical applications, specially if constellation-expanding or convolution codes are used. Compared to these techniques, THP becomes a potential competitor because it places the feedback filter to the transmitter, where decision errors are impossible. In addition, the significant transmit power increase that could result from placing this filter previous to transmission is overcome by a non-linear modulo operation, which is employed to bound the value of the transmitted symbol.

Next generation telecommunication services require an extensive use of the resources due to evident bandwidth limitation in the current spectrum. The key resides in determining techniques that allow the system to place more and more information to the channel using a limited portion of frequency. This mathematically converges to a high capacity system which, at the end, is able to attain a larger number of served users with better quality of service (QoS) —in terms of both high speed data rates and increased reliability. On the one hand, multiple-input multiple-output (MIMO) systems

have been regarded as one of the most promising research areas of wireless communications due to the increase of higher degrees of diversity, fostering the use of these systems in order to enhance the capacity and the spectral efficiency with respect to classical single-input single-output (SISO) systems. The order of capacity attained is enabled by the fact that, in a rich environment, the signal from each individual input appears highly uncorrelated at each of the channel outputs and, as a result, the receiver obtains different signatures from the same signal. Goldsmith [GJJV03] has shown very presumable limits in terms of channel capacity for several communications problems easily modeled by a MIMO system. On the other hand, in view of the increased complexity with the order of the MIMO channel, an orthogonal frequency division multiplexing (OFDM) is usually considered for the physical layer due to its numerous advantages. Firstly, it can easily adapt to severe temporal channel conditions and it is robust in front of ISI and fading caused by the multi-path propagation. Secondly, it has high spectral efficiency and an efficient implementation using the fast Fourier transform (FFT). Surprisingly, Cioffi [RC98] has shown that using multi-carrier approaches on frequency-selective MIMO channels is a capacity-lossless structure. Therefore, the combination of MIMO and OFDM becomes, definitely, a cutting technology. It is foreseen that some of current and future telecommunication standards implement these techniques. In fact, MIMO has been already employed in third generation cellular systems and wireless networks like IEEE 802.11n and IEEE 802.16e Wimax standards. Also, OFDM has been well adopted in nowadays systems such as the European digital terrestrial television broadcasting or the WLAN IEEE 802.11a and ETSI Hiperlan II standards.

Most of multi-user systems show an asymmetric structure in condition of complexity, motivated by [GS99]: a highly-complex transmitter or base station serves a set of low-complex receivers or mobile terminals. This heterogeneous structure, consequence of the tight design of the mobile terminals in terms of cost and power consumption —many devices run on batteries—, is often modeled as the so-called broadcast channel (BC). Introduced in 1972 by Thomas M. Cover [Cov72], this channel models many situations of past and recent wireless and wire line telecommunication systems such as broadcasting terrestrial digital television information from a transmitter to multiple receivers in a given local area, the transmission of user information through a bundle of digital subscriber lines (DSLs), the transmission of voice and data on the cellular wireless network, the data stream downlink channel of a WLAN or the broadcast information sent by the home automation top box to several home machines through powerline communication (PLC). What is more, recent researched systems, such as the use of relays in cellular networks —a relay is a low-cost and low-transmit power element that receives and forwards data from the base station to the users via wireless channels, with the aim of boosting coverage and link capacity in regions with significant shadowing, cf. [CTHC08]— follow the broadcast formulation. In any case, the first evident question that arises in multi-user BC communication is how to separate the users, a concept that refers to multi-user multiplexing. While traditional systems make use of time and frequency to overcome multiplexing —for instance the GSM cellular network—; modern systems employ more sophisticated and spectrally efficient techniques such as spread spectrum techniques —like code domain multiplexing (CDM) in UMTS or EDGE, or frequency hopping in Bluetooth— and orthogonal multi-tone techniques —such as orthogonal frequency division multiple access (OFDMA) in the IEEE 802.16e Wimax standard. However, a promising approach is to employ the space, thanks to the MIMO paradigm, to spatially multiplex the users: spatial domain multiplexing (SDM). The main impairment of these systems is represented by the multi-user interference (MUI) arising from the simultaneous transmission of parallel data streams over the same frequency —and time and code— band. Luckily, there is a candidate with strong points to efficiently overcome MUI. Guess which? THP.

Chapter 2

Objectives and Contributions

The main objective of this thesis is to design, study and provide high detailed formal and critical background on techniques based on THP for the multi-user BC in view of the fact that it is a potential candidate for next generation of telecommunication systems involving wireless transmission and multi-users with asymmetric complexity. Albeit, this objective has been properly split into three parts in a very well-defined working plan. This methodology not only provides a productive and didactic manner to cope with the study, but also has served to solve many problems and acquire indispensable knowledge in a cultivated and incremental philosophy. The three objectives correspond to parts II, III and IV in this thesis:

SISO The first defined objective is to design the THP system willing to reduce the ISI in a single-user SISO frequency-selective channel, the original scenario of application first figured by Tomlinson and Harashima. The main premise of this objective has been to study the leitmotiv of THP —namely the feedback filter, the modulo operation, the statistics of the transmitted signal and the linearization of the model— in a very detailed level with the use of vectorial notation.

On the one hand, the understanding of the modulo operation requires its study from an algebraic point of view, because this non-linear operation directly affects the statistics of the transmitted signal, as well as the recoverability of the information at the receiver. In this sense, this theory allows the model to be linearized and, therefore, to be treated mathematically with efficiency. Also, the similarities between this operation and the process of quantization in many systems makes possible to analyze it while making use of the quantization theory. This has led to the introduction of a dithering signal to, first, control the statistics of the transmitted signal and, second, to reduce the peak-to-average power ratio (PAR) at the transmitter.

On the other hand, the SISO choice is justified with the use of vectorial notation in the process of optimization. The initial planing is to optimize the THP system with a power transmit constraint, under minimum mean-square error (MMSE) and zero-forcing (ZF) criteria, as well as the use of two matrix factorization techniques that have been comprehensively studied in the literature: Cholesky factorization and QR decomposition. By making so, many concepts related to matrices and derivatives, are in a such formal level that turn out to be a powerful tool to be applied in the next parts.

MIMO-OFDM The second aim is to design the THP system to reduce the inter-channel interference (ICI) in a single-user MIMO-OFDM frequency-selective channel. This objective is well justified with the wide use of MIMO and OFDM in nowadays and future telecommunication systems,

due to their numerous advantages.

The introduction of MIMO to the problem mathematically generates a channel with more capacity, which is translated to spatial diversity. However, a new type of interference, the ICI, is present in the resulting communication system. The introduction of OFDM is defended by the fact that, while it adapts to severe temporal channel conditions, it is of a reduced complexity and makes the problem scalable. The combination of MIMO and OFDM gives a high degree of flexibility to the system in placing the information in space and frequency. As a result, one of the main implications of this part is to overcome the problem of power allocation, an essential problem in modern communication systems. For the MIMO channel, the majorization and the Schur-convexity theories become generalized mathematical tools because provide true solutions to many power allocation problems.

The optimization of THP is done, as well, with a power transmit constraint, under MMSE and ZF criteria, in addition to Cholesky factorization an QR decomposition techniques.

MISO-OFDM BC Lastly, the main scope of the present work is to focus on how THP performs in a multi-user system as a precoding technique to reduce the MUI, the particular type of interference that appears in systems where multiple users coexist and share the available resources to communicate. This is the case of SDM, where the capacity offered by MIMO channels is properly adjusted to place multiple users who simultaneously transmit.

The objective is, hence, to design the THP system to spatially multiplex users in a MIMO frequency-selective channel together with the use of OFDM. In addition, this part considers the case of asymmetric complexity, in which the receiver is kept as simple as possible to obtain a low-complexity and low-cost solution. In other words, the receiver simply performs an automatic gain instead of linear filter processing, leading to the so-called multiple-input single-output (MISO) BC.

Again, the power allocation plays a fundamental part in BCs given the flexibility achieved thanks to MIMO and OFDM. However, in multi-user systems it is expected to require more sophisticated algorithms and concepts due to the fact that the system must manage the resources to fairly serve multiple users. This arises ethical questions related to the QoS —notice that users might be using different services which may need different requirements— and the very important concept of user ordering in precoding. All these questions must be translated to practical communication considerations, e.g., the power allocation. Under this scope, the power allocation algorithm must be designed from an information theory point of view and puts on the table theories related to user rate capacities and the multiple-access channel (MAC)-BC duality tool. It is well-known that the BC power allocation problem is non-convex, which means that it must be solved exhaustively and therefore without practicability. That is why in the last years many authors make use of the MAC-BC duality, as the MAC problem is much easer to solve because it allows convex formulation and, thus, efficient known solutions.

Finally, it will be studied the optimization of THP under MMSE and ZF criteria, as well as the possibility of Cholesky factorization and QR decomposition, all of them under a power transmit constraint.

The presented methodology of study turns out to be very capable if used in a constructive way. In other words, the signal level concepts analyzed in the part II, such as the dithering and the modulo operation, will apply to the following cases and many ideas will be reused, light-weighting the com-

plexity and focusing on other topics in the next parts. The same is expected to happen between part III and part IV, as both deal with power allocation.

Finally, the following points provide some of the contributions done in this thesis:

1. Analysis of the modulo operation under an algebraic point of view (subsection 5.2.1).
2. Study of temporal SISO THP in vectorial notation —i.e. the basic premise for which many SISO concepts will apply to the MIMO and BC parts (chapter 6).
3. Use of a dither signal to control the statistics of the transmitted signal (subsection 5.2.3).
4. Providing consistent mathematical backgrounds to justify the use of Cholesky factorization (subsection 6.3.2).
5. Use of Cholesky factorization and QR decomposition for the temporal SISO channel (sections 6.3 and 6.6).
6. Checking the performance of THP in nowadays communication standards by making use of temporal diversity (section 7.3).
7. Study of spatial MIMO THP in conjunction with OFDM (part III).
8. Providing consistent mathematical backgrounds for the optimization of the MIMO-OFDM power allocation (section 10.3).
9. Optimization of the power allocation for the ZF criterion (section 11.2).
10. Checking the performance of THP versus different parameters of the MIMO-OFDM channel (sections 12.3 and 12.4).
11. Discussion of a priori user ordering in multi-user systems (section 15.1).
12. Application of information theory known results to draw down the multi-user THP filters (section 15.2).
13. Discussion of applicability of QR decomposition in uncooperative scenarios (section 16.3).
14. Study of the iterative power allocation algorithm in MISO-OFDM channels (section 17.1).
15. Checking the performance of THP versus different parameters of the MISO-OFDM channel (sections 17.3, 17.4 and 17.5).

Chapter 3

Foundations

The present chapter summarizes some of the basis or groundwork concepts required to read (and write) this thesis. It involves many information theory and communication concepts that, throughout the dissertation, are referred to. This means, not only that are required to understand many ideas, but also that their comprehension and profound apprehension have been indispensable to the author. More specifically, the following sections introduce concepts related to DPC, CSI and channel capacity, as well as a brief review on equalization and precoding techniques.

3.1 Dirty Paper Coding

THP has gained a wide acceptance among the telecommunication society because it is considered a low-complex and efficient implementation of the well-known DPC, motivated by Costa [Cos83] in 1983. It has been demonstrated that many important communication problems —among them the majority of broadcast channels, the precoding for ISI channels and digital watermarking— can be modeled by the channel presented by Costa, usually referred as the dirty paper channel. This channel model considers a communication scheme where part of the noise is known as side information by the system and this part of the noise is usually known as interference. Costa's work is of such a theoretical level that no practical implementations are proposed.

The essence of DPC is not to directly transmit the user information —usually a small number of possible symbols—, but a larger set of codes instead. This codebook must be designed to represent the user information in such a way that there is no loss with respect to the additive white Gaussian noise (AWGN) channel —i.e. as if the interference was zero. Costa shows the existence of such codes and, recently, Shamai [ESZ05] outlines that whatever is the interference —this includes ISI, ICI and MUI, regardless its statistics—, if known to the transmitter at high signal-to-noise ratio (SNR) and appropriate coding is used, there is communication without performance loss compared to the AWGN channel.

Hence, regarding the structure of THP —a feedback loop plus modulo operation—, one realizes that it is performing iterative precoding of the information, at the signal level. In other words, each symbol is precoded taking into consideration the interference it will receive when being transmitted through the channel and, in addition, limited in amplitude by a modulo operation. This modulo operation —analogous to modulo arithmetic with real numbers— defines a coset of signals which plays the paper of codebook. Mathematically speaking, if the communication channel relates its output, Y, to its input, X by $Y = X + S + N$, where S is the interference known to the transmitter

and N is the Gaussian random noise, the THP would read

$$X = \lfloor U - \alpha S \rfloor_{\mathcal{C}},$$

with $\mathcal{C}$ the coset of signals, U the user signal and α the Costa parameter. Though its implementation is discussed in the body of the thesis —in fact, the feedback filter gives α and the modulo operation $\mathcal{C}$—, this simple equation shows the kernel of THP. Notice that in cases where the interference depends on the transmitted signal —which is the case of ISI, ICI and MUI—, the precoding must be done in an iterative fashion. However, there is a bare consideration to be made: DPC —and thus THP— works if and only if the knowledge of the interference is correct. Since the interference is given by the transmitted signal —which is without doubt known— and the channel response at the moment of transmission, the full knowledge is translated to the concept of CSI. In other words, the success of THP is extremely linked to the success of channel estimation.

3.2 Channel State Information

CSI refers to the degree of knowledge about the channel. Most signal processing techniques involving channel inversion rely on the capacity of the system in providing accurate measures of the value of the channel. This leads to the channel estimation theory, which is not treated in this thesis. Usually, the estimation of time-varying frequency-selective channels —this means that the frequency response must be estimated with some frequency step periodically at each certain period of time— is done by sending known information. Usually, this information is a pseudo-random sequence and it might be send in both links or one direction link and make use of channel reciprocity. The key in channel estimation is the capability of the estimator in giving accurate values as well its capacity in adapting to channel changes. Another consideration is which resources the system employs for the estimation, which might be reserved —like the case of OFDM pilots which transmit at a given number of carriers— or shared with data using, for example, spread spectrum techniques.

In the case of fixed or low mobility wireless networks —e.g. television, personal and local networks— achieving CSI is not a critical point. However, designing techniques that assume the knowledge of the channel realization (i.e. complete CSI), suffer from significant performance loss when the estimator cannot adapt to severe temporal channel conditions, which might be the case of services to users with high mobility. As a result, complete CSI is suboptimal (shown in [Lia05]) in some wireless scenarios with high channel variation and robust techniques must be designed. In this line, the conditional mean estimator (cf. [LTV06] for a good reference) is used to obtain the optimal filters when, not the realization of the channel is known, but the conditioned channel probability density function (PDF) (i.e. partial CSI). Recent works on THP that consider partial CSI are [DBU07] and [BF08], while [ZL07], [FTK07b] and [FTK07a] analyze the effect of channel mismatch in techniques that assume complete CSI.

3.3 Capacity: Rates and Regions

The channel capacity is the measure of the amount of information that can be reliably transmitted over a communication channel. By the noisy-channel coding theorem, however contaminated with noise and interference a communication channel may be, it is possible to communicate information nearly error-free up to a given maximum rate through the channel. This surprising result, sometimes

called the fundamental theorem of information theory was first presented by Shannon in 1948 [Sha48]. Information theory, developed Shannon during World War II, besides defining the notion of channel capacity, provides a mathematical model by which it can be computed. The key result states that the capacity of the channel is given by the maximum of the mutual information between the input and output of the channel, where the maximization is with respect to the input distribution. The channel capacity has been studied for several types of channels and one of the most important results is that the mutual information is maximized when the statistics of the channel input are Gaussian. This would require to design systems whose transmitted signal is Gaussian distributed, which might not be feasible. However, the expressions of channel capacity are widely used regarding that they provide a tight upper bound of the amount of information that can be sent through that channel. Mathematically speaking, many of the developments and considerations when optimizing target functions can be, without loss of generality, on their bounds.

On the one hand, the channel capacity for the SISO channel is well-known (e.g. [Cov72]), while the expression of the capacity for MIMO channels have received much attention with the last wave of MIMO systems. It is worthy to cite one of the first references in literature is Foschini [FG98] for MIMO channels in fading environment, Scharf [SM00] for his nice demonstration of the channel capacity relating rate and interference in a geometric fashion, and Goldsmith [GJJV03], for being one of the most complete reference on capacity for the MIMO channel.

On the other hand, characterizing the performance of multi-user systems is a more challenging problem compared to single-user systems. The maximum error-free transmission rate for a given power budget is a commonly adopted criterion, although find the minimum power to satisfy a target rate could be another approach. From the former point of view, single-user capacities are sufficient to characterize the multi-user channel if time domain multiplexing (TDM) is employed or random access methods are used in the uplink. However, if more sophisticated signal techniques are used, the full characterization of the channel requires the determination of the K-dimensional capacity region —with K the number of users—, where each point represent the set of K achievable rates associated to a given power allocation. The BC channel has been analyzed in terms of channel capacity region by Goldsmith [VJG03], Caire and Shamai [CS03], Cioffi [YC04].

3.4 Equalization and Precoding Techniques

This last section of the introductory part briefly summarizes the most important techniques used to face the effect of a channel. The use of vectored transmission notation defines a unified framework in which any channel may be represented by a matrix, $\boldsymbol{H}$ —from the the simple ISI SISO channel to the complex MIMO BC with coded OFDM—, so the objective of any equalization or precoding technique is to process, filter and manipulate the information and the signals, prior and after the transmission, to recover the transmitted information as reliable as possible. Hence, a concept that maybe relates any technique is the one of **perfect recoverability** of a MIMO channel.

A MIMO channel is said to be perfect left-recoverable if and only if all its inputs are perfect recoverable, i.e., there exists a non-negative integer, k_i, and an $R \times 1$ polynomial vector $\boldsymbol{g}(D)$ such that

$$\boldsymbol{g}(D)\boldsymbol{H}(D) = D^{k_i}\boldsymbol{e}_i, \tag{3.1}$$

where D denotes the formal D-transform (cf. [OS75]) and $\boldsymbol{e}_i$ is the i-th column if the identity matrix. In other words, the MIMO channel described by $\boldsymbol{H}(D)$ can be perfectly recovered with a delay of

k_i samples on the i-th output of the channel. Notice that this relation can be analogously stated by right-multiplying the channel matrix by a $T \times 1$ polynomial vector $\boldsymbol{f}(D)$, for which the $R \times T$ channel matrix would be right-recoverable. This formulation is the basis of the Bezóut Identity. Notice, additionally, that (3.1) is time-dependent (instant given by D), and that when processing a finite length block, i.e. finite length response (FIR) filtering, $\boldsymbol{f}$ and $\boldsymbol{g}$ become matrices.

The first traditional technique to equalize a channel —i.e. compensate its effects— is LE. Barbarossa [BS00], Scaglione [SSB+02] and Cioffi [PCL03] form a very extensive resource of LE for MIMO systems. The key resides in finding the optimal transmit and receive linear filters, $\boldsymbol{F}$ and $\boldsymbol{G}$, that equalize the channel matrix. The motivation of LE is given by the fact that the singular value decomposition (SVD) decomposes the channel into three matrices, from whom the outer ones can be compensated by the optimal filters, so that the resulting recovered channel is diagonal. The main drawback of LE is the noise enhancement —bigger when designing under ZF condition, e.g. [JUN05]— at the receiver.

With the aim of reducing the noise enhancement, DFE is often used at the place of LE. DFE consists in placing a feedback loop at the receiver where the detection is done iteratively (cf. [RG07]) —notice that since the maximum likelihood detection is not linear, DFE is considered a non-linear technique. As a consequence, the receive filter is implemented in its inverted form. However, error propagation can be a major concern in practical applications with DFE structure, especially if constellation-expanding codes, or convolution codes are used in concatenation with DFE, since the error rate on the inner DFE is worse prior to the decoder. Another technique that operates mainly at the transmitter side is V-BLAST [WFGV98]. This technique is analogous to DFE in multi-user scenarios for which the users are decoded iteratively.

Non-linear techniques such as THP have gained presence in the research society. First Forney [FE91] and later Fischer ([WFVH04] and [FWLH02]) give a performance comparison among different techniques for the MIMO channel including THP, placing the coding design at the transmitter and receiver —it is the case when the coding is done at the transmitter that we call it precoding, as it is done prior to transmission. Also, by construction, the equivalent matrix $\boldsymbol{F}$ is implemented in its inverse version.

Other processing techniques are based on the factorization of the channel, besides the SVD. These are, of importance, the Cholesky factorization [KJUB07], claimed to be a low-complexity technique for THP that avoids the high number of channel inversions when the number of user increases; and the QR decomposition extensively used by Cioffi in [GC00] and [GC02] to reduce the cross-talk interference in multi-user DSL bundles. On the one hand, the Cholesky factorization decomposes a matrix into to lower triangular matrices, very desirable in THP where the iterative precoding nature requires the feedback matrix to be lower triangular. On the other hand, QR factorizes the channel matrix into a triangular matrix and a unitary matrix. This implies that detection must be again done iteratively, which results to be unsuitable in non-cooperative scenarios such as the BC.

Finally, Zhang in [KWZ02] and [ZK05] provides a FIR solution to the Bezóut identity, for which it is known as LaBést (Layered Bézout Spatce-Time) equalization. Equation (3.1) does not give any information on the length in time —that is, the range of D— of the channel matrix neither the filters. As a consequence, the solution might the of infinite length. LaBést simply consists in truncating the infinite length solution in an iterative manner (layered). The key to Bezóut is defining the necessary algebra where the channel —which is treated as a polynomial matrix— can be perfectly recovered. With the aim of citing, Kung in [WK04] provided LE based on LaBést with imprecise CSI.

Part II

Temporal THP for the Single-User SISO Frequency-Selective Channel

Chapter 4

Introduction to Channel Equalization

4.1 Motivation

Tomlinson-Harashima precoding (THP) was first designed to reduce the inter-symbol interference (ISI) for the frequency-selective channel, [Tom71] and [HM72] in the early seventies. Despite of this, the non-linear technique introduced by Tomlinson and Harashima has recently received much attention as it is considered an implementation of dirty paper codng (DPC), motivated by [Cos83] in 1983. It has been demonstrated that the so-called dirty paper channel models well various important communication problems, among them precoding for ISI channels, digital watermarking and various broadcast schemes. This channel model considers a power constrained communication where part of the noise is known as side information, which is usually referred to as interference. The essence of these precoding techniques is not to directly transmit the user information, but a larger set of codes (DPC) or signals (THP) is designed to represent the user information so making use of side information. In fact, outlined by [ESZ05], knowledge about the state of the channel might be used by the transmitter in a dirty-channel scenario to achieve capacities near the additive white Gaussian noise (AWGN) channel.

The efficient utilization of a communication channel requires the simultaneously reduction of the interference and the independent random noise. Equalization methods try to mitigate these effects and several techniques have been described in the literature. Linear equalizers achieve relatively good performance, e.g. [JUN05], but they often show limited effectiveness and other structures must be designed. Decision feedback equalization (DFE) is intrinsically motivated by the ability to reduce the noise enhancement in linear equalization [RG07] and finite-length results shown better performance in front of linear equalization, as claimed in [ADC95].

However, error propagation can be a major concern in practical applications with DFE structure, especially if constellation-expanding codes, or convolution codes are used in concatenation with DFE, since the error rate on the inner DFE is worse prior to the decoder. The main idea of THP is to achieve an asymmetric design of the signal processing, [GS99], by operating the DFE at the transmitter side, where decision errors are indeed impossible. Evidently, the CSI knowledge paradigm is reinforced here and, in addition, the significant transmit-power increase that could result from placing the feedback filter at the transmitter is overcome by a modulo arithmetic, non-linearity, which is employed to bound the value of the transmitted symbol.

4.2 Objectives Outline

In this work, user information is initially precoded at the transmitter side with a feedback filter, $\boldsymbol{B}$, which makes uses of side information of the state of the channel —in this case, channel lags introduce ISI, though the mathematical approach allows to adopt other interferences such as inter-channel interference (ICI) and multi-user interference (MUI)— to prepare the transmitted symbol. The channel introduces both ISI and additive noise so that its output is post-equalized by the receiver filter $\boldsymbol{G}$. The use of the non-linearity is motivated, as outlined above, by the possible power increase that the feedback filter may introduce when the channel shows a deep fade. This operation is performed at both transmitter and receiver sides and a common known randomness or dither signal is also used to control the statistics of the transmitter signal and reduce the noise introduced by the non-linear operation.

A joint transmit-receive optimization of THP, detailed in chapter 6 under both minimum mean-square error (MMSE) and zero-forcing (ZF), is performed to face a single-user frequency-selective Rayleigh single-input single-output (SISO) channel. In addition, complete CSI is assumed as both transmitter and receiver have complete knowledge about the realization of the channel by frequently training the down and up links. The traditional optimization using Lagrange multipliers leads the following feedback and receiver filter expressions for both MMSE and ZF THP:

$$\begin{aligned} \boldsymbol{B} &= \mathrm{diag}^{-1}(\boldsymbol{\Phi}^{-1})(\boldsymbol{U}^T \odot \boldsymbol{\Phi}^{-1}) \\ \boldsymbol{G} &= \boldsymbol{B}\boldsymbol{\Phi}\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1} \end{aligned}$$

where $\boldsymbol{\Phi}$ is a design matrix that takes the following expression depending on MMSE or ZF criteria:

$$\begin{aligned} \boldsymbol{\Phi}_{MMSE} &= ({\boldsymbol{R_x}}^{-1} + \boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}\boldsymbol{H})^{-1} \\ \boldsymbol{\Phi}_{ZF} &= (\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}\boldsymbol{H})^{-1}. \end{aligned}$$

In former expressions, $\boldsymbol{U}$ is an all ones upper triangular matrix, $\boldsymbol{H}$ is the channel matrix, whereas $\boldsymbol{R_x}$ and $\boldsymbol{R_n}$ are the autocorrelation matrices of the transmitted signal and noise, respectively.

To illustrate the scope of THP in the single-user SISO channel, the uncoded symbol error rate (SER) curves for both MMSE and ZF THP compared to the AWGN channel limit have been obtained in a typical indoor office scenario —refer to third column in table 7.1. As it can be appreciated in figure 4.1, THP techniques outperform at high signal-to-noise ratio (SNR) compared to the reference AWGN channel thanks to the intelligent use of the multiple paths that the rich temporal diversity channel offers to the system. At low SNR, however, the effect of the modulo operation springs out reducing the gain in front of the AWGN channel.

The remainder of this part is organized as follows. In chapter 5 the complete model of the system is justified, and it is intended to provide a deep detail on the key points in THP such as the modulo operation and the use of the dither signal. Chapter 6 discusses the joint transmit-receiver optimization of the feedback and receiver filters under both MMSE and ZF criteria, making use of Lagrange multipliers as well as the very well-known result in signal processing of the orthogonality principle. Finally, some interesting results are depicted over chapter 7 where the design presented in this part is simulated in different wireless scenarios in order to analyze its performance as a precoding technique, as well as studying the collateral losses introduced by the modulo operation which will be present in any THP system in this thesis, regardless its use.

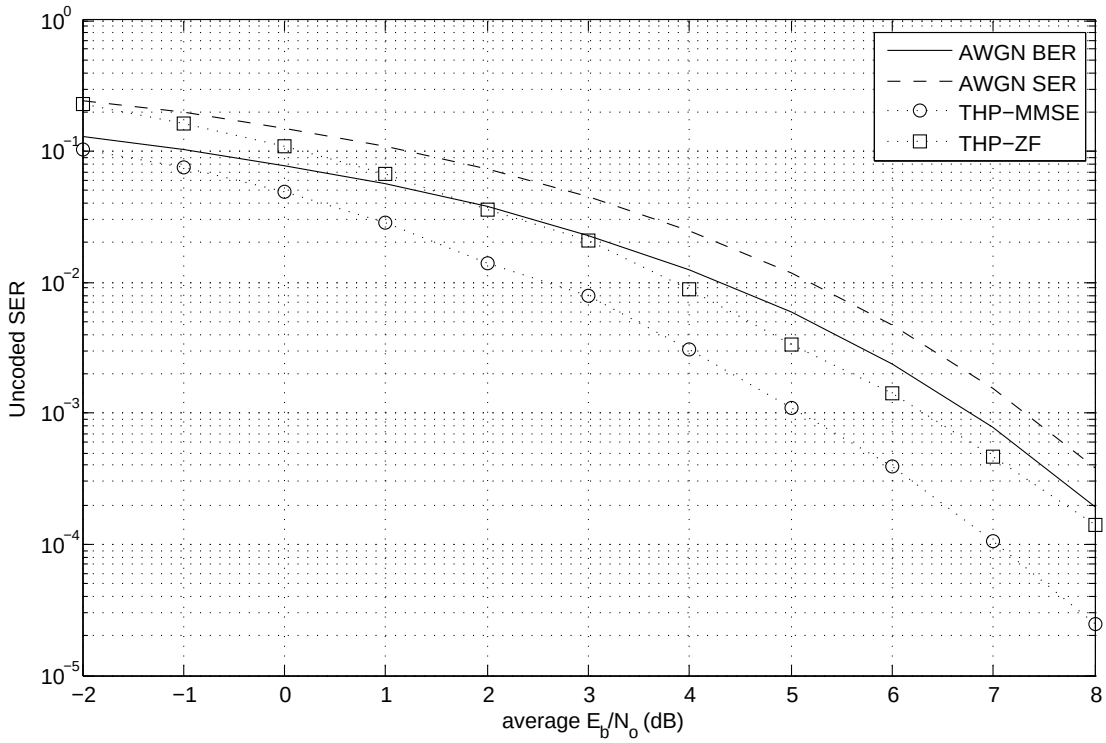


Figure 4.1: SER for the THP SISO channel, with high temporal diversity, versus average E_b/N_o.

Chapter 5

The SISO Frequency-Selective Channel

Uncoded user binary information, b_n, bijectively labels two-dimensional symbols belonging to a certain constellation, $s \in \mathbb{A} \subset \mathbb{C}$, which are grouped by a N-block symbol such as $\boldsymbol{s}_k = [s(kN)\ s(kN+1)\ \ldots\ s(kN+N-1)]^T \in \mathbb{C}^N$, where k indicates the k-th block. Signal processing techniques will be consequently applied to select the N optimal transmitted signals, grouped in $\boldsymbol{x}_k$, to face the channel and, from the received signal, process it to obtain the estimation of $\boldsymbol{s}_k$, $\hat{\boldsymbol{s}}_k$.

Following sections provide a description and justification on the channel, transmitter and receiver models. As a matter of notation, the subindex k will be omitted and the subindex m will denote the $m-th$ temporal symbol inside a vector, e.g. the scalar $s_m = s(kN+m-1)$, $1 \leq m \leq N$.

5.1 Channel Vectorization

Maybe one of the most relevant justification in studying the SISO channel is the adoption of the vectorial notation. The available signal at the output of a vector channel, i.e. at the input of the receiver, is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}. \tag{5.1}$$

where $\boldsymbol{x} \in \mathbb{C}^N$ is the transmitted signal (input), $\boldsymbol{y}$ is the received signal (output) and $\boldsymbol{n}$ the added noise, both 1-column vectors in $\mathbb{C}^L$.

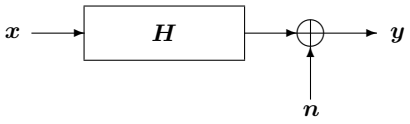


Figure 5.1: SISO channel

On the one hand, the channel matrix is set as a $L \times N$ matrix. In general, $L > N$ so its rank will be given by the number of linearly independent rows. This matrix can represent mostly all digital channels, like a flat-fading multiple-input multiple-output (MIMO) channel, a frequency-selective SISO channel or both. In the particular case of frequency-selective channel, each user is affected at least by his or her own former transmitted symbols, inducing to undesired ISI. The scalar channel impulse

response is given by the formal definition $H(D)$, $H(D) = \sum_\ell h_\ell D^\ell$ —the z-transform of a sequence or impulse response if adopted by the D-transform with $D = z^{-1}$, cf. [OS75]—, where h_ℓ denotes the channel realization of the ℓ-th lag, statistically modeled as a Rayleigh random variable. It will be assumed that the channel is causal with memory K, i.e. that $H(D) = h_0 + ...h_K D^K$. Since

$$y_m = \sum_{\ell=0}^{K} h_\ell x_{m-\ell} + n_m, \tag{5.2}$$

the channel matrix is built setting as columns the corresponding channel vectors at a certain instant of time. This means that the channel seen as a set of columns is time-variant. However, for practical purposes, here it will be assumed that the coherence time, i.e. the time during which the channel remains statistically invariant, is about the order or larger than the block transmission duration, $\tau_{coh} > NT$. As well, the transmitted block is sufficiently large to hold $D_s << NT$, being T the symbol duration and D_s the channel's delay spread —practically, $K < N$. Under this constraints, $\boldsymbol{H}$ is a tall-matrix such as

$$\boldsymbol{H} = \begin{bmatrix} h_0 & 0 & \dots & 0 \\ \vdots & h_0 & & \vdots \\ h_K & \vdots & \ddots & 0 \\ 0 & h_K & & h_0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & h_K \end{bmatrix}. \tag{5.3}$$

On the other hand, the noise signal is a vector of i.i.d. Gaussian complex variables with σ_n^2 variance. Hence, its correlation matrix reads $\boldsymbol{R_n} = \sigma_n^2 \boldsymbol{I}_L \in \mathbb{R}_+^L$.

5.2 Transmitter Model

DPC

Yet explained in the introductory part of this thesis, namely part I, THP is well known for being a DPC implementation, because it has a very simple and low complex precoding structure done at signal level rather than coding level. Is it important to keep in mind once more that the DPC approach, introduced by M. Costa in [Cos83] relies on the non-causal knowledge, by the transmitter, of the interference. Whatever is the interference nature, the knowledge of it is usually translated to the knowledge of the transmitted signal and the channel realization (CSI), since any other signals that cannot be estimated or predicted are considered noise.

This section focuses on the original scenario where the interference is ISI-like. The transmitter consists in a DFE or precoding, as depicted in figure 5.2. Mentioned above, the feedback matrix $\boldsymbol{B} \in \mathbb{C}^{N \times N}$ is in charge of diminishing the ISI, while the modulo operation is in charge of limiting the transmitted power. Specifically, the sequence $\boldsymbol{s}$ is processed by the modulo operator and fed back to obtain the transmitted signal $\boldsymbol{x}$ with limited energy per channel use, $E[\boldsymbol{x}^H \boldsymbol{x}] < \mathcal{E}$.

5.2.1 On the Modulo Operation

The modulo operation, $\boldsymbol{M}(\boldsymbol{w})$, in introduced to reduce the signal power increase that could be generated by the feedback loop to cancel out any interference. We assume that the inputs of the modulo operation, w_m, $\forall m$, are independently and identically uniformed distributed, and it can be shown

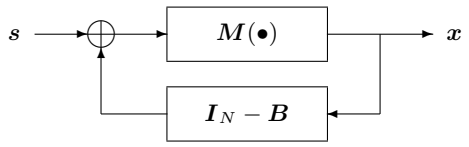


Figure 5.2: SISO Tomlinson-Harashima precoder

that the modulo operator output will also be independent and uniform. Therefore, the modulo operator $\boldsymbol{M}(\boldsymbol{w})$ consists in applying individual scalar modulo operations to each data symbol. That is, $\boldsymbol{M}(\boldsymbol{w}) = \text{diag}\left(M(w_1), \ldots, M(w_N)\right)$.

This individual operation was first presented by Tomlinson and Harashima, [Tom71, HM72], for one dimensional constellation, e.g. pulse amplitude modulation (PAM) constellation. For instance, in a PAM alphabet such as $\mathbb{A} = \{s_m \in \mathbb{R} : s_m \in \{\pm 1, \pm 3 \ldots \pm (M-1)\}\}$, the constellation was bounded and limited at, intuitively, $\pm M$; that is, setting the symmetry axis at half distance between symbols.

Mazo and Salz [MS76] first introduced the modulo operation for a two dimensional constellation, with symmetry at both dimensions. In this case of major concern, any input symbol may be regarded as a complex number, $s_m \in \mathbb{C}$, with an in-phase and quadrature components: $s_m = s_m^I + \jmath s_m^Q$. A further extension has been made by [Wei94], discussing a key point in THP: the design of the modulo operation since, indeed, it is constellation-dependent. Here we will address the modulo operation design for a QPSK constellation, though it is easily extensible to any other constellation. Since this particular constellation is formed by two symmetric symbols, we will define the complex modulo set as a bounded complex rectangular region where all transmitted symbols will lie on. Formally speaking,

$$\mathbb{M} \triangleq \left\{ x_m = x_m^I + \jmath x_m^Q \in \mathbb{C} : |x_m^I| \leq \tau_I/2, |x_m^Q| \leq \tau_Q/2, \right\}, \tag{5.4}$$

where the values of τ_I and τ_Q define the length of the complex set in the in-phase and quadrature components, respectively; and x_m^I and x_m^Q are the in-phase and quadrature components of the m-th transmitted symbol. It is evident that these parameters will play an important role in the transmitted power and they are critical parameters in the performance of the system. A very large value for τ_I and τ_Q would let the transmitter to achieve big peak-to-average power ratio (PAR) when an interference is needed to be canceled without the use of the modulo operation; whereas an small value would tight the average transmitted power but, whether the bounds are closed to a certain symbol, this one will probably show higher error probability due to the folding effect. Usually, and following the idea of Tomlinson and Harashima, the bounds are placed at a distance with respect to the outer symbol equal to the half distance of this symbol to its neighbor. In QPSK, clearly, $\tau_I = \tau_Q = 2\sqrt{2E_s}$.

The mathematical formulation of the modulo operation reads as follows:

$$M(w_m) = M(w_m^I + \jmath w_m^Q) = w_m - \left\lfloor \frac{w_m^I}{\tau_I} + \frac{1}{2} \right\rfloor \tau_I - \jmath \left\lfloor \frac{w_m^Q}{\tau_Q} + \frac{1}{2} \right\rfloor \tau_Q, \tag{5.5}$$

where $\lfloor \bullet \rfloor$ denotes the closest lower integer. Hence, from expression (5.5), it can be observed that the output of the modulo operator will be always restricted to the defined set $\mathbb{M}$, i.e. $\boldsymbol{x} \in \mathbb{M}^N$.

The modulo operation is a congruence relation of all possible values in $\mathbb{C}$ whose congruence class is the set $\mathbb{M}$. In other words, every possible input of the modulo operator, denoted by $w_m \in \mathbb{C}$, has

its particular representation in $\mathbb{M}$. Moreover, it defines a surjective function,

$$\begin{aligned} M : \mathbb{C} &\longrightarrow \mathbb{M} \\ w_m &\longmapsto x_m = w_m - a_m, \end{aligned}$$

since its values span its whole co-domain; that is, for every output $x_m \in \mathbb{M}$, there is at least one input element $w_m \in \mathbb{C}$ accomplishing the relation. The auxiliary complex scalar symbol a_m is automatically subtracted by the modulo operation. From (5.5), we can state that $a_m \in \mathbb{C}$ is build as a complex number whose real and imaginary parts are multiple of τ_I and τ_Q respectively. In other words, it follows $a_m = k_m^I \tau_I + \jmath k_m^Q \tau_Q$, $k_m^I, k_m^Q \in \mathbb{Z}$.

Modular arithmetic can be handled mathematically by introducing a congruence relation on the complex numbers that is compatible with the operations of the ring defined by the addition, subtraction, and multiplication. Let $w_m, w_n \in \mathbb{C}$. Then, the modulo addition $\oplus_M$ holds

$$M(w_m + w_n) = x_m \oplus_M x_n.$$

Proof. It is easy to see that, from the definition of M as a surjective function, w_n is allowed to be represented by

$$w_m = x_m + a_m.$$

Thus, writing each complex number with its in-phase and quadrature components:

$$\begin{aligned} M(w_m + w_n) &= M[(x_m + a_m) + (x_n + a_n)] \\ &= M[(x_m + x_n) + (a_m + a_n)] \\ &= M[(x_m + x_n) + (k_m^I \tau_I + \jmath k_m^Q \tau_Q + k_n^I \tau_I + \jmath k_n^Q \tau_Q)] \\ &= M[(x_m + x_n) + (k_m^I + k_n^I)\tau_I + \jmath(k_m^Q + k_n^Q)\tau_Q]. \end{aligned}$$

Since the term $(k_m^I + k_n^I)\tau_I$ is real multiple of τ_I, whereas $\jmath(k_m^Q + k_n^Q)\tau_Q$ is a pure imaginary term multiple of τ_Q, they will be removed by the modulo operator, finally giving:

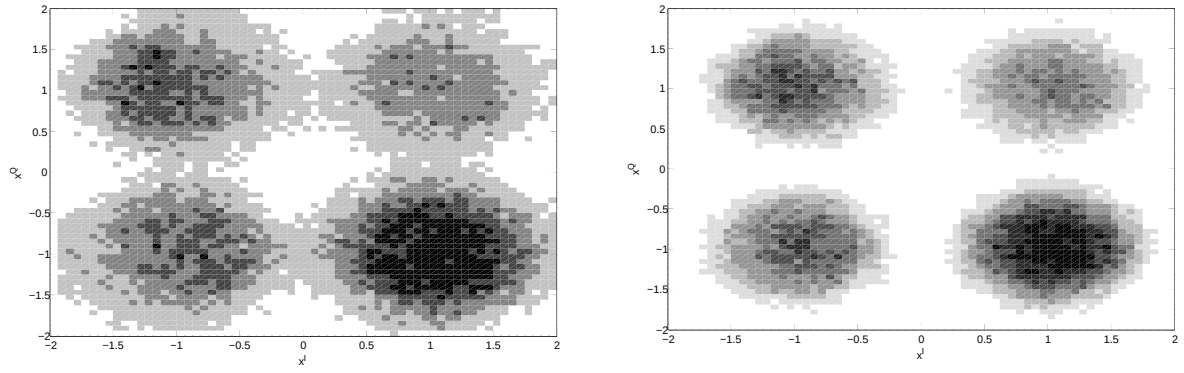
$$\begin{aligned} M(w_m + w_n) &= M[(x_m + x_n)] \\ &= x_1 \oplus_M x_2. \end{aligned}$$

□

However, the modulo operation cannot be distributed over the product in complex numbers. This is easy to see since the product of two complex numbers derives pure real components from two imaginary parts and pure imaginary components from cross products. For instance, when multiplying w_m by w_n, it appears the real term $-x_m^Q k_n^Q \tau_Q$ which may not be multiple neither τ_I or τ_Q. This means that the precoding loop must avoid scaling the input data signal by any factor, rather only subtracting its interference. In other words, **the main diagonal of the feedback filter must be all ones**, so that $b_{mm} = 1$ for all m.

5.2.2 First Approach to the Statistics of the Transmitted Signal

An interesting mathematical problem is to determine the probability distribution of $\boldsymbol{x}$ obeying the nonlinear recursion in figure 5.2. In general, one would not expect a probability density to exist, although



(a) (b)

Figure 5.3: Transmitted signal density plot for (a) ZF and (b) MMSE non-dithered THP.

the distribution could be very singular. As done by [MS76], $\boldsymbol{M}$ is regarded as a function whose input is the addition of $\boldsymbol{s}$ and the output of the feedback filter, denoted by $\boldsymbol{\theta}$. Using the definitions made above, $\boldsymbol{x} = \boldsymbol{M}(\boldsymbol{w}) = \boldsymbol{M}(\boldsymbol{s} + \boldsymbol{\theta})$, with $\boldsymbol{\theta} \triangleq (\boldsymbol{I}_N - \boldsymbol{B})\boldsymbol{x}$. Notice that it is not mathematically feasible to obtain the statistics of $\boldsymbol{\theta}$ because of the following reasons. Firstly, it depends on the feedback filter whose value will vary with the channel to precode at a certain instant of time; and, secondly, it depends on the transmitted symbol which, not only it is fed back by itself, but also it is the output of a non-linear operation. Under this situation, a first simulation tries to study these effects using the optimization filters obtained in chapter 6.

Up to 1200 Rayleigh channel realizations have been generated for a QPSK modulation for an average E_b/N_o of 0 dB. As depicted in figure 5.3, they obey roughly the channel statistics, which are Gaussian in both in-phase and quadrature components. Notice that in the ZF, since the feedback filter directly tries to cancel out the interference introduced by the channel, will show more variation with respect to the MMSE solution, which is more conservative and takes the noise into account. In this sense, the increase in transmitter power is higher in the ZF scheme. The second-order statistics have got a big importance, not only as a tool to compare the performance of different systems, but only, in our case, on the design of the optimum filters.

Though the transmitted signal conserves the Gaussian nature of the channel, as we increase the symbol energy and, therefore, probably the interference, the clipping effect on the bound of the modulo operation becomes notable. Trying to constantly estimate the statistics of the transmitted signal could result in waisting processing resources. A clever solution is to consider that the variance of the transmitted signal will always be upper-bounded by the uniform distribution over the modulo set (5.4). This leads the concept of dithering in THP to be introduced.

5.2.3 The Use of Dithering

Dithering was first presented by B. Widrow in [Sch64] in order to study its effect on the noise of quantization. Here we may notice the analogy between the THP transmitted signal and the quantization noise or error. As induced from (5.5), the output of the modulo operation is indeed the quantization of the input signal subtracted to itself. Namely, the quantization error reads $\epsilon = w - Q(w) = w - w_Q$ or, in our notation, $x = w - a$. Therefore any result on the quantization error may apply to the transmitted signal.

The early objective of dithering was achieving statistical independence between the input signal, w, and the quantization noise, x. In other words, $x(t, w) = x(t)$, result shown by Widrow in [Wid61].

The former condition derives, in the one-dimensional case, to the quantization theorem.

Theorem 1 (One-Dimensional Quantization Theorem)**.** *Let w be the input of the quantization process $Q(\bullet)$. We assume it is a continuous random variable with probability density function (PDF) $f_w(w)$ and characteristic function*

$$\phi_w(\lambda) = E[e^{\jmath\lambda x}],$$

where E denotes the mathematical expectation. If $\phi_w(\lambda) = 0$ for all $|\lambda| \geq 2\pi/\tau$, being τ the modulo range, then the probability density of the quantization error, $f_x(x)$ is uniform in the interval $[-\tau/2, \tau/2]$.

Though this is a very hard condition on the input signal, it can be always met if the proper dither signal is used. A dither signal is a second signal which is added to the input of the quantizer and then subtracted after the quantizing operation. In a communication system, this subtraction operation must take place in the receiver and implies the use of a known dither waveform and the requirement of synchronous information being transmitted to the receiver or the removal of the dither signal by filtering. Schuchman in [Sch64] and later Sripad and Snyder in [SS77] stated a necessary and sufficient condition in oder to obtain a uniform distribution of the quantization error:

Proposition 1. *The characteristic function of the input random variable satisfies*

$$\phi_w(2\pi n/\tau) = 0 \ \forall n \neq 0 \tag{5.6}$$

if and only if the density function of the quantization error, x, is uniform, i.e.

$$f_x(x) = \begin{cases} 1/\tau, & -\tau/2 \leq x < \tau/2 \\ 0, & \textit{otherwise.} \end{cases}$$

If this proposition holds, then the mean and variance of x are given by $E[x] = 0$ and $E[x^2] = \tau^2/12$. Obviously this condition is weaker that the one stated in the theorem. In our problem, this is equivalent to make the general assumption that the output elements are uniformly distributed over the $\mathbb{M}$ rectangle. Assuming that the in-phase and quadrature components are statistically independent, the variance is computed as $\sigma_x^2 = \sigma_{x^I}\sigma_{x^Q}$ being

$$\sigma_{x^I} = \sqrt{\int (x^I)^2 f_{x^I}(x^I) dx^I} = \sqrt{2\int_0^{\tau_I/2} (x^I)^2 \frac{1}{\tau_I}} = \frac{\tau_I}{\sqrt{6}},$$

and similarly for x^Q. Hence, the average transmitted power becomes, with the transmitted energy constraint,

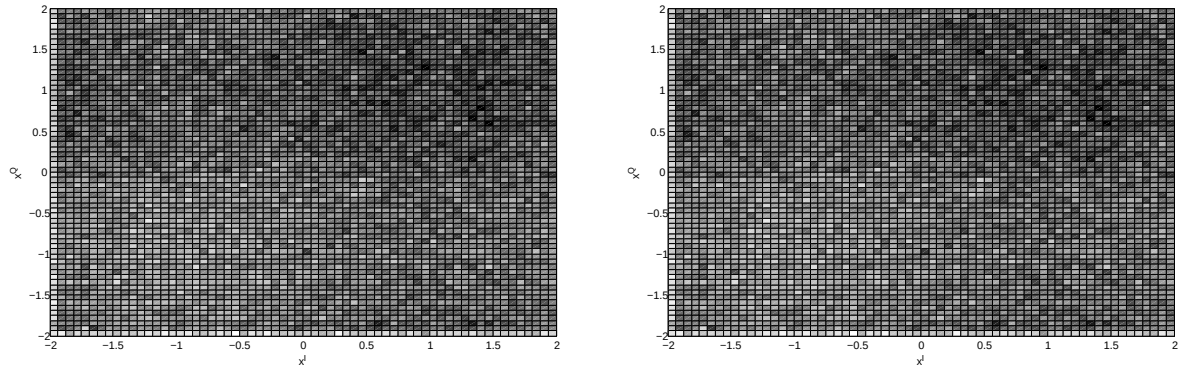
$$\sigma_x^2 = \frac{1}{6}\tau_I\tau_Q \leq \mathcal{E},$$

and, as a result, the correlation matrix of the signal $\boldsymbol{x}$ reads $\boldsymbol{R_x} = \frac{1}{6}\tau_I\tau_Q\boldsymbol{I}_N \in \mathbb{R}_+^N$.

The same simulation, at average E_b/N_o has been performed using the dithering signal proposed in this section. This is shown in figure 5.4 for both ZF and MMSE solutions, which in this case are very similar because the uniformity masks the specific behaviors of both solutions. As it can be seen, the transmitted signal is roughly uniformly distributed over the modulo set, in this example with $\tau = 4$.

The final model becomes $\boldsymbol{x} = \boldsymbol{M}(\boldsymbol{w} - \boldsymbol{u}) = \boldsymbol{M}(\boldsymbol{s} + (\boldsymbol{I}_N - \boldsymbol{B})\boldsymbol{x} - \boldsymbol{u})$, where $\boldsymbol{u}$ is the proper dither signal, shared by the transmitter and the receiver, and distributed uniformly over the modulo set:

$$f_u(u) = \frac{1}{\tau^I\tau^Q} \ |u^I| < \tau^I/2, \ |u^Q| < \tau^Q/2.$$



(a) (b)

Figure 5.4: Transmitted signal density plot for (a) ZF and (b) MMSE dithered THP.

This distribution accomplishes the condition (5.6) since its characteristic function is (for one component):

$$\phi_u(\lambda) = E[e^{\jmath\lambda u}] = \int_{-\infty}^{\infty} f_u(u)e^{\jmath\lambda u}du = \frac{\sin(\tau\lambda/2)}{\tau\lambda/2},$$

and hence,

$$\phi_u(2\pi n/\tau) = \frac{\sin(\pi n)}{\pi n} = 0 \ \forall n \neq 0,$$

which is a very well-known result in telecommunications. If not reinforced, all THP systems proposed in the sequel will be dithered due to its advantages.

5.3 Equivalent Transmitter Model

As justified in former section, the output of the modulo operator is simply the subtraction of its input and an auxiliary term denoted by $\boldsymbol{a}$, which ensures that all scalar entries of the output symbol are elements of $\mathbb{M}$. When taking this observation, we end up with the equivalent linear representation of the transmitter loop depicted in figure 5.5. As the auxiliary signal is included automatically by the modulo operator, we may follow from figure 5.2, adding the dithering signal justified in the former section:

$$\boldsymbol{x} = \boldsymbol{M}(\boldsymbol{w} - \boldsymbol{u}) = \boldsymbol{w} - \boldsymbol{a} - \boldsymbol{u} = \boldsymbol{s} + (\boldsymbol{I}_N - \boldsymbol{B})\boldsymbol{x} - \boldsymbol{a} - \boldsymbol{u},$$

which derives to

$$\boldsymbol{B}\boldsymbol{x} = \boldsymbol{s} - \boldsymbol{a} - \boldsymbol{u} \triangleq \boldsymbol{d}. \tag{5.7}$$

With this approach, we will use $\boldsymbol{d}$ as the desired signal in the following optimizations. Therefore, signals $\boldsymbol{x}$ and $\boldsymbol{d}$ become related by $\boldsymbol{x} = \boldsymbol{B}^{-1}\boldsymbol{d}$.

Notice that **to ensure the realizability of the feedback loop, matrix $\boldsymbol{B}$ needs to be lower triangular with all ones at the main diagonal**. This property is often called temporal causality, as only data symbols which have already been precoded —in other words, fed back and passed through the modulo operation— are fed. That is, the complex scalar symbol x_1 is constructed without feedback and it is equal to s_1, so the first row of $(\boldsymbol{I}_N - \boldsymbol{B})$ is zero, and consequently $\boldsymbol{B}$'s first element is one; whereas transmitted symbol x_m depends on s_m and $s_1, \ldots, s_{m-1}$ but not on itself, so the last element of the last row of $\boldsymbol{B}$ is one. We could also assume alternative orderings, like the reverse ordering and the resulting feedback matrix would be upper triangular with ones at the main diagonal instead.

The structure imposed to the feedback filter is crucial in the design of any THP, whatever is the

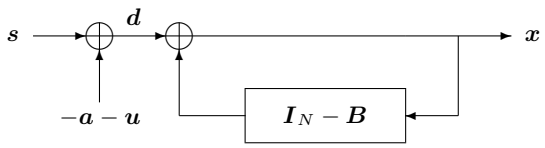


Figure 5.5: SISO Tomlinson-Harashima equivalent precoder

interference nature. Since the precoding is done at the transmitter side, what really interferes comes from the precoded signals, not the unprecoded data. Therefore, though in the temporal THP this is called temporal causality, in other systems where the interference comes from other sources —for instance other users— this kind of causality must apply as well. Specifically, this will be seen in the next two parts where, despite knowing the interference at the same time, an specific precoding order must be followed, leading always to a triangular feedback matrix.

5.4 Receiver Model

Finally, the receiver model can be appreciated in figure 5.6. Though the channel memory leads a longer length of the received sequence compared to its input, system receiver is designed to take the L available symbols to obtain the best N symbols. To do so, the received signal $\boldsymbol{y}$ is linearly processed by the forward matrix $\boldsymbol{G} \in \mathbb{C}^{N \times L}$. As a result, the available information at the input of the decider

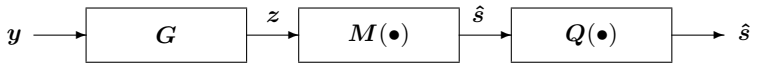


Figure 5.6: SISO Tomlinson-Harashima receiver

$\boldsymbol{Q}(\bullet)$ is $\hat{\boldsymbol{s}} = \boldsymbol{M}(\boldsymbol{G}\boldsymbol{y}) = \boldsymbol{M}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{x} + \boldsymbol{G}\boldsymbol{n})$. Notice that the noise is affected by the receiver filter and the modulo operation. At high SNR, the variance of the noise is very low compared to the signal variance, so the range modulo arithmetic set $\mathbb{M}$ is small compared to the typical deviation of $\boldsymbol{G}\boldsymbol{n}$. However, at low SNR, the modulo operation may play an important role on the statistics of the noise.

Regarding the operations done at the transmitter, firstly, the dither signal will be added by the decoder as it is a known sequence; and, secondly, the modulo operation is in charge of compensating the non-linear effect introduced by the transmitter. Accordingly, the operation will add the same signal $\boldsymbol{a}$ in absence of noise. Thus, the equivalent linear receiver model is shown in figure 5.7. The equivalent representation of the modulo operation at both sides, allows the problem to be overcome as a communication system from signal $\boldsymbol{d}$ to its estimate, $\hat{\boldsymbol{d}}$, where the signal power is automatically limited at the transmitter. In other words, signal $\boldsymbol{x}$ is power-bounded. Joining figures 5.5, 5.1 and 5.7, the expression that describes the signal processing and communication in the Tomlinson-Harashima equivalent system becomes

$$\hat{\boldsymbol{d}} = \boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}) = \boldsymbol{G}\boldsymbol{H}\boldsymbol{B}^{-1}\boldsymbol{d} + \boldsymbol{G}\boldsymbol{n}. \tag{5.8}$$

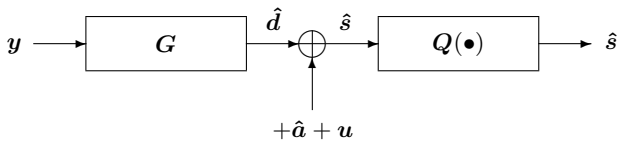


Figure 5.7: SISO Tomlinson-Harashima equivalent receiver

Chapter 6

Optimization Techniques

6.1 MMSE using the Orthogonality Principle

The THP model described in (5.8) is firstly optimized under the MMSE criterion. Finite-length filters will be used to process data at the transmitter, that is the feedback filter, and receiver. The causality in the feedback filter must be ensured, as explained in former section. The MSE, ε^2, is defined by $\varepsilon^2 \triangleq E\left[|\boldsymbol{d}-\hat{\boldsymbol{d}}|^2\right]$. The former expression can be read, using $\boldsymbol{d}=\boldsymbol{B}\boldsymbol{x}$ from (5.7) and $\hat{\boldsymbol{d}}=\boldsymbol{G}\boldsymbol{y}$ from (5.8), as

$$\begin{aligned}
\varepsilon^2(\boldsymbol{B},\boldsymbol{G}) &= E\left[|\boldsymbol{B}\boldsymbol{x}-\boldsymbol{G}\boldsymbol{y}|^2\right] \\
&= E\left[(\boldsymbol{B}\boldsymbol{x}-\boldsymbol{G}\boldsymbol{y})^H(\boldsymbol{B}\boldsymbol{x}-\boldsymbol{G}\boldsymbol{y})\right] \\
&= E\left[(\boldsymbol{x}^H\boldsymbol{B}^H-\boldsymbol{y}^H\boldsymbol{G}^H)(\boldsymbol{B}\boldsymbol{x}-\boldsymbol{G}\boldsymbol{y})\right] \\
&= E\left[\boldsymbol{x}^H\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{x}\right]-E\left[\boldsymbol{y}^H\boldsymbol{G}^H\boldsymbol{B}\boldsymbol{x}\right]-E\left[\boldsymbol{x}^H\boldsymbol{B}^H\boldsymbol{G}\boldsymbol{y}\right]+E\left[\boldsymbol{y}^H\boldsymbol{G}^H\boldsymbol{G}\boldsymbol{y}\right]. \qquad (6.1)
\end{aligned}$$

This four terms may be expressed using the trace operator using the following well-known result in signal processing. Let $\boldsymbol{x}\in\mathbb{C}^{1\times N}$ be a row vector, $\boldsymbol{A}\in\mathbb{C}^{N\times M}$ a matrix and $\boldsymbol{y}\in\mathbb{C}^{M\times 1}$ a column vector; then, the scalar quadratic form can be computed as the following expression:

$$\boldsymbol{x}\boldsymbol{A}\boldsymbol{y}=\mathrm{tr}(\boldsymbol{A}\boldsymbol{y}\boldsymbol{x})=\sum_{i=1}^{N}\sum_{j=1}^{M}x_i a_{ij} y_j,$$

being a_{ij} the ij-th element in matrix $\boldsymbol{A}$. This result allows the MMSE to be written as

$$\begin{aligned}
\varepsilon^2 &= E\left[\mathrm{tr}(\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{x}\boldsymbol{x}^H)\right]-E\left[\mathrm{tr}(\boldsymbol{G}^H\boldsymbol{B}\boldsymbol{x}\boldsymbol{y}^H)\right]-E\left[\mathrm{tr}(\boldsymbol{B}^H\boldsymbol{G}\boldsymbol{y}\boldsymbol{x}^H)\right]+E\left[\mathrm{tr}(\boldsymbol{G}^H\boldsymbol{G}\boldsymbol{y}\boldsymbol{y}^H)\right] \\
&= \mathrm{tr}(\boldsymbol{B}^H\boldsymbol{B}E[\boldsymbol{x}\boldsymbol{x}^H])-\mathrm{tr}(\boldsymbol{G}^H\boldsymbol{B}E[\boldsymbol{x}\boldsymbol{y}^H])-\mathrm{tr}(\boldsymbol{B}^H\boldsymbol{G}E[\boldsymbol{y}\boldsymbol{x}^H])+\mathrm{tr}(\boldsymbol{G}^H\boldsymbol{G}E[\boldsymbol{y}\boldsymbol{y}^H]) \\
&= \mathrm{tr}(\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{x}})-\mathrm{tr}(\boldsymbol{G}^H\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{x}\boldsymbol{y}})-\mathrm{tr}(\boldsymbol{B}^H\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{y}\boldsymbol{x}})+\mathrm{tr}(\boldsymbol{G}^H\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{y}}).
\end{aligned}$$

Here, the correlation matrices between the transmitted signal and the received signal have been defined, and obtained using (5.8), as

$$\begin{aligned} \boldsymbol{R_x} &\triangleq E\left[\boldsymbol{x}\boldsymbol{x}^H\right] \\ \boldsymbol{R_{xy}} &\triangleq E\left[\boldsymbol{x}\boldsymbol{y}^H\right] = \boldsymbol{R_x}\boldsymbol{H}^H \quad (6.2) \\ \boldsymbol{R_{yx}} &\triangleq E\left[\boldsymbol{y}\boldsymbol{x}^H\right] = \boldsymbol{H}\boldsymbol{R_x} \quad (6.3) \\ \boldsymbol{R_y} &\triangleq E\left[\boldsymbol{y}\boldsymbol{y}^H\right] = \boldsymbol{H}\boldsymbol{R_x}\boldsymbol{H}^H + \boldsymbol{R_n}. \quad (6.4) \end{aligned}$$

Finally, using the commutative trace property,

$$\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A}), \quad (6.5)$$

e.g. [MN99] being $\boldsymbol{A}$ and $\boldsymbol{B}$ be two matrix whose inner product exists, the MMSE target function to be minimized becomes

$$\varepsilon^2(\boldsymbol{B},\boldsymbol{G}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{R_x}\boldsymbol{B}^H) - \mathrm{tr}(\boldsymbol{B}\boldsymbol{R_{xy}}\boldsymbol{G}^H) - \mathrm{tr}(\boldsymbol{G}\boldsymbol{R_{yx}}\boldsymbol{B}^H) + \mathrm{tr}(\boldsymbol{G}\boldsymbol{R_y}\boldsymbol{G}^H). \quad (6.6)$$

The optimization problem consisting in minimizing (6.6) can be held by means of the orthogonality principle. When taking the derivative with respect to the receiver filter in the expression of the mean-square error (MSE), a relationship between the feedback filter and the receiver filter may be found. First, we express the MMSE in terms of the error signal, i.e. $\boldsymbol{e} \triangleq \boldsymbol{d} - \hat{\boldsymbol{d}}$, and the receiver filter transposed conjugate:

$$\varepsilon^2(\boldsymbol{B},\boldsymbol{G}) = E\left[\boldsymbol{e}^H\boldsymbol{e}\right] = E\left[(\boldsymbol{G}\boldsymbol{y} - \boldsymbol{B}\boldsymbol{x})^H\boldsymbol{e}\right]) = E\left[\boldsymbol{y}^H\boldsymbol{G}^H\boldsymbol{e}\right] - E\left[(\boldsymbol{B}\boldsymbol{x})^H\boldsymbol{e}\right].$$

Since the second part does not depend on the receiver filter, it will disappear with the derivative. Here we make use again of the trace operator properties to express the MMSE as

$$\varepsilon^2(\boldsymbol{B},\boldsymbol{G}) = E\left[\mathrm{tr}(\boldsymbol{G}^H\boldsymbol{e}\boldsymbol{y}^H)\right] + \varepsilon^2(\boldsymbol{B}). \quad (6.7)$$

Now we introduce an important result consisting in the derivative of the trace operator with respect to a matrix, e.g. [MN99]. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two matrix whose inner product exists. Hence,

$$\nabla_{\boldsymbol{A}}\mathrm{tr}(\boldsymbol{B}\boldsymbol{A}) = \boldsymbol{B}^T. \quad (6.8)$$

This result, applied to (6.7) follows

$$\nabla_{\boldsymbol{G}^H}\left(E\left[\mathrm{tr}(\boldsymbol{G}^H\boldsymbol{e}\boldsymbol{y}^H)\right] + \varepsilon^2(\boldsymbol{B})\right) = \nabla_{\boldsymbol{G}^H}E\left[\mathrm{tr}(\boldsymbol{e}\boldsymbol{y}^H\boldsymbol{G}^H)\right] = E\left[(\boldsymbol{e}\boldsymbol{y}^H)^T\right] = \boldsymbol{0}_{L\times N} \in \{0\}^{L\times N},$$

where $\boldsymbol{0}_{L\times N}$ is an $L \times N$ all zeros matrix. We can omit the transpose symbol in the relationship to write the **orthogonality principle** in this problem:

$$E\left[\boldsymbol{e}\boldsymbol{y}^H\right] = \boldsymbol{0}_{N\times L}. \quad (6.9)$$

From (6.9) a matrix relationship between the feedback filter and the forward filter can be obtained,

by means of two correlation matrices, as follows

$$\begin{aligned} E\left[\boldsymbol{e}\boldsymbol{y}^H\right] &= \mathbf{0}_{N\times L} \\ E\left[(\boldsymbol{d}-\hat{\boldsymbol{d}})\boldsymbol{y}^H\right] &= \mathbf{0}_{N\times L} \\ E\left[\boldsymbol{d}\boldsymbol{y}^H\right] &= E\left[\hat{\boldsymbol{d}}\boldsymbol{y}^H\right] \\ E\left[\boldsymbol{B}\boldsymbol{x}\boldsymbol{y}^H\right] &= E\left[\boldsymbol{G}\boldsymbol{y}\boldsymbol{y}^H\right] \\ \boldsymbol{B}E\left[\boldsymbol{x}\boldsymbol{y}^H\right] &= \boldsymbol{G}E\left[\boldsymbol{y}\boldsymbol{y}^H\right] \\ \boldsymbol{B}\boldsymbol{R}_{\boldsymbol{xy}} &= \boldsymbol{G}\boldsymbol{R}_{\boldsymbol{y}}. \end{aligned} \tag{6.10}$$

From the former expression, the receiver filter and its transposed conjugate read, respectively:

$$\boldsymbol{G} = \boldsymbol{B}\boldsymbol{R}_{\boldsymbol{xy}}{\boldsymbol{R}_{\boldsymbol{y}}}^{-1}, \tag{6.11}$$

$$\boldsymbol{G}^H = {\boldsymbol{R}_{\boldsymbol{y}}}^{-1}\boldsymbol{R}_{\boldsymbol{yx}}\boldsymbol{B}^H; \tag{6.12}$$

since the autocorrelation matrix is symmetric and real valued and ${\boldsymbol{R}_{\boldsymbol{xy}}}^H = E\left[\boldsymbol{x}\boldsymbol{y}^H\right]^H = E\left[\boldsymbol{y}\boldsymbol{x}^H\right] = \boldsymbol{R}_{\boldsymbol{yx}}$. As a result, the MSE can be exclusively expressed as a function of the feedback filter using the orthogonality principle. By plugging (6.11) and (6.12) into (6.6):

$$\begin{aligned} \varepsilon^2(\boldsymbol{B}) &= \operatorname{tr}(\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{x}}\boldsymbol{B}^H) - \operatorname{tr}(\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{xy}}{\boldsymbol{R}_{\boldsymbol{y}}}^{-1}\boldsymbol{R}_{\boldsymbol{yx}}\boldsymbol{B}^H) - \operatorname{tr}(\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{xy}}{\boldsymbol{R}_{\boldsymbol{y}}}^{-1}\boldsymbol{R}_{\boldsymbol{yx}}\boldsymbol{B}^H) \\ &+ \operatorname{tr}(\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{xy}}{\boldsymbol{R}_{\boldsymbol{y}}}^{-1}\boldsymbol{R}_{\boldsymbol{y}}{\boldsymbol{R}_{\boldsymbol{y}}}^{-1}\boldsymbol{R}_{\boldsymbol{yx}}\boldsymbol{B}^H). \end{aligned}$$

Noticing that the three last terms are the same, two common factors can be taken at both sides in terms of the feedback filter and its transposed conjugate. In other words,

$$\begin{aligned} \varepsilon^2(\boldsymbol{B}) &= \operatorname{tr}(\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{x}}\boldsymbol{B}^H) - \operatorname{tr}(\boldsymbol{B}\boldsymbol{R}_{\boldsymbol{xy}}{\boldsymbol{R}_{\boldsymbol{y}}}^{-1}\boldsymbol{R}_{\boldsymbol{yx}}\boldsymbol{B}^H) \\ &= \operatorname{tr}\left(\boldsymbol{B}(\boldsymbol{R}_{\boldsymbol{x}} - \boldsymbol{R}_{\boldsymbol{xy}}{\boldsymbol{R}_{\boldsymbol{y}}}^{-1}\boldsymbol{R}_{\boldsymbol{yx}})\boldsymbol{B}^H\right). \end{aligned}$$

Here we might define the matrix $\boldsymbol{\Phi}$ as the terms in between the feedback filters in the former expression:

$$\boldsymbol{\Phi} \triangleq \boldsymbol{R}_{\boldsymbol{x}} - \boldsymbol{R}_{\boldsymbol{xy}}{\boldsymbol{R}_{\boldsymbol{y}}}^{-1}\boldsymbol{R}_{\boldsymbol{yx}}. \tag{6.13}$$

The inversion lemma,

$$(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}+\boldsymbol{C}^{-1})^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}, \tag{6.14}$$

reduces the computation of matrix $\boldsymbol{\Phi}$. Using (6.3), (6.2) and (6.4), this matrix becomes

$$\boldsymbol{\Phi} = \boldsymbol{R}_{\boldsymbol{x}} - \boldsymbol{R}_{\boldsymbol{x}}\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{R}_{\boldsymbol{x}}\boldsymbol{H}^H + \boldsymbol{R}_{\boldsymbol{n}})^{-1}\boldsymbol{H}\boldsymbol{R}_{\boldsymbol{x}};$$

in which we identify the following matrices of the inversion lemma $\boldsymbol{A} = \boldsymbol{R}_{\boldsymbol{x}}$, $\boldsymbol{B} = \boldsymbol{H}^H$, $\boldsymbol{C} = {\boldsymbol{R}_{\boldsymbol{n}}}^{-1}$, and $\boldsymbol{D} = \boldsymbol{H}$. Therefore, plugging these results into (6.14) applied on the left sense, we may obtain

$$\boldsymbol{\Phi} = ({\boldsymbol{R}_{\boldsymbol{x}}}^{-1} + \boldsymbol{H}^H{\boldsymbol{R}_{\boldsymbol{n}}}^{-1}\boldsymbol{H})^{-1}. \tag{6.15}$$

Up to this point, minimize ε^2 is equivalent to minimize the following quadratic form

$$\varepsilon^2(\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{\Phi}\boldsymbol{B}^H) \tag{6.16}$$

subject to $\boldsymbol{B}$ being lower triangular with ones at the main diagonal (cf. 5.3). This condition may be written introducing $\boldsymbol{U}$ as an $N \times N$ upper diagonal matrix with all ones,

$$U \triangleq \begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{bmatrix}. \tag{6.17}$$

Hence, the feedback filter holds

$$\boldsymbol{B} \odot \boldsymbol{U} = \boldsymbol{I}_N. \tag{6.18}$$

It results interesting to express former constraint in terms of $\boldsymbol{B}^H$ by transpose conjugating (6.18), that is, $\boldsymbol{B}^H \odot \boldsymbol{U}^T = \boldsymbol{I}_N$. To determine the optimal filter, then, we shall make use of the Lagrange duality in this scalar convex problem with one $N \times N$ matrix constraint, [Fle87]. The Lagrangian reads

$$\mathcal{L}(\boldsymbol{B}) = \varepsilon^2(\boldsymbol{B}) - 2\mathrm{Re}\ \mathrm{tr}\left(\boldsymbol{\Lambda}(\boldsymbol{B}^H \odot \boldsymbol{U}^T - \boldsymbol{I}_N)\right), \tag{6.19}$$

with $\boldsymbol{\Lambda}$ the Lagrange matrix multiplier and $\varepsilon^2(\boldsymbol{B})$ the expression found in (6.16). Before the derivation, we shall show the distribution of the Hadamard product with respect to the inner matrix product. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two square matrices and $\boldsymbol{C}$ and $\boldsymbol{D}$ two diagonal matrices. Then,

$$\boldsymbol{C}(\boldsymbol{A} \odot \boldsymbol{B})\boldsymbol{D} = (\boldsymbol{C}\boldsymbol{A}) \odot \boldsymbol{B}\boldsymbol{D},$$

cf. [Mil07]. In our problem, $\boldsymbol{D}$ is the identity, $\boldsymbol{C}$ is the Lagrange multiplier matrix and $\boldsymbol{A}$ and $\boldsymbol{B}$ are the feedback filter and the all ones diagonal upper matrix. Applying this to the Lagrange duality function, (6.19) becomes

$$\mathcal{L}(\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{\Phi}\boldsymbol{B}^H) - 2\mathrm{Re}\ \mathrm{tr}\left((\boldsymbol{\Lambda} \odot \boldsymbol{U}^T)\boldsymbol{B}^H - \boldsymbol{I}_N)\right).$$

The derivative with respect to $\boldsymbol{B}^H$ will derive the optimal feedback filter in terms of $\boldsymbol{\Phi}$ and the Lagrange multiplier. Making use, again, of (6.8) the following relation

$$\nabla_{\boldsymbol{B}^H}\mathcal{L}(\boldsymbol{B}) = (\boldsymbol{B}\boldsymbol{\Phi})^T - (\boldsymbol{\Lambda} \odot \boldsymbol{U}^T)^T = \boldsymbol{0}_{N\times N} \tag{6.20}$$

gives a first result on the feedback filter:

$$\boldsymbol{B} = (\boldsymbol{\Lambda} \odot \boldsymbol{U}^T)\boldsymbol{\Phi}^{-1} = \boldsymbol{\Lambda}(\boldsymbol{U}^T \odot \boldsymbol{\Phi}^{-1}).$$

Finally, the constraint is applied as follows to obtain the Lagrange multiplier (6.18):

$$\begin{aligned} \boldsymbol{B} \odot \boldsymbol{U} &= (\boldsymbol{\Lambda} \odot \boldsymbol{U}^T)\boldsymbol{\Phi}^{-1} \odot \boldsymbol{U} \\ &= \boldsymbol{\Lambda}(\boldsymbol{U}^T \odot \boldsymbol{\Phi}^{-1} \odot \boldsymbol{U}) = \boldsymbol{I}_N. \end{aligned}$$

That is, $\boldsymbol{\Lambda} = \mathrm{diag}^{-1}(\boldsymbol{\Phi}^{-1})$. The effect of Hadamard multiplying by $\boldsymbol{U}^T$ and $\boldsymbol{U}$, at any order, results

in building a diagonal matrix whose diagonal is the diagonal of the multiplied matrix. As a result, the feedback matrix for the MMSE THP reads

$$\boldsymbol{B} = \mathrm{diag}^{-1}(\boldsymbol{\Phi}^{-1})(\boldsymbol{U}^T \odot \boldsymbol{\Phi}^{-1}). \tag{6.21}$$

The optimal forward filter is, from the orthogonality principle:

$$\boldsymbol{G} = \boldsymbol{B}\boldsymbol{\Phi}\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}. \tag{6.22}$$

This last expression is a compact version of the receiver filter and reduces the computation complexity of, instead, applying directly (6.11). The development has been performed as follows. From the orthogonality principle, the forward filter is related to the feedback filter, which is yet determined in function of matrix $\boldsymbol{\Phi}$, by means of the cross-correlation matrix between the transmitted and the received signal, as well as the inverse of the correlation matrix of the received signal. When plugging the results in (6.2) and (6.4) into (6.11), it looks like

$$\boldsymbol{G} = \boldsymbol{B}\boldsymbol{R_x}\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{R_x}\boldsymbol{H}^H + \boldsymbol{R_n})^{-1}. \tag{6.23}$$

Now, the expression corresponding to the inverse of the received signal correlation matrix is treated by the inverse lemma (6.14) in the forward direction to obtain

$$\boldsymbol{G} = \boldsymbol{B}\boldsymbol{R_x}\boldsymbol{H}^H \left({\boldsymbol{R_n}}^{-1} - {\boldsymbol{R_n}}^{-1}\boldsymbol{H}(\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}\boldsymbol{H} + {\boldsymbol{R_x}}^{-1})^{-1}\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1} \right).$$

If we place the first $\boldsymbol{H}^H$ element in former expression inside the outer parenthesis, the product $\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}$ becomes a common right factor. In other words,

$$\boldsymbol{G} = \boldsymbol{B}\boldsymbol{R_x} \left(\boldsymbol{I}_N - \boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}\boldsymbol{H}(\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}\boldsymbol{H} + {\boldsymbol{R_x}}^{-1})^{-1} \right) \boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}.$$

Noticing that the following term in former equation allows the inverse lemma to be applied backward to be written as

$$\boldsymbol{I}_N\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}\boldsymbol{H}(\boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}\boldsymbol{H} + {\boldsymbol{R_x}}^{-1})^{-1} = (\boldsymbol{I}_N + \boldsymbol{H}^H {\boldsymbol{R_n}}^{-1}\boldsymbol{H}\boldsymbol{R_x})^{-1},$$

the transmitted signal correlation matrix can enter the parenthesis by right multiplying —notice that the parenthesis is inverted so the correlation matrix enters as its inverse— all terms, so that (6.1) becomes $\boldsymbol{\Phi}$, and (6.22) is thus justified.

6.2 MMSE with Lagrange Duality

The solution of the minimum MSE for THP in a symmetric approach can be also determined directly applying the Lagrange duality to minimize ε^2 subject to the constraint (6.18). The result is obviously equivalent since the orthogonality principle has been derived from the derivative of the MMSE with respect to the receiver filter.

However, we shall address here the MMSE solution as a traditional optimization problem. The MSE is written as a function of the feedback and receiver filters using $\boldsymbol{d} = \boldsymbol{B}\boldsymbol{x}$ from (5.7), $\hat{\boldsymbol{d}} =$

$\boldsymbol{GHx} + \boldsymbol{Gn}$ from (5.8) and a similar development as in 6.1, as

$$\begin{aligned}\varepsilon^2(\boldsymbol{B},\boldsymbol{G}) &= E\left[|\boldsymbol{d}-\hat{\boldsymbol{d}}|^2\right] \\ &= E\left[|\boldsymbol{Bx}-\boldsymbol{GHx}+\boldsymbol{Gn}|^2\right] \\ &= \mathrm{tr}(\boldsymbol{BR_xB}^H) - \mathrm{tr}(\boldsymbol{BR_xH}^H\boldsymbol{G}^H) - \mathrm{tr}(\boldsymbol{GHR_xB}^H) + \\ &\quad \mathrm{tr}(\boldsymbol{GHR_xH}^H\boldsymbol{G}^H) + \mathrm{tr}(\boldsymbol{GR_nG}^H).\end{aligned}$$

The Lagrangian now requires the former expression of the MMSE and the constraint as follows:

$$\mathcal{L}(\boldsymbol{B},\boldsymbol{G}) = \varepsilon^2(\boldsymbol{B},\boldsymbol{G}) - 2\mathrm{Re}\ \mathrm{tr}\left(\boldsymbol{\Lambda}(\boldsymbol{B}^H\odot\boldsymbol{U}^T-\boldsymbol{I}_N)\right).$$

Taking into consideration both (6.5) and (6.8) results on the trace operator, the Lagrange function may be derived with respect to $\boldsymbol{B}^H$ and $\boldsymbol{G}^H$ to obtain these two equations

$$\begin{aligned}\nabla_{\boldsymbol{B}^H}\mathcal{L}(\boldsymbol{B},\boldsymbol{G}) &= \boldsymbol{R_x}^T\boldsymbol{B}^T - \boldsymbol{R_x}^T\boldsymbol{H}^T\boldsymbol{G}^T - \boldsymbol{\Lambda}^T\odot\boldsymbol{U} = \boldsymbol{0}_{N\times N} \\ \nabla_{\boldsymbol{G}^H}\mathcal{L}(\boldsymbol{B},\boldsymbol{G}) &= -\boldsymbol{H}^*\boldsymbol{R_x}^T\boldsymbol{B}^T + \boldsymbol{H}^*\boldsymbol{R_x}^T\boldsymbol{H}^T\boldsymbol{G}^T + \boldsymbol{R_n}^T\boldsymbol{G}^T = \boldsymbol{0}_{L\times N}.\end{aligned} \tag{6.24}$$

On the one hand, from the derivative with respect to the receiver filter, after transposing and taking $\boldsymbol{G}$ as a common left factor, we shall read $\boldsymbol{BR_xH}^H = \boldsymbol{G}(\boldsymbol{HR_xH}^H + \boldsymbol{R_n})$, or, equivalently, $\boldsymbol{G} = \boldsymbol{BR_xH}^H(\boldsymbol{HR_xH}^H + \boldsymbol{R_n})^{-1}$, which is exactly the same expression as in (6.23). When plugging this result into the derivative with respect to the feedback filter, after transposing all the equation, we obtain

$$\boldsymbol{BR_x} - \boldsymbol{BR_xH}^H(\boldsymbol{HR_xH}^H + \boldsymbol{R_n})^{-1}\boldsymbol{HR_x} = \boldsymbol{\Lambda}\odot\boldsymbol{U}^T.$$

It is then easy to see that if the feedback filter is taken as a left common factor in the left part of former equation, the following equation is obtained

$$\boldsymbol{B}[\boldsymbol{R_x} - \boldsymbol{BR_xH}^H(\boldsymbol{HR_xH}^H + \boldsymbol{R_n})^{-1}\boldsymbol{HR_x}] = \boldsymbol{\Lambda}\odot\boldsymbol{U}^T,$$

which is equivalent to (6.20) regarding the definition of matrix $\boldsymbol{\Phi}$ in (6.13). Therefore, the same filter will be determined here. That is, $\boldsymbol{B} = \mathrm{diag}^{-1}(\boldsymbol{\Phi}^{-1})(\boldsymbol{U}^T\odot\boldsymbol{\Phi}^{-1})$.

On the other hand, notice that if we right multiply the transposed derivative with respect to the feedback filter by the transposed conjugate channel matrix, it reads

$$\boldsymbol{BR_xH}^H - \boldsymbol{GHR_xH}^H - (\boldsymbol{\Lambda}\odot\boldsymbol{U}^T)\boldsymbol{H}^H = \boldsymbol{0}_{N\times L}.$$

The derivative with respect to the receiver filter, after transposing, becomes

$$-\boldsymbol{BR_xH}^H + \boldsymbol{GHR_xH}^H + \boldsymbol{GR_n} = \boldsymbol{0}_{N\times L}.$$

Finally, adding these two last equations, we obtain $\boldsymbol{GR_n} - (\boldsymbol{\Lambda}\odot\boldsymbol{U}^T)\boldsymbol{H}^H = \boldsymbol{0}_{N\times L}$, which, since $(\boldsymbol{\Lambda}\odot\boldsymbol{U}^T) = \boldsymbol{B\Phi}$ from (6.2), holds $\boldsymbol{G} = \boldsymbol{B\Phi H}^H\boldsymbol{R_n}^{-1}$.

6.3 Cholesky Factorization

An interesting approach consists of using an algebraic result known as Cholesky factorization. Let $\boldsymbol{X}$ be a Hermitian and definite positive square matrix —that is, all the eigenvalues are positive. Then,

it allows to be factorized as

$$\boldsymbol{X} = \boldsymbol{L}\boldsymbol{L}^H, \tag{6.25}$$

where $\boldsymbol{L}$ is a lower triangular matrix whose diagonal entries are all positive. This matrix is unique, given a matrix there is only one triangular matrix that allows this representation. Further more, if a matrix can be factorized like this, the matrix is then Hermitian and positive definite. In our work, the design matrix $\boldsymbol{\Phi}$ is factorized using the Cholesky decomposition.

6.3.1 Mathematical Results

Before entering into detail about how applying the Cholesky factorization to THP, several mathematical results are clarified in the following list.

(i) The squared Frobenius norm of a matrix is $||\boldsymbol{X}||^2$ and it is computed as the $\text{tr}(\boldsymbol{X}\boldsymbol{X}^H)$.

(ii) The product of two lower triangular matrices results another lower triangular matrix whose diagonal entries are the product of the diagonal entries of these two matrices.

Proof. Let $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ be two lower triangular matrices. Both admit the following decomposition $\boldsymbol{L}_i = \boldsymbol{D}_i + \boldsymbol{S}_i$, where $\boldsymbol{D}$ is a diagonal matrix containing the diagonal entries of $\boldsymbol{L}$ and $\boldsymbol{S}$ is an strictly inferior matrix. Then,

$$\boldsymbol{L}_1\boldsymbol{L}_2 = (\boldsymbol{D}_1 + \boldsymbol{S}_1)\cdot(\boldsymbol{D}_2 + \boldsymbol{S}_2) = \boldsymbol{D}_1\boldsymbol{D}_2 + \boldsymbol{D}_1\boldsymbol{S}_2 + \boldsymbol{S}_1\boldsymbol{D}_2 + \boldsymbol{S}_1\boldsymbol{S}_2.$$

Since the three last terms are all of them strictly inferior,

$$\text{diag}\,(\boldsymbol{L}_1\boldsymbol{L}_2) = \text{diag}\,(\boldsymbol{D}_1\boldsymbol{D}_2) = \text{diag}\,(\boldsymbol{D}_1)\text{diag}\,(\boldsymbol{D}_2).$$

□

(iii) Singular value decomposition (SVD). Let $\boldsymbol{X}$ be a complex $M \times N$ matrix. Then there exists a factorization of the form

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H,$$

where $\boldsymbol{U}$ is an $M \times M$ unitary matrix, $\boldsymbol{\Sigma}$ is an $M \times N$ matrix with nonnegative numbers on the diagonal and zeros off the diagonal, and $\boldsymbol{V}^H$ denotes the conjugate transpose of $\boldsymbol{V}$, an $N \times N$ unitary matrix. Matrix $\boldsymbol{\Sigma}$ contains the κ singular values of matrix $\boldsymbol{X}$, being κ its rank.

(iv) Eigenvalue decomposition (EVD). Also called spectral decomposition, it is the factorization of a matrix into a canonical form, whereby the matrix is represented in terms of its eigenvalues and eigenvectors. The fundamental theory of this decomposition lyes in the fact that any vector $\boldsymbol{q}$ is an eigenvector of a squared matrix $\boldsymbol{X}$ if and only if it satisfies $\boldsymbol{X}\boldsymbol{q} = \lambda\boldsymbol{q}$, where the scalar λ is the associated eigenvalue. If the set of eigenvectors, $\{\boldsymbol{q}_i\}_{i=1}^{\kappa}$ is linearly independent, then the matrix can be written as

$$\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1},$$

where the $i-th$ row of $\boldsymbol{Q}$ is the corresponding eigenvector $\boldsymbol{q}_i$ and $\boldsymbol{\Lambda}$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues, λ_i.

(v) Normal Matrix. A complex square matrix is a normal matrix if $\boldsymbol{X}\boldsymbol{X}^H = \boldsymbol{X}^H\boldsymbol{X}$. Notice that an Hermitian matrix (when $\boldsymbol{X}^H = \boldsymbol{X}$) is always normal.

(vi) Relation between SVD and EVD. The singular value decomposition is very general in the sense that it can be applied to any rectangular matrix. The EVD, on the other hand, can only be applied to certain classes of square matrices. Nevertheless, the two decompositions are related. In the special case that the matrix is normal, the singular values and the singular vectors coincide with the eigenvalues and eigenvectors of $\boldsymbol{X}$.

Proof. If a matrix is normal, its SVD accomplishes:

$$\boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{V}\boldsymbol{\Sigma}^2\boldsymbol{V}^H = \boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{U}^H = \boldsymbol{X}\boldsymbol{X}^H. \tag{6.26}$$

Consequently, the squares of the non-zero singular values of $\boldsymbol{X}$ are equal to the non-zero eigenvalues of either $\boldsymbol{X}^H\boldsymbol{X}$ or $\boldsymbol{X}\boldsymbol{X}^H$. □

(vii) Frobenius Norm of a singular value decomposed matrix. Let $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$ be the SVD of matrix $\boldsymbol{X}$. Then, its Frobenius norm can be written as $||\boldsymbol{X}||^2 = \sum_{i=1}^{\kappa}\sigma_i^2$, where κ is the matrix rank and σ_i are its singular values.

Proof. Noticing that the left and right matrices in the singular valued decomposition are unitary, i.e. $\boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{I}_M$ and $\boldsymbol{V}^H\boldsymbol{V} = \boldsymbol{I}_N$, and that the trace of two matrices can be shifted as $\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A})$, we can write

$$||\boldsymbol{X}||^2 = \mathrm{tr}(\boldsymbol{X}\boldsymbol{X}^H) = \mathrm{tr}(\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H\boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{U}^H) = \mathrm{tr}(\boldsymbol{U}\boldsymbol{\Sigma}^2\boldsymbol{U}^H) = \mathrm{tr}(\boldsymbol{\Sigma}^2\boldsymbol{U}^H\boldsymbol{U}) = \sum_{i=1}^{\kappa}\sigma_i^2.$$

□

(viii) The sum of eigenvalues is equal to the trace of the matrix. In other words, $\mathrm{tr}\boldsymbol{X} = \sum_{i=1}^{\kappa}\lambda_i$.

(ix) Weyl's Inequality. Written by [Wey49] in 1949, it states the eigenvalues and singular values of a square matrix have an equality relationship given by the following theorem.

Theorem 2 (The Weyl's Inequality)**.** *Let $\varphi(\lambda)$ be an increasing function of the positive argument λ: $\varphi(\lambda) \geq \varphi(\lambda')\forall\lambda \geq \lambda' \geq 0$, such that $\varphi(\lambda^\xi)$ is a convex function of ξ and $\varphi(0) \to 0$. Then, the squared of the eigenvalues λ_i of matrix $\boldsymbol{X}$, denote by λ_i, and the singular values of the Hermitian matrix obtained by $\boldsymbol{X}\boldsymbol{X}^H$, denote σ_i, arranged in decreasing order satisfy the inequality $\varphi(\lambda_1) + \ldots + \varphi(\lambda_\kappa) \leq \varphi(\sigma_1) + \ldots + \varphi(\sigma_\kappa)$. In particular:*

$$\sum_{i=1}^{\kappa}\lambda_i^s \leq \sum_{i=1}^{\kappa}\sigma_i^s, \tag{6.27}$$

for any real number s.

6.3.2 Process of Optimization

With all the points stated above, we can follow this development of the MSE —we temporally omit the constraint on the feedback matrix because it is implicitly set in the factorization— to determine the optimal feedback matrix. This development is now done in the single-user SISO case but, thanks to using the matrix and vectorial notation, it will be a very analog and powerful tool for parts III and IV for the MIMO channel and broadcast channel (BC).

From the general system equation, the MSE can be written in terms of the feedback filter and the well-known matrix $\boldsymbol{\Phi}$ (6.15) as $\varepsilon^2(\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{\Phi}\boldsymbol{B}^H)$. Let $\boldsymbol{\Phi} = \boldsymbol{L}\boldsymbol{L}^H$ be the Cholesky decomposition (6.25) of this matrix. Then, by (i), the MSE becomes

$$\varepsilon^2(\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{L}\boldsymbol{L}^H\boldsymbol{B}^H) = ||\boldsymbol{B}\boldsymbol{L}||^2. \tag{6.28}$$

Let now be the SVD decomposition (iii) of $\boldsymbol{B}\boldsymbol{L}$, $\boldsymbol{B}\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$. Then, using (vii) it holds

$$\varepsilon^2(\boldsymbol{B}) = ||\boldsymbol{B}\boldsymbol{L}||^2 = \sum_{i=1}^{\kappa} \sigma_i^2(\boldsymbol{B}\boldsymbol{L}). \tag{6.29}$$

Let the set $\{\lambda_i^2\}_{i=1}^N$ be the squared eigenvalues of the matrix product $\boldsymbol{B}\boldsymbol{L}$. By (ix), we can state that

$$\varepsilon^2(\boldsymbol{B}) = \sum_{i=1}^{N} \sigma_i^2(\boldsymbol{B}\boldsymbol{L}) \geq \sum_{i=1}^{N} \lambda_i^2(\boldsymbol{B}\boldsymbol{L}). \tag{6.30}$$

The known result in (viii) and the fact that it is a lower triangular matrix allows us to read

$$\varepsilon^2(\boldsymbol{B}) \geq \sum_{i=1}^{N} \lambda_i^2(\boldsymbol{B}\boldsymbol{L}) = \sum_{i=1}^{N} (\boldsymbol{B}\boldsymbol{L})_{ii}^2; \tag{6.31}$$

which finally taking into account that $\boldsymbol{B}$ has one main diagonal and (ii), the MSE becomes

$$\varepsilon^2(\boldsymbol{B}) \geq \sum_{i=1}^{N} l_{ii}^2, \tag{6.32}$$

where l_{ii} are the diagonal elements of matrix $\boldsymbol{L}$. Since the MSE is lower bounded, its minimum is achieved with the equality. By (v) and (vi), the equality satisfies when the enclosed matrix, $\boldsymbol{B}\boldsymbol{L}$ is normal. The only way that a triangular matrix is normal —notice that the product of a lower triangular matrix by an upper triangular matrix is a full matrix— is that this matrix is diagonal [cf. (ii)]. Therefore, we set the feedback filter as the inverse of the Cholesky matrix with the main diagonal scaled:

$$\boldsymbol{B} = \mathrm{diag}\ (l_{11} \dots l_{NN})\boldsymbol{L}^{-1}.$$

Finally, the receive filter that minimizes the MSE is given as well by (6.22), which finally reads

$$\boldsymbol{G} = \mathrm{diag}\ (l_{11} \dots l_{NN})\boldsymbol{L}^H\boldsymbol{H}^H{\boldsymbol{R}_{\boldsymbol{n}}}^{-1}.$$

6.4 MMSE-ZF with Lagrange Duality

The system model described in (5.8) is attractive to be treated as a ZF technique in the transmitter side as initially presented by Tomlinson and Harashima. Under this constraint, the optimization problem is reduced to select the filter at the receiver that minimizes the residual MSE due to the noise, that is why we refer to it as MMSE-ZF. The ZF condition can be seen as $\hat{\boldsymbol{d}}|_{\boldsymbol{n}\,=\,\mathbf{0}_{L\times 1}} = \boldsymbol{d}$, or equivalently, as usual, all the signal processing chain is the identity matrix, $\boldsymbol{G}\boldsymbol{H}\boldsymbol{B}^{-1} = \boldsymbol{I}_N$, resulting the ZF condition

$$\boldsymbol{G}\boldsymbol{H} = \boldsymbol{B}. \tag{6.33}$$

Given this relation, the MSE simply becomes $\varepsilon^2(\boldsymbol{G}) = \mathrm{tr}(\boldsymbol{G}\boldsymbol{R_n}\boldsymbol{G}^H)$. Nevertheless, the same condition on the feedback filter must be applied as in the MMSE criterion as justified. Specifically, we take both (6.18) and (6.33) in their transposed conjugated version. In other words, $\boldsymbol{U}^T \odot \boldsymbol{B}^H = \boldsymbol{I}_N$ and $\boldsymbol{H}^H\boldsymbol{G}^H = \boldsymbol{B}^H$. With this, the Lagrangian reads

$$\begin{aligned}\mathcal{L}(\boldsymbol{B},\boldsymbol{G}) &= \varepsilon^2(\boldsymbol{G}) - 2\mathrm{Re}\ \mathrm{tr}\left(\boldsymbol{\Lambda}(\boldsymbol{B}^H \odot \boldsymbol{U}^T - \boldsymbol{I}_N)\right) - 2\mathrm{Re}\ \mathrm{tr}\left(\boldsymbol{\Omega}(\boldsymbol{H}^H\boldsymbol{G}^H - \boldsymbol{B}^H)\right)\\ &= \mathrm{tr}(\boldsymbol{G}\boldsymbol{R_n}\boldsymbol{G}^H) - 2\mathrm{Re}\ \mathrm{tr}\left(\boldsymbol{\Lambda}(\boldsymbol{B}^H \odot \boldsymbol{U}^T - \boldsymbol{I}_N)\right) - 2\mathrm{Re}\ \mathrm{tr}\left(\boldsymbol{\Omega}(\boldsymbol{H}^H\boldsymbol{G}^H - \boldsymbol{B}^H)\right)\end{aligned}$$

Here $\boldsymbol{\Lambda}, \boldsymbol{\Omega} \in \mathbb{C}^{N\times N}$ are the Lagrange multipliers for both conditions. The derivatives with respect to both transposed conjugated filters are obtained using the known properties on the trace operator and the Hadamard product as follows:

$$\nabla_{\boldsymbol{B}^H}\mathcal{L}(\boldsymbol{B},\boldsymbol{G}) = -\boldsymbol{\Lambda}^T \odot \boldsymbol{U} + \boldsymbol{\Omega}^T = \boldsymbol{0}_{N\times N} \tag{6.34}$$

$$\nabla_{\boldsymbol{G}^H}\mathcal{L}(\boldsymbol{B},\boldsymbol{G}) = \boldsymbol{R_n}^T\boldsymbol{G}^T - \boldsymbol{H}^*\boldsymbol{\Omega}^T = \boldsymbol{0}_{L\times N}. \tag{6.35}$$

From the derivative with respect to the feedback filter $\boldsymbol{B}$, the following relationship between the two multipliers may be obtained: $\boldsymbol{\Lambda} \odot \boldsymbol{U}^T = \boldsymbol{\Omega}$. This result is then plugged into the transposed derivative with respect to the receiver filter to obtain a first expression of $\boldsymbol{G}$:

$$\boldsymbol{G} = (\boldsymbol{\Lambda} \odot \boldsymbol{U}^T)\boldsymbol{H}^H\boldsymbol{R_n}^{-1}, \tag{6.36}$$

which allows the feedback filter to be written, from the ZF condition in (6.33), as $\boldsymbol{B} = (\boldsymbol{\Lambda} \odot \boldsymbol{U}^T)\boldsymbol{H}^H\boldsymbol{R_n}^{-1}\boldsymbol{H}$. For commodity, we define here the inverse of the ZF matrix $\boldsymbol{\Phi}_{ZF}$ as

$$\boldsymbol{\Phi}_{ZF}^{-1} \triangleq \boldsymbol{H}^H\boldsymbol{R_n}^{-1}\boldsymbol{H}. \tag{6.37}$$

With this and applying the condition on the feedback filter (6.18), the Lagrange multiplier is solved and determined:

$$\begin{aligned}\boldsymbol{B} \odot \boldsymbol{U} &= \boldsymbol{I}_N\\ (\boldsymbol{\Lambda} \odot \boldsymbol{U}^T)\boldsymbol{\Phi}_{ZF}^{-1} \odot \boldsymbol{U} &= \boldsymbol{I}_N\\ \boldsymbol{\Lambda}(\boldsymbol{U}^T \odot \boldsymbol{\Phi}_{ZF}^{-1} \odot \boldsymbol{U}) &= \boldsymbol{I}_N\\ \boldsymbol{\Lambda} &= (\boldsymbol{U}^T \odot \boldsymbol{\Phi}_{ZF}^{-1} \odot \boldsymbol{U})^{-1}\\ \boldsymbol{\Lambda} &= \mathrm{diag}^{-1}(\boldsymbol{\Phi}_{ZF}^{-1}).\end{aligned}$$

Taking this result into (6.36), the optimal receiver filter for the ZF THP becomes

$$\boldsymbol{G} = \mathrm{diag}^{-1}(\boldsymbol{\Phi}_{ZF}^{-1})(\boldsymbol{U}^T \odot \boldsymbol{H}^H\boldsymbol{R_n}^{-1}),$$

whereas the optimal feedback filter may be now easily determined through (6.33).

Notice that the optimal feedback filter can be also expressed as

$$\boldsymbol{B} = \mathrm{diag}^{-1}(\boldsymbol{\Phi}_{ZF}^{-1})(\boldsymbol{U}^T \odot \boldsymbol{\Phi}_{ZF}^{-1}), \tag{6.38}$$

which is the same structure as the MMSE in (6.21). Given the feedback filter, the receiver filter

$\boldsymbol{G}_{ZF-THP}$ can be determined using the ZF condition applying the pseudo-inverse of the channel as

$$\boldsymbol{G} = \boldsymbol{B}(\boldsymbol{H}^H\boldsymbol{H})^{-1}\boldsymbol{H}^H. \tag{6.39}$$

6.5 MMSE-ZF derived from MMSE at high SNR

A very elegant methodology to determine the ZF filters is to employ a usual foresight in communications consisting in taking the MMSE solution at high SNR, because the ZF is known to neglect the effect of the noise and concentrate in uniquely compensating the channel. Let ξ be a parameter that reflects the inverse of the SNR, that is $\xi \triangleq \frac{\sigma_n^2}{\sigma_x^2}$. The samples of the transmitted signal and noise are assumed to be statistically independent, so that, as mentioned above, their correlation matrices become white: $\boldsymbol{R_x} = \sigma_x^2\boldsymbol{I}_N$ and $\boldsymbol{R_n} = \sigma_n^2\boldsymbol{I}_L$.

Firstly, the ZF condition is derived from the orthogonality principle (6.9). Reminding (6.2), $\boldsymbol{R_{xy}} = \boldsymbol{R_x}\boldsymbol{H}^H$, and taking the limit on the received signal autocorrelation matrix:

$$\begin{aligned}
\lim_{\xi\to 0}\boldsymbol{R_y} &= \lim_{\xi\to 0}\boldsymbol{H}\boldsymbol{R_x}\boldsymbol{H}^H + \boldsymbol{R_n} \\
&= \lim_{\xi\to 0}\sigma_x^2\boldsymbol{H}\boldsymbol{H}^H + \sigma_n^2\boldsymbol{I}_L \\
&= \lim_{\xi\to 0}\sigma_x^2(\boldsymbol{H}\boldsymbol{H}^H + \sigma_n^2\sigma_x^{-2}\boldsymbol{I}_L) \\
&= \lim_{\xi\to 0}\sigma_x^2(\boldsymbol{H}\boldsymbol{H}^H + \xi\boldsymbol{I}_L) \\
&= \boldsymbol{H}\boldsymbol{R_x}\boldsymbol{H}^H,
\end{aligned}$$

this last result allows the limit of (6.10) to be read as

$$\begin{aligned}
\lim_{\xi\to 0}\boldsymbol{B}\boldsymbol{R_{xy}} &= \lim_{\xi\to 0}\boldsymbol{G}\boldsymbol{R_y} \\
\boldsymbol{B}\boldsymbol{R_x}\boldsymbol{H}^H &= \boldsymbol{G}\boldsymbol{H}\boldsymbol{R_x}\boldsymbol{H}^H \\
\boldsymbol{B} &= \boldsymbol{G}\boldsymbol{H},
\end{aligned}$$

which is nothing more than the ZF condition obtained in (6.33). Making use of the pseudo-inverse of the channel matrix, the same expression as in (6.39) is obtained, $\boldsymbol{G} = \boldsymbol{B}(\boldsymbol{H}^H\boldsymbol{H})^{-1}\boldsymbol{H}^H$.

Secondly, the MMSE matrix $\boldsymbol{\Phi}$ defined in (6.15) derives to the expression of the ZF matrix, $\boldsymbol{\Phi}_{ZF}$ obtained in (6.37). In other words,

$$\begin{aligned}
\boldsymbol{\Phi}_{ZF} &= \lim_{\xi\to 0}\boldsymbol{\Phi} \\
&= \lim_{\xi\to 0}(\boldsymbol{R_x}^{-1} + \boldsymbol{H}^H\boldsymbol{R_n}^{-1}\boldsymbol{H})^{-1} \\
&= \lim_{\xi\to 0}(\sigma_x^{-2}\boldsymbol{I}_N + \sigma_n^{-2}\boldsymbol{H}^H\boldsymbol{H})^{-1} \\
&= \lim_{\xi\to 0}\sigma_n^2(\sigma_n^2\sigma_x^{-2}\boldsymbol{I}_N + \boldsymbol{H}^H\boldsymbol{H})^{-1} \\
&= \lim_{\xi\to 0}\sigma_n^2(\xi\boldsymbol{I}_N + \boldsymbol{H}^H\boldsymbol{H})^{-1} \\
&= \boldsymbol{H}^H\boldsymbol{R_n}^{-1}\boldsymbol{H},
\end{aligned}$$

which is, as expected, the same expression as in (6.37). Finally, the ZF feedback filter holds the same

structure as the MMSE solution with this new matrix, as obtained as well in (6.38):

$$\begin{aligned} \boldsymbol{B} &= \lim_{\xi \to 0} \text{diag}^{-1}(\boldsymbol{\Phi} - 1)(\boldsymbol{U}^T \odot \boldsymbol{\Phi}^{-1}) \\ &= \text{diag}^{-1}(\boldsymbol{\Phi}_{ZF}^{-1})(\boldsymbol{U}^T \odot \boldsymbol{\Phi}_{ZF}^{-1}). \end{aligned}$$

6.6 QR Decomposition

Finally, in schemes where the precoding or pre cancelation is done in an iterative manner, factorizations such as the QR decomposition have been employed by tradition. Nice examples of its utilization are V-BLAST (e.g. [WFGV98]) and cross-talk cancellation for digital subscriber lines (DSLs) (e.g. [GC02]), whose feedback filters require a lower triangular structure as well. In any case, the channel matrix is directly decomposed using the QR decomposition as follows:

$$H = \boldsymbol{Q}\boldsymbol{R}. \tag{6.40}$$

Here, $\boldsymbol{Q} \in \mathbb{C}^{L \times N}$ is an orthonormal matrix, i.e. $\boldsymbol{Q}^H\boldsymbol{Q} = \boldsymbol{I}_N$ and $\boldsymbol{R}$ is the $N \times N$ right lower triangular matrix. In view of this result, the feedback matrix is in charge of canceling the lower triangular part —and then, the precoding is based on the pivot technique— and the receiver filter the orthonormal matrix part. In other words, the feedback filter reads

$$\boldsymbol{B}_{QR} = \text{diag } (r_{11}^{-1} \dots r_{NN}^{-1})\boldsymbol{R},$$

whereas the receive filter is obtained from the ZF condition as

$$\boldsymbol{G}_{QR} = \text{diag } (r_{11}^{-1} \dots r_{NN}^{-1})\boldsymbol{Q}^H.$$

The scaling factor on the main diagonal is done to force that the feedback filter has one main diagonal as requested by design. This approach is simple and clear, since we force the channel to be decomposed into a triangular matrix (whose inverse is also triangular), the structure that we require for $\boldsymbol{B}$. However it is not optimum in the sense of MSE but, as it can be seen in the simulation, it achieves good performance with respect to the MMSE-ZF solution.

Chapter 7

Performance Analysis

The simplicity inherent within a single-user SISO system limits the results to analyze in quantity, but not in quality. Each of the developments and backgrounds introduced in this part should master the reader on such contents by their applicability to many systems and by noticing that all this work will lightweight future studies on THP. In the following channels studied in this thesis, new parameters will play a significant part, but the former chapters —specially on the modulo operation and the use of dithering— as well as the following results —such as the performance losses and the practical effect of dithering— are of a big boundlessness.

7.1 Metrics

Various simulations in MATLAB have been done to check the performance of the joint transmit-receive optimization scheme of THP with complete CSI in the SISO Rayleigh channel. The ℓ-th lag of the channel, h_ℓ, is obtained as a zero-mean Gaussian complex number whose variance is given by the PDP of the environment. In order words, let PDP_ℓ be an exponential power delay profile for the Rayleigh channel as $\mathrm{PDP}_\ell(T/D_s) = e^{-\ell T/D_s}$, where T/D_s is the symbol period over channel delay spread ratio, so the ℓ-th impulse response sample becomes

$$h_\ell = \mathcal{N}(0, \mathrm{PDP}_\ell/2) + \jmath\mathcal{N}(0, \mathrm{PDP}_\ell/2).$$

Hence, for instance, we will have a very low ratio (approaching zero) for very dispersive channels, whereas a channel with low power delays will have larger ratios. In the sequel, we will consider three environments with $D_s = T$, $D_s = 2T$ and $D_s = 5T$, accomplishing always the premise that $D_s << NT$. The power delay profiles resulting from these rations are depicted in table 7.1. The first column shows a channel with very low temporal diversity, so the system is not expected to achieve very good improvements with respect to the AWGN channel. Instead, the third column corresponds to a typical office environment with notable number of multi paths with no line of sight —this model has been adopted in HIPERLAN radio network standard in [MS98].

MATLAB simulations on a QPSK modulation on both MMSE and ZF designs are done to obtain performance results on the uncoded or raw SER and the normalized mean-square error (NMSE) is computed in terms of the user signal as

$$\bar{\varepsilon^2} = \frac{|\boldsymbol{s} - \hat{\boldsymbol{s}}|^2}{|\boldsymbol{s}|^2},$$

ℓ	$D_s = T$	$D_s = 2T$	$D_s = 5T$
0	0	0	0
1	-4.3	-2.1	-0.8
2	-8.7	-4.3	-1.7
3	-13.0	-6.5	-2.6
4		-8.7	-3.4
5		-10.8	-4.3
6		-13.0	-5.2
7			-6.0
8			-6.9
9			-7.8
10			-8.6
11			-9.5
12			-10.4
13			-11.2
14			-13.0

Table 7.1: Rayleigh channel exponential PDP in dB for different delay spread over symbol duration ratio.

while the SER is computed by Monte-Carlo method.

7.2 Performance Losses of THP

Prior to analyzing how THP performs for the single-user SISO channel in terms of SER and MSE, we shall present a particular classification of types of losses caused by precoding, as THP, introduced by [SL96]. These losses try to quantify the effect that THP has as a result of its particularities such as the modulo operation. Interestingly, if not treated, these three losses will be existent in any THP system, including the MIMO and BC parts of this thesis.

In THP, three types of losses shall be distinguished:

1. **Shaping loss.** This loss is associated to the statistics of the channel input signal, i.e. the transmitted signal. After the modulo operation, the channel input will be uniformly distributed over the modulo set. As we know, to achieve the AWGN channel capacity, the channel input must be Gaussian distributed. Thus, the uniformly distributed channel input introduces shaping loss. Accordingly, there is also a shaping loss at the receiver due to the receive filter, so the input of the detector might not be Gaussian. This effect dominates at high SNR.

2. **Power loss.** In THP, the power loss is caused by the precoding at the transmitter. Since the precoding tries to compensate the interference introduced by the channel, a power increase is arbitrary expected. Moreover, in dithered THP, the transmitted power is fixed by the uniform distribution to $\tau^2/6$, which is always greater than any other distribution. Compared to a non-precoded QPSK transmission, the power loss is roughly of 1.25 dB.

3. **Modulo loss.** This loss is caused by the modulo operations at the system, both transmitter and receiver. Due to the noise, the signal at the boundary of the constellation may be folded into the opposite boundary of the constellation. When this happens, an error that would not have occurred in regular channel is produced. This effect is notable at low SNR, as seen in high diversity scenarios.

In order to appreciate the former losses on a practical point of view, a full simulation of the THP over the channel with delay spread two times the symbol duration has been performed to obtain the

uncoded SER versus average E_b/N_o in dB. The result is depicted in figure 7.1 and allows the reader to attain the following considerations.

First of all, as expected, the MMSE solution is better than the ZF solution at any point of SNR. This is because the MMSE takes into consideration the noise signal introduced by the channel when determining the feedback and receive filters; whereas the ZF solution first tries to fully equalize the channel, producing the usual enhance of noise power at reception. Secondly, at low SNR, the effect of modulo loss is appreciated when comparing the SER curves to the AWGN channel ones, as the latter is better. As explained above, this is due to the folding effect caused by the presence of the modulo operation, again more significant in the ZF solution. Finally, at high SNR, the SER of both ZF and MMSE improve with respect to the AWGN channel due to the system's capacity in making an intelligent use of the diversity. This effect will be seen in more detail in the next section. However, it is worthy to reinforce that the higher the SNR, the higher the shaping loss at the input of the detector because the Gaussianity is less with the absence of noise. Again, this effect is more important in the ZF, not because the noise can be neglected at high SNR —which should approach both solutions—, but because the MMSE solution is more clever because it will not equalize the full chain, it will obtain the optimal filters to minimize the MSE, which is translated to minimize the SER.

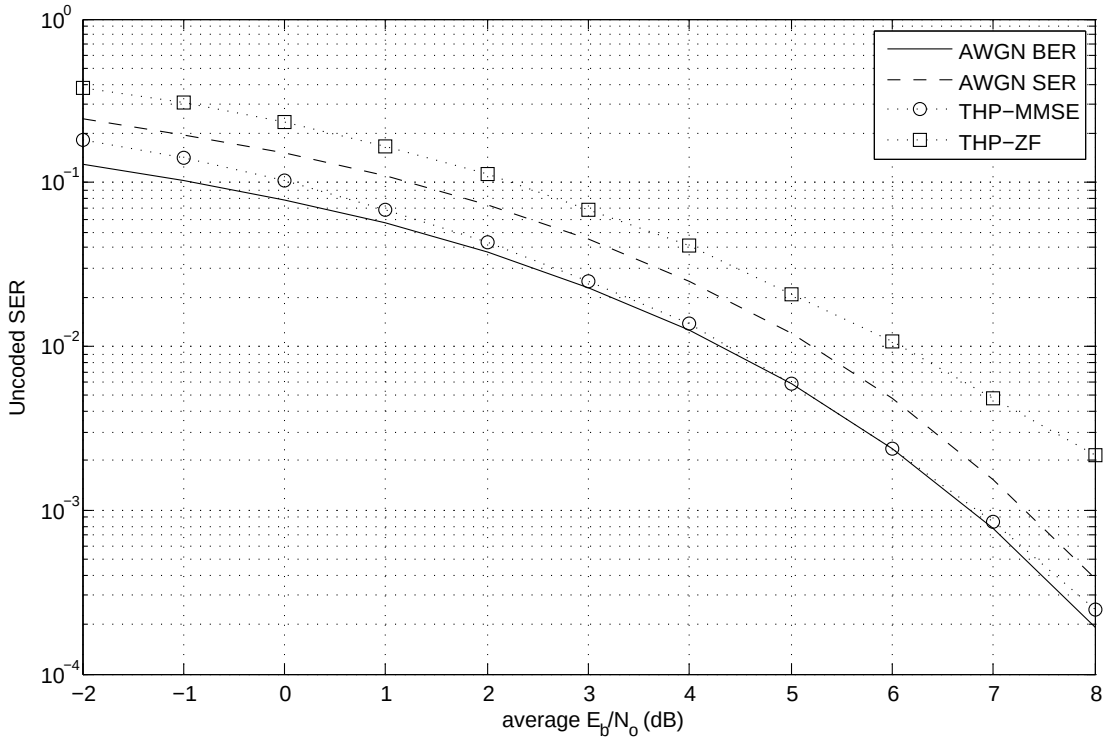


Figure 7.1: SER for the THP SISO channel, with $D_s = 2T$, versus average E_b/N_o.

7.3 The Effect of Temporal Diversity

The most relevant result in temporal THP is how the system adapts to the environment and makes a clever use of the diversity that the channel provides. To do so, the three channels described by their PDP in table 7.1 have been subject to simulation with THP for both MMSE and ZF solutions, and figures 7.2 and 7.3 respectively depict the uncoded SER and the MSE versus the average SNR.

The higher the diversity level —in this case it corresponds to channel (c)—, the better the performance is achieved by the temporal THP in terms of both SER and MSE. This is because the Tomlinson-Harashima feedback loop not only diminishes the apparent ISI, but also feeds back the user information (precoding) by favoring the transmission in a way such that the information is constructively repeated. Traditionally the ISI channel has been designed under ZF constraint because the information received through other channel lags has been considered as interference. However, the MMSE solution outperforms the ZF —this can be seen in almost all the simulations in this part— because regards the temporal lags as a possible source of diversity. And it does so.

Finally, as a conclusion, in a simple scenario with very low diversity, e.g. line of sight, it results not useful to use THP since it would not outperform compared to another signal processing technique, which would even show less precoding losses. In this sense, as it will be seen, the THP has only sense in scenarios where more than one signal is transmitted and provokes a mix of them, as it does the temporal ISI, the spatial ICI and the spatial MUI.

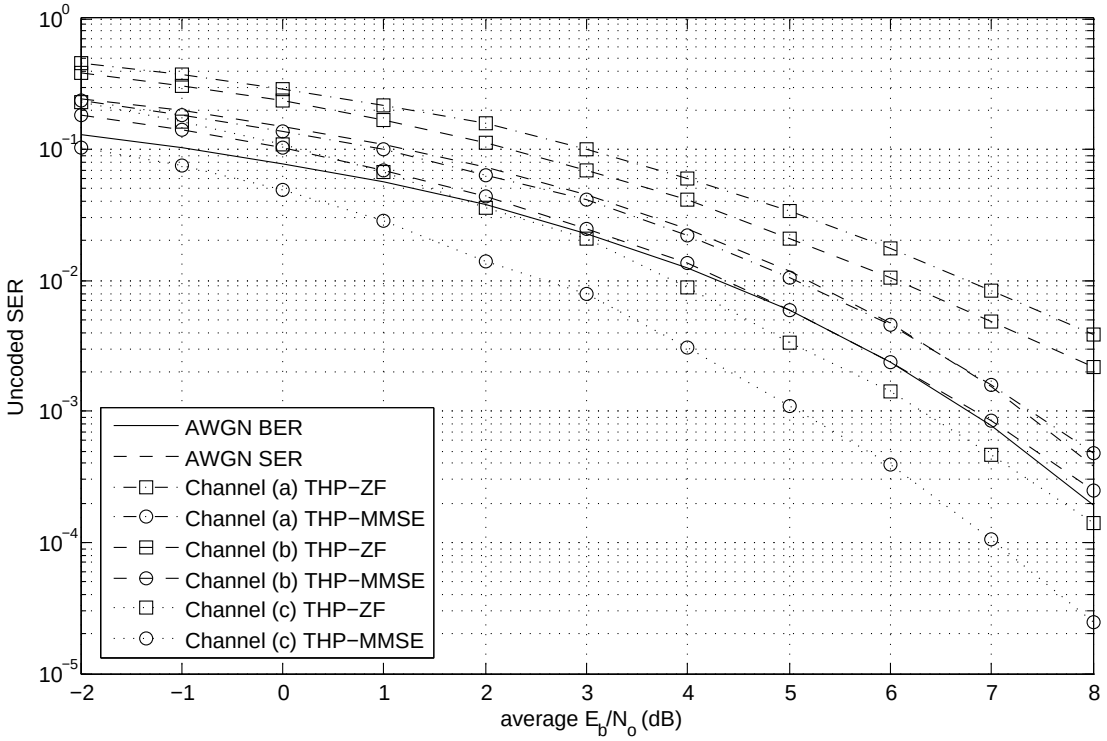


Figure 7.2: SER for the THP SISO channel, with (a) $D_s = T$, (b) $D_s = 2T$ and (c) $D_s = 5T$, versus average E_b/N_o.

7.4 The Dithered THP as a Peak to Average Power Ratio Reduction

To study the effect of the modulo operation as well as the dithering signals, figure 7.4 shows the average transmitted power for the dithered and the non-dithered THP. As it is expected, the use of dither signal forces the transmitted signal to be always uniformly distributed over the modulo set —see figure 5.4 in 5.2.2—. The result is that the dithered THP shows roughly constant transmitted power and it is able to control the peak power transmissions to be approximate the same as in average, while the

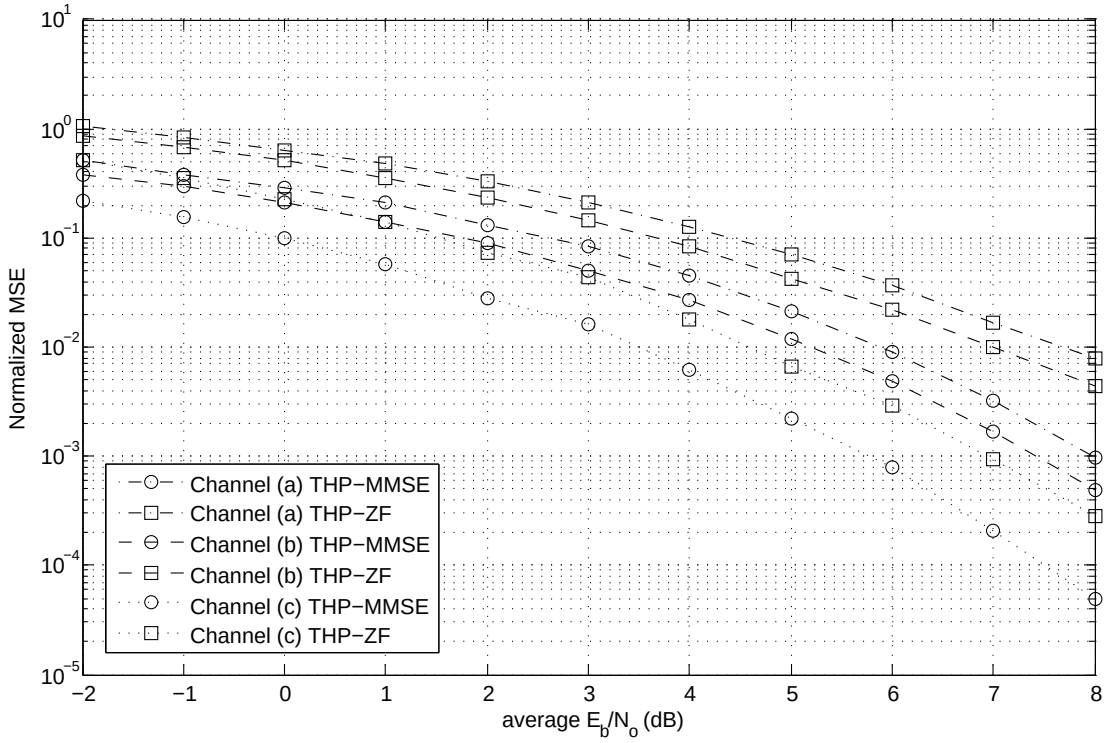


Figure 7.3: NMSE for the THP SISO channel, with (a) $D_s = T$, (b) $D_s = 2T$ and (c) $D_s = 5T$, versus average E_b/N_o.

non-dithered version would need different transmitted powers depending on the channel realization and the former transmitted symbols given a certain state. This concept has recently been applied by [LLLS07] in multi carrier systems as an improvement of [CC04] for DSL lines.

In account of the fact that if no precoding were used, the QPSK transmission would require only a mean energy per channel use lower that if the precoding is used, the impact of the dithering is an increased transmitted power, i.e. as power loss, as introduced in 7.2. However, the dithering plays a double role in the system: a part from controlling the transmitted power, what is more important, it fixes the statistics of the input of the channel. Whether these statistics are not known, the optimization problem could not be performed in such accurate way. In this case, the transmitter and the receiver should monitor the statistics and constantly adjust the signal processing filters, which might not be feasible. The reason is that, not only the channel is variant, but also the transmitted signals are dependent on the user information sequence.

7.5 Performance of Factorization Techniques

A simulation on the SISO channel has been done to get results on the SER and the MSE for the ZF at transmitter, ZF with QR, MMSE-ZF, MMSE and MMSE with Cholesky techniques. The QR and the Cholesky factorization are quite better compared to their respective solutions. Please refer to figures 7.5 and 7.6 for a complete plot of SER and NMSE versus the average SNR. At high SNR, the gain in performance with respect to the solutions without factorization is more relevant rather than at low SNR. The reason is that the factorization applied to the channel is intelligent regarding the signal processing structure. The optimal solutions —MMSE and MMSE-ZF— are considered to be optimal for any channel matrix that, in general, does not allow any type of factorization. Therefore, the

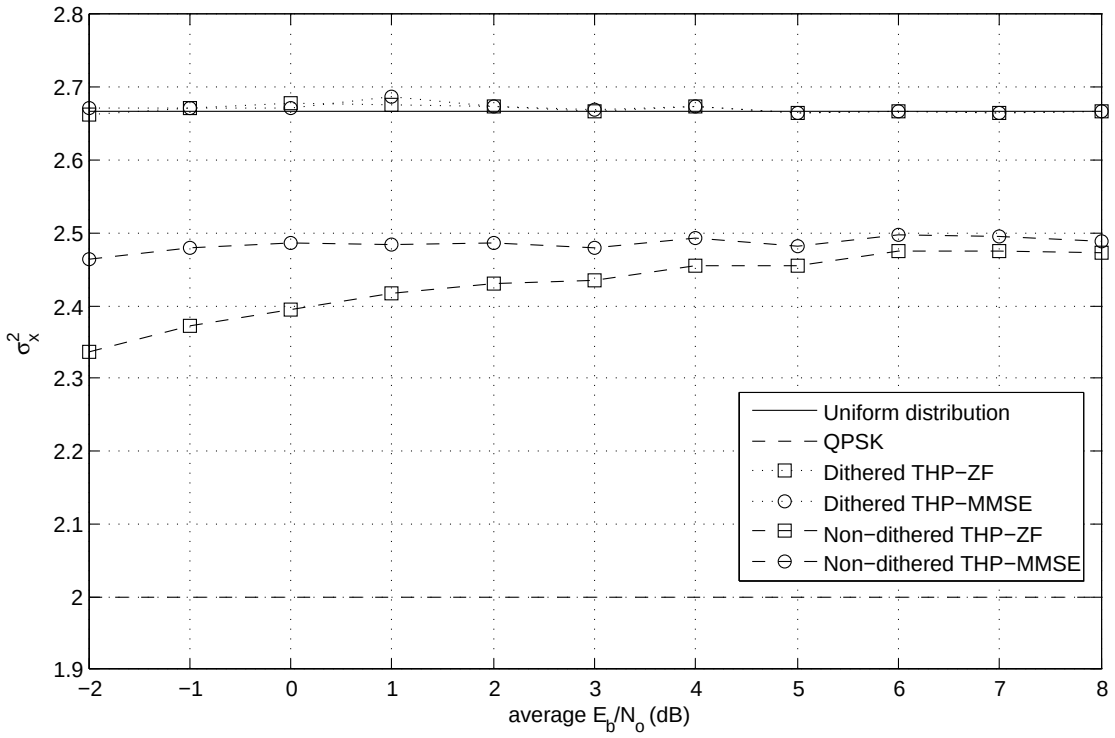


Figure 7.4: Average transmitted signal power for the dithered and non-dithered THP.

restriction on the feedback matrix is required, which removes a degree of freedom in the formulation of the problem. However, when a lower triangular factorization is done, this particular restriction is intrinsically present in the factorization and, as a result, the system can perform better.

Despite of this, the behavior of the four solutions is very similar, so the Cholesky factorization and the QR decomposition filters are expected to follow the same consequences when altering the temporal diversity of the channel. The interest in using this kind of factorization is fostered by complexity considerations, which are omitted in this thesis for the simple fact that the most of the signal processing –that is, the feedback loop- is placed in the transmitted side.

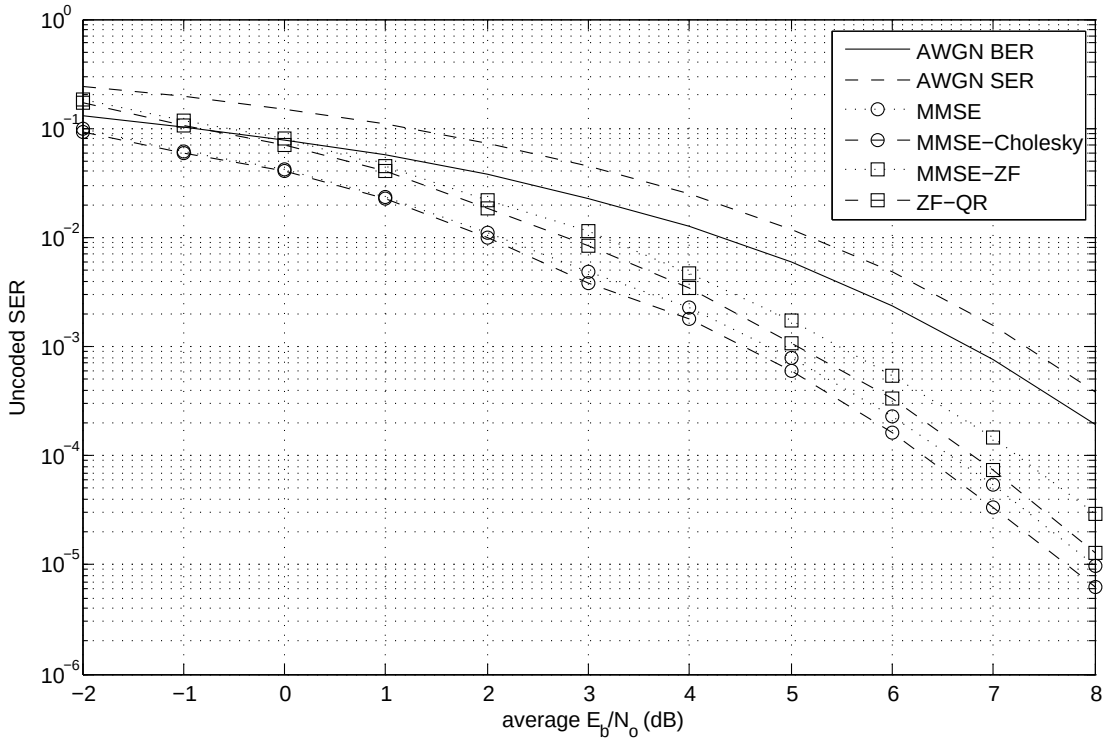


Figure 7.5: SER for the THP SISO channel, with $D_s = 5T$, versus average E_b/N_o.

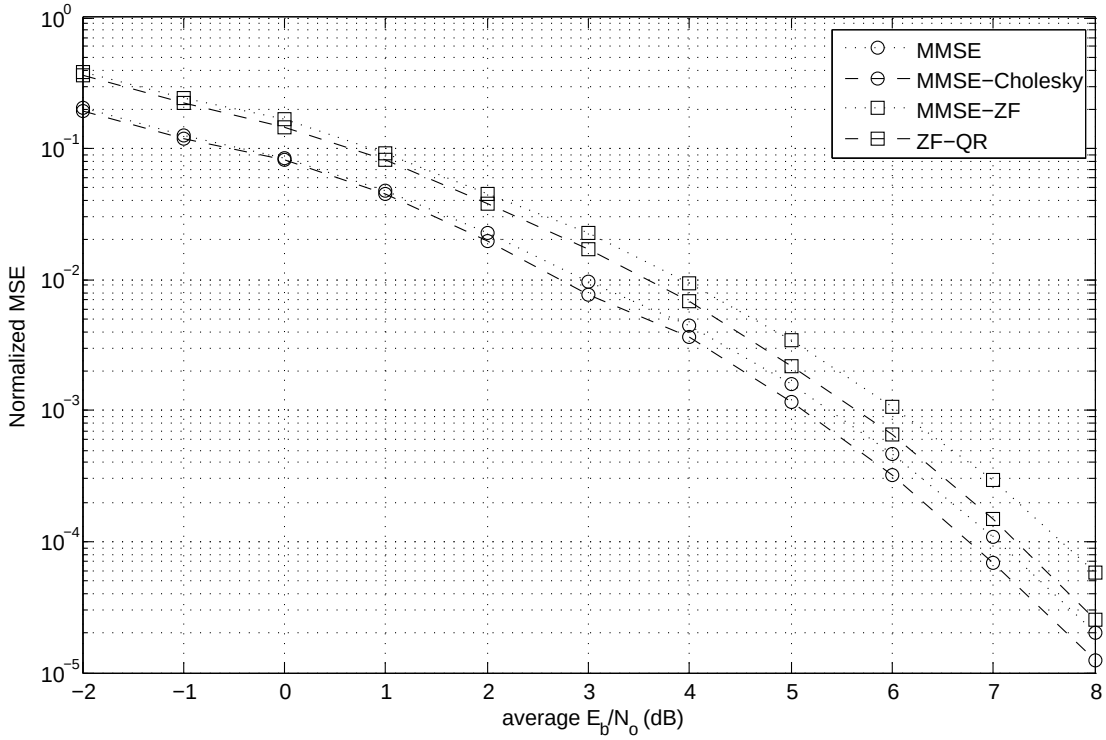


Figure 7.6: NMSE for the THP SISO channel, with $D_s = 5T$, versus average E_b/N_o.

Part III

Spatial THP for the Single-User MIMO Frequency-Selective Channel

Chapter 8

Introduction to MIMO and OFDM

8.1 Motivation

On the one hand, multiple-input multiple-output (MIMO) systems have been regarded as one of the most promising research areas of wireless communications. The increase of higher degrees of diversity fosters the use of MIMO systems in order to enhance the system capacity and the spectral efficiency with respect to single-input single-output (SISO) systems. The diversity attained is enabled by the fact that, in a rich environment, the signal from each individual sub channel —we refer to sub channel or branch as the channel between any given input and output— appears highly uncorrelated at each of the channel outputs. When this happens, the receiver obtains different signatures or channel effects and it can use these differences to simultaneously and at the same frequency separate them or, in a constructive way, recover them. Goldsmith [GJJV03] has shown very presumable limits in terms of channel capacity for several communications problems easily modeled by a MIMO system. Though there are several systems than can be analyzed as MIMO, usually the system is equipped with several transmitting antennas and receiving antennas, describing a MIMO system through spatial diversity.

On the other hand, in view of the high complexity when the order of the channel increases, in order to attain scalability problems, a orthogonal frequency division multiplexing (OFDM) will be considered for the physical layer due to its numerous advantages. Firstly, it can easily adapt to severe temporal channel conditions and it is robust in front of inter-symbol interference (ISI) and fading caused by the multi-path propagation, allowing the system to be treated as a set of flat-fading MIMO channels. Secondly, it has high spectral efficiency and an efficient implementation using the fast Fourier transform (FFT). Finally, since in terms of spectral efficiency a MIMO system should be designed to approach the capacity of the channel, the use of a multi-carrier approach on frequency-selective MIMO channels has been shown (cf. [RC98]) to be a capacity-lossless structure.

8.2 Objectives Outline

Hence, the scope of this part is to design a joint transmit-receive optimization based on Tomlinson-Harashima precoding (THP) for the single-user MIMO-OFDM channel. With the aim of presenting the capable performance range of THP in the MIMO-OFDM channel, a first overview on the system is given in the following lines. The THP technique has been tested in a single-user MIMO-OFDM system equipped with two transmitting and two receiving antennas and 8 orthogonal carriers, so that two QPSK symbols are spatially multiplexed —that is, each antenna transmits one symbol per channel

use— and each MIMO symbol is copied onto the 8 carriers, which makes the OFDM modulation acting as a diversity technique. This is done to show the system's performance in a real scenario: the uncoded symbol error rate (SER) does not reflect the error rate at application level, since a correcting code might be used in the middle ware. In this work, instead, the OFDM is used as a diversity technique so the probability of detecting an error after combining all the carrier under a RAKE [Cio02] receiver diminishes considerably, as it happens with a code.

The optimal filters have been found under minimum mean-square error (MMSE), zero-forcing (ZF), Cholesky and QR decomposition techniques, analogous to the SISO case. As a result, the optimal feedback and receive filters have very similar expressions —please refer to chapters 10 and 11 for complete expressions— with a block diagonal structure due to the OFDM modulation. However, with the introduction of MIMO, as motivated in the next sections, a precoding filter is required after the Tomlinson-Harashima feedback loop and before facing the channel. This filter is in charge of power allocating the signal along the transmitting antennas and OFDM carriers. This new adding deserves attention to be cited here. It is shown that the optimal precoding filter in the MMSE sense accomplishes, at the n-th carrier:

$$\boldsymbol{F}_n = \boldsymbol{\mathcal{U}}_n \boldsymbol{\Sigma}_n,$$

where $\boldsymbol{\mathcal{U}}_n$ has as columns the eigenvectors of $\boldsymbol{R}_{\boldsymbol{H}_n}$ and $\boldsymbol{\Sigma}_n$ is a non-squared matrix with one diagonal whose squared i-th element is:

$$\sigma_{n,i}^2 = \left(\frac{\mathcal{E}/\sigma_v^2 + \sum_{\forall j} \sum_{\forall m} \lambda_{m,j}^{-1}}{\sum_{\forall j} \sum_{\forall m} \lambda_{m,j}^{-1/2}} \lambda_{n,i}^{-1/2} - \frac{1}{\lambda_{n,i}} \right)^+,$$

for the MMSE cases, and

$$\sigma_{n,i}^2 = \frac{\mathcal{E}/\sigma_v^2}{\sum_{\forall m} \sum_{\forall j} \lambda_{m,j}^{-1/2}} \lambda_{n,i}^{-1/2}$$

for the ZF cases. In these expressions, $\mathcal{E}$ is the total available transmit energy per channel use and $\lambda_{m,j}$ is the j-th eigenvalue corresponding to the eigenvalue decomposition (EVD) of the channel correlation matrix $\boldsymbol{R}_{\boldsymbol{H}_n}$ at the n-th carrier. $\boldsymbol{R}_{\boldsymbol{H}_n}$ is a matrix that reflects the conditions of the channel in terms of channel gain/attenuation and noise. As a consequence, $\boldsymbol{\mathcal{U}}_n$ gives direction to the information to optimally face the channel and $\boldsymbol{\Sigma}_n$ gives the corresponding weights in which the information is placed. Though this is a known result, it has been never applied to a complete OFDM-MIMO THP system.

Before entering into detail, figure 8.1 depicts the SER of the system described above for the four studied cases. As it can be appreciated, the MMSE and the Cholesky factorization perform equally, as the MMSE-ZF and the QR decomposition do. The most relevant result is that the MMSE solutions perform much better than the ZF solutions, and that the difference in performance is notably bigger compared to the SISO case. This is due to the distinct power allocation policy that the MMSE and ZF follow, i.e. $\sigma_{n,1}^2$, because the ZF power allocation waist the resources to achieve the ZF condition, whereas the MMSE power allocation distributes the power in a more intelligent manner.

8.3 Previous Work

When referring to linear equalization (LE) for the MIMO channel, Scaglione and Barbarossa [BS00] first address the optimal receive and transmit filter when complete CSI is available at both sides. Further in [SSB+02], they derive closed-forms for the optimal terminal filters by scalarizing —that is, the matrix equations are substituted by scalar ones— the problem exploiting the channel EVD and

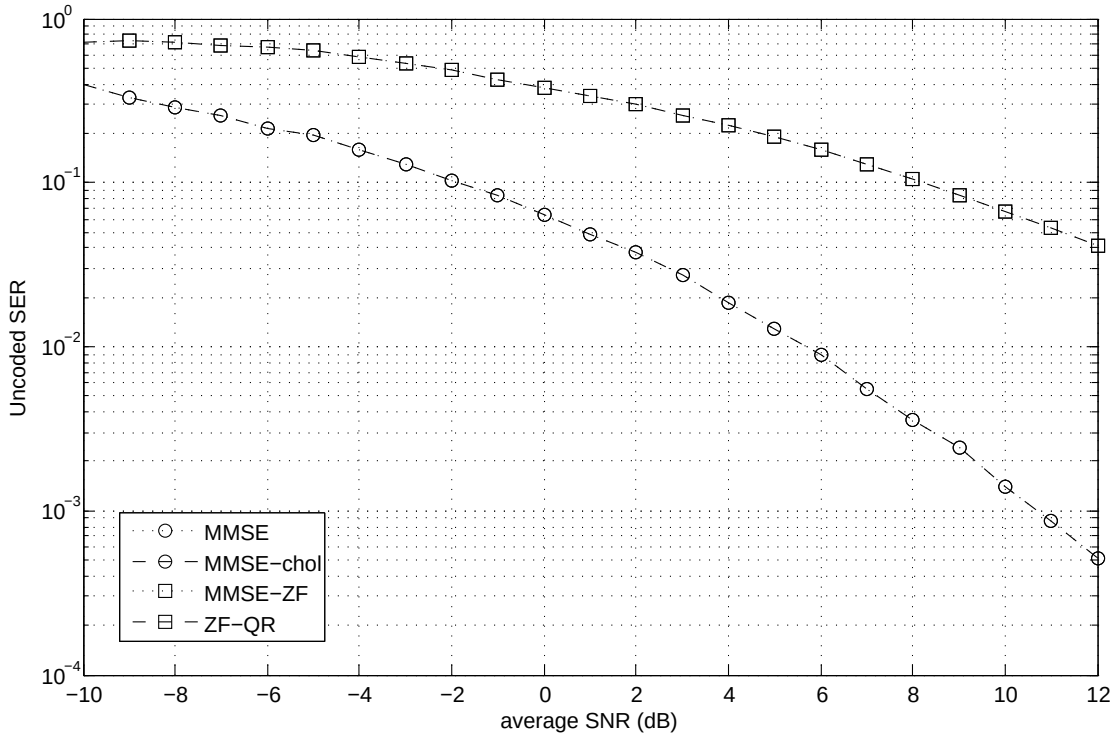


Figure 8.1: SER for the $N = 8$ carriers THP 2×2 MIMO channel with $\rho_R = \rho_T = 0.2$ correlation parameters and uniform noise versus average SNR.

minimizing the overall mean-square error (MSE) under several constraints.

Cioffi and Lagunas in [PCL03] derived a unified framework for LE in multi-carrier systems, like OFDM, when the parameter to be optimized is a certain objective or cost function of the MSE. The leitmotiv behind their work is the difficulty in determining the optimal filters in a non-convex problem (e.g. [BV04]), which is usually the case when matrix equalization is considered. Interestingly, one may find several optimization criteria based on minimizing the MSE, the SER, the package error rate (PER) or maximizing the channel capacity or the cut-off rate, which provides a lower bound on the channel capacity.

In this line, a general framework is presented by [SS04] for both linear and non-linear precoding and equalization for the flat-fading MIMO channel. They design the optimal receive and transmitting filters under the MMSE criterion, while basing the power allocation on the results in [SSB$^+$02]. More recently, —and noting that THP is being still under research— the THP is designed for a MIMO system with the use of the Cholesky factorization (6.25) in [SD08] also under a general framework for different objective functions to be optimized.

Here, we assume complete CSI at both transmitter and receiver sides and the MSE is used as a target function. The power allocation is jointly optimized over the whole bandwidth and over the transmitting antennas; whereas for simplicity the bit loading and the coding scheme are assumed to be the same in all the branches, i.e. equal-rate transmission as employed in standards like the European HIPERLAN/2 and the U.S. IEEE 802.11 for WLANs.

Chapter 9

The MIMO-OFDM Frequency-Selective Channel

9.1 MIMO-OFDM Channels

The MIMO channels arise many different scenarios such as when a bundle of twisted pairs in digital subscriber lines (DSLs) is treated as a whole, when multiple antennas are used at both sides of a wireless link, or simply when a frequency-selective channel is properly modeled by using, for instance, filter banks.

The system considered has T inputs and R outputs, which, in spatial diversity, would correspond to T transmit and R receive antennas. The frequency-selective channel equation is given by (e.g. [SSB+02]):

$$\boldsymbol{y}(m) = \sum_{\ell} \boldsymbol{H}(m, m-\ell)\boldsymbol{x}(\ell) + \boldsymbol{n}(m), \tag{9.1}$$

where m denotes the discrete-time temporal index, $\boldsymbol{x} \in \mathbb{C}^T$ is the baseband vector signal wrapping the T transmit signals $(x_i, 1 \leq i \leq T)$, $\boldsymbol{y} \in \mathbb{C}^R$ is the vector wrapping the R received signals from the channel $(y_j, 1 \leq j \leq R)$, $\boldsymbol{n} \in \mathbb{C}^R$ is the vector containing the R noise signals at each received branch and $\boldsymbol{H} \in \mathbb{C}^{R\times T}$ is the frequency-selective MIMO channel matrix whose (i,j)-th entry contains the channel impulse response of the path between input i and output j, $h_{ij}(m, m-\ell)$, at time m. If the channel discrete-time time-varying impulse response $\boldsymbol{H}$ is causal and has finite memory K, the former expression (9.1) can be written in the form of finite-length vectors by stacking M consecutive transmitted signals, $L = N + K$ consecutive received signals and noise signals and defining $\mathcal{H}$ as an $RL \times TN$ block banded matrix, similarly constructed as in (5.3). That is, $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$. If the channel is also time-invariant during the transmission of the N block, i.e. $\boldsymbol{H}(m, m-\ell) = \boldsymbol{H}(\ell)$, $\mathcal{H}$ becomes a block Toeplitz matrix:

$$\mathcal{H} = \begin{bmatrix} \boldsymbol{H}(0) & 0 & \dots & 0 \\ \vdots & \boldsymbol{H}(0) & & \vdots \\ \boldsymbol{H}(K) & \vdots & \ddots & 0 \\ 0 & \boldsymbol{H}(K) & & \boldsymbol{H}(0) \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \boldsymbol{H}(K) \end{bmatrix}.$$

Comparing with respect to the signal processing applied to SISO channel in part II, it is easily ap-

preciated that now the order of the signal processing filters increases proportionally to the dimensions of the MIMO channel. For this reason, the frequency-selective nature of the channel is dealt with by taking a multi-carrier OFDM approach, without loss of generality (e.g. [RC98]).

The main characteristic of OFDM is that the whole transmission bandwidth is divided into parallel streams, over which a different consecutive symbol is transmitted. When the length of the cycle prefix in the OFDM modulation is higher or equal than the channel memory, K, each of the streams is only affected by a multiplicative gain. As a result, and defining the OFDM block length as N —that is, the bandwidth is divided into N sub carriers—, the coupled problem is converted into N separate $R \times T$ individual flat-fading MIMO channels.

Specifically, let $\boldsymbol{x}_n \in \mathbb{C}^T, 0 \leq n < N$ be the transmitted MIMO signal placed on the n-th OFDM band. The modulation consists of applying an unitary inverse FFT and the insertion of a cycle prefix of P samples, so the MIMO-OFDM symbol has $N+P$ samples. Let $\boldsymbol{z}$ be the global MIMO transmitted signal:

$$\boldsymbol{z}(m) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \boldsymbol{x}_n e^{\jmath \frac{2\pi}{N} nm}, \tag{9.2}$$

where m is the sample index, $-P \geq m < N$. The received MIMO signal, $\boldsymbol{y}$, is the convolution of the OFDM symbol, $\boldsymbol{z}$ with the MIMO channel, $\boldsymbol{H}$, plus the $R \times 1$ additive noise. Notice that the power of each of the R receiving noises, $n_j(m)$ $1 \leq j \leq R$, will show different values depending on the receive conditions at each of the channel outputs. From (9.1), we can write the channel impulse response as $\boldsymbol{H}(D) = \sum_{\ell=0}^{K} \boldsymbol{H}(\ell) D^\ell$, and by $D = e^{\jmath \frac{2\pi}{N} n}$ we obtain the frequency response at the n-th carrier, that is

$$\boldsymbol{H}_n = \sum_{\ell=0}^{K} \boldsymbol{H}(\ell) e^{\jmath \frac{2\pi}{N} \ell n}. \tag{9.3}$$

The cycle prefix is extracted to the received signal $\boldsymbol{r} \in \mathbb{C}^R$ and the unitary FFT is calculated. When $M \geq K$ ([Pro00]), the received signal at the n-th sub carrier becomes

$$\boldsymbol{y}_n = \boldsymbol{H}_n \boldsymbol{x}_n + \boldsymbol{n}_n. \tag{9.4}$$

Notice that the noise statistics are not altered since the OFDM demodulation is a unitary transformation. Therefore, the frequency-selective MIMO channel can be decoupled into N MIMO flat-fading channels given by (9.4), for $1 \leq n \leq N$. As a result, the Tomlinson-Harashima system is split into N sub-systems that process each sub carrier independently. In other words, THP acts N times to diminish the inter-channel interference (ICI) inherent on each MIMO channel and places these N processed signals to the OFDM modulator. To illustrate the structure of the modulation and MIMO channel, figure 9.1 depicts the scheme for spatial ($R \times T$) MIMO diversity. The transmitter generates T signals stacked on the vector $\boldsymbol{x}(n) = [x_1(n) \dots x_T(n)]^T$, which is OFDM modulated according to (9.2) and placed over the T transmitting antennas. The receiver firstly performs the demodulation, leading to the equivalent channel written in (9.4). Finally, the overall system is now an $NR \times NT$ problem with highly decreased complexity.

9.2 Diversity Model

Yet several multi element systems follow the channel model described above (e.g. beam forming), in this work and exclusively for simulation purposes a spatial diversity model achieved with multiple antennas is considered. To do so, we make use of a separate correlation model between branches (e.g.

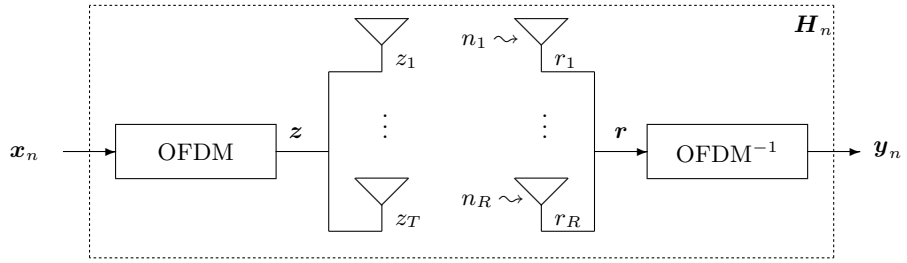


Figure 9.1: MIMO-OFDM channel

[FK00]) where the correlation between channels (i, j) and (k, l) is given by the product of the spatial correlation at the receiver, $r_R(i, j)$ and the spatial correlation at the transmitter $r_T(k, l)$, which are, by simplicity, assumed to be equal over the whole bandwidth. Hence, each matrix is assumed to be zero-mean circularly symmetric complex Gaussian distributed with a separable spatial correlation function so that

$$\boldsymbol{H}_n = \boldsymbol{R}_R^{H/2} \boldsymbol{H}_\omega \boldsymbol{R}_T^{1/2}, \tag{9.5}$$

where the square correlation matrices, $\boldsymbol{R}_R \in \mathbb{R}_+^{R\times R}$ and $\boldsymbol{R}_T \in \mathbb{R}_+^{T\times T}$ are defined as $[\boldsymbol{R}_R]_{ij} \triangleq r_R(i, j)$ and $[\boldsymbol{R}_T]_{kl} \triangleq r_T(k, l)$, and $\boldsymbol{H}_\omega$ is a matrix of i.i.d. circularly symmetric complex Gaussian variables whose variance represent the gain or attenuation at the n-th band. For simplicity, the numerical evaluation of the performance of the presented algorithms will be carried out for an AR(1) model of the spatial correlation:

$$\begin{aligned} r_R(i, j) &= \frac{1}{\sqrt{R}} \rho_R^{|i-j|} \\ r_T(k, l) &= \frac{1}{\sqrt{T}} \rho_T^{|k-l|}, \end{aligned}$$

with $0 \leq \rho_R, \rho_T \leq 1$ being the correlation coefficients at the receiver and transmitter respectively.

9.3 MIMO-OFDM Strategies

The use of OFDM on the MIMO channel turns out to be a good method of increasing the diversity order of the system. Notice that now, the system is able to place T symbols over N different bands. Yet, the diversity attained by means of the spatially correlation will be different than the one attained on the frequency domain.

The transmitter is in charge of preparing the N transmitting $T \times 1$ vectors, $\boldsymbol{x}_n$, from the user information. At each carrier, the matrix channel $\boldsymbol{H}_n$ has $\kappa_n = \text{rank}(\boldsymbol{H}_n)$ channel eigenmodes or sub channels (i.e. non-vanishing singular values) so the size of the information signal to be processed has to be M_n, with $M_n \leq \kappa_n = \min\{R, T\}$. This fixes the THP matrices $\{\boldsymbol{B}_n\}_n$ dimensions to $M_n \times M_n$, as well as the receive matrices $\{\boldsymbol{G}_n\}_n$ to $M_n \times R$. By comodity, we define M as the total number of symbols transmitted over all the bands, i.e., $M = \sum_{n=1}^{N} M_n$.

In the SISO channel, the user signal given to the precoder is directly itself, whereas the MIMO channel requires the definition of the strategy γ as

$$\begin{aligned} \gamma : \mathbb{C} &\longrightarrow \mathbb{C}^M \\ s(m) &\longmapsto \tilde{\boldsymbol{s}} = \gamma(s(m)), \end{aligned}$$

that is, how many information we want to transmit per channel use. We might distinguish the following extreme cases:

Maximum Throughput This strategy consists of placing different user information on the channel. This approach achieves the maximum channel use and makes use of the diversity by transmitting more information at the same time. In this particular scheme, a single channel use is able to place M information symbols over the whole bandwidth and MIMO channel. At a given sub band, the user vector signal would become $\tilde{\boldsymbol{s}}_n = [s(m) \dots s(m+M-1)]^T$.

Maximum Diversity The second strategy consists of the contrary: a unique information signal is copied onto the M_n symbols and onto the N carriers with the aim of exploiting the channel diversity. For instance, in this case γ would perform $\tilde{\boldsymbol{s}}_n = [\overbrace{s(m) \dots s(m)}^{R}]^T$, leading a throughput of at most M symbols per channel use.

Any other strategy would be bounded by this two cases and, in general, the channel use will be limited between 1 and M symbols per channel use. It is important that the system does not achieves such diversity and throughput without any penalty of resources. This is seen at two levels: space and frequency. On the one hand, to achieve spatial diversity or throughput, the system needs to put more energy compared to the SISO case or, in fairness of comparison, the same energy must be split onto the T transmitting antennas, having only a portion of power at each eigen mode. On the other hand, the OFDM modulation splits, at its turn, the whole bandwidth into N parallel channels which, obviously, see their bandwidth reduced by approximately a factor N. This is translated to the rate at which the user access the channel, since the OFDM symbol duration is increased. This last consideration is further analyzed in 12.4.

9.4 Power Allocation

With the use of OFDM, the MIMO system requires a power allocation in a double sense. Firstly, the power allocation over the whole bandwidth (i.e. the N carriers) can be done from an information theory point of view (cf. to [Sha48]) through the classical water-filling method (e.g. [Gal68]) by maximizing the channel capacity, or through other methods more related to the SER (e.g. [TdPE02]) or the MSE. Secondly, the design of the power allocation over the T transmitting antennas can be done under several criteria; commonly speaking about the minimization of the MSE, the minimization of the SER (e.g. [BF07]) or the maximization of the channel capacity (e.g. [GJJV03]).

Beamforming is a signal processing technique that controls the power allocated on each channel input. When receiving a signal, beamforming can increase the receiver sensitivity in the direction of wanted signals and decrease the sensitivity in the direction of interference and noise. In this work, this is done at the same level for both space and frequency through the forward matrix at the transmitter, $\boldsymbol{F}_n \in \mathbb{C}^{T \times M_n}$. Specifically, it is assumed that the THP system is given a certain amount of energy for the whole bandwidth, $\mathcal{E}$, to be used in the N MIMO channels described in (9.4). Let $\boldsymbol{v}_n$ be the output signal of the THP generated using the user signal $\tilde{\boldsymbol{s}}_n$ and the corresponding modulo operation and feedback filter $\boldsymbol{B}_n$. With the use of dithering, all the M signals will show average power of $\tau^2/6$, being τ the size of the modulo set (cf. 5.2.2). The total transmitted energy by channel use is given by means of the introduced forward transmit filter since $\boldsymbol{x}_n = \boldsymbol{F}_n \boldsymbol{v}_n$, $||\boldsymbol{x}_n||^2 = \mathrm{tr}(\boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H)$, where $\boldsymbol{R}_{\boldsymbol{v}_n}$ represents the correlation matrix of the signal $\boldsymbol{v}_n$, which becomes to be formed by i.i.d. white variables. In the SISO channel a power restriction was not required in the optimization process

since the modulo operation was strictly in charge of limiting the transmitted power to a given value. However, in the MIMO channel the following restriction must be considered in all the optimization processes. The formulation of this constraint for the MIMO-OFDM reads

$$\sum_n \mathrm{tr}(\boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H) \leq \mathcal{E}. \tag{9.6}$$

The forward filter will be therefore jointly optimized with the rest of the processing system, which is described in the next section.

9.5 Signal Processing

The signal model is analogous to the SISO case but now the processing is done separately on the N carriers and with order M_n. The full system proposed in this chapter is detailed in figures 9.2 and 9.3, which depict the transmitter and the receiver respectively.

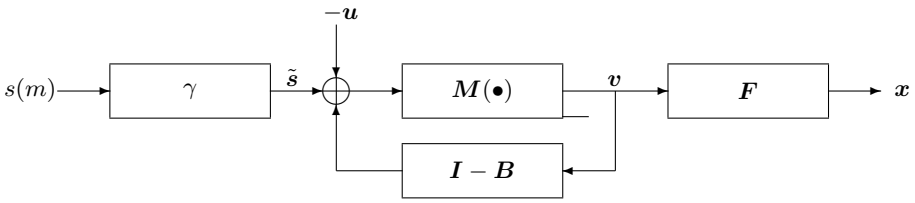


Figure 9.2: MIMO-OFDM Tomlinson-Harashima precoder

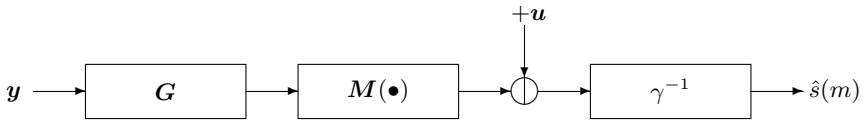


Figure 9.3: MIMO-OFDM Tomlinson-Harashima receiver

The modulo operation can be treated linearly using the auxiliary signal at each carrier $\boldsymbol{a} = [\boldsymbol{a}_1^T \dots \boldsymbol{a}_N^T]^T \in \mathbb{C}^M$ and the proper dithering signal, $\boldsymbol{u} = [\boldsymbol{u}_1^T \dots \boldsymbol{u}_N^T]^T \in \mathbb{C}^M$, will be used at both sides as justified in 5.2.3. Therefore, defining $\boldsymbol{d} \triangleq \tilde{\boldsymbol{s}} - \boldsymbol{a} - \boldsymbol{u}$, the overall equation reads

$$\hat{\boldsymbol{d}} = \boldsymbol{G}\mathcal{H}\boldsymbol{F}\boldsymbol{B}^{-1}\boldsymbol{d} + \boldsymbol{G}\boldsymbol{n}, \tag{9.7}$$

where now, $\boldsymbol{d} \in \mathbb{C}^M = [\boldsymbol{d}_1^T \dots \boldsymbol{d}_N^T]^T$, $\boldsymbol{n} \in \mathbb{C}^{NR} = [\boldsymbol{n}_1^T \dots \boldsymbol{n}_N^T]^T$ and, with the use of OFDM, the signal processing filters can be structured as:

$$\begin{aligned} \boldsymbol{B} &= \mathrm{diag}(\boldsymbol{B}_1 \dots \boldsymbol{B}_N) \in \mathbb{C}^{M \times M} \\ \boldsymbol{F} &= \mathrm{diag}(\boldsymbol{F}_1 \dots \boldsymbol{F}_N) \in \mathbb{C}^{NT \times M} \\ \boldsymbol{G} &= \mathrm{diag}(\boldsymbol{G}_1 \dots \boldsymbol{G}_N) \in \mathbb{C}^{M \times NR}, \end{aligned}$$

while the channel matrix is now $\mathcal{H} = \mathrm{diag}(\boldsymbol{H}_1 \dots \boldsymbol{H}_N)$. Following this result, the input-output relation

in (9.7) is rewritten as a set of equations for each carrier. That is,

$$\hat{\boldsymbol{d}}_n = \boldsymbol{G}_n \boldsymbol{H}_n \boldsymbol{F}_n \boldsymbol{B}_n^{-1} \boldsymbol{d}_n + \boldsymbol{G}_n \boldsymbol{n}_n \tag{9.8}$$

Of course, this carrier-uncooperative model assumes that the carriers are completely orthogonal. In this particular scheme, we will always transmit M_n symbols at each carrier given by the rank of the channel response at that frequency; whereas a cooperative approach would consider the whole budget M to reallocate the symbols among the carriers in an intelligent way, having sense that this general model has a potential better performance. However, from a mathematical point of view, the non-cooperative scheme is more general since the carrier-cooperative model is obtained by particularizing $N = 1$, as if only one frequency was used. Thus, in the sequel, the carrier non-cooperative matrix signal model is considered, without loss of generality.

Chapter 10

Optimal MMSE Filters

The joint transmit-receive matrix design is in general a complicated non-convex problem. However, as previously mentioned, for some specific design criterion, the original complicated problem is greatly simplified because the channel turns out to be diagonalized, which allows the scalarization of the problem. Examples of this are the minimization of the sum of MSEs of all spatial sub channels, the minimization of the determinant of the MSE matrix and the maximization of the mutual information. Recall that (cf. [PCL03]) the diagonalizing structure is not optimal for other interesting design criterion such as the minimum SER. In this chapter, we require of the majorization theory (cf. [MO79] as a complete reference on the topic) and the so-called Schur-convex and Schur-concave functions. However, though requiring of new theories and new matrix precoding filters, the formulation and mathematical results in part II referring to the use of dithering, the linearization of the model, and the orthogonality principle turn out to be analogous to some results in the current part.

Without loss of generality, the study is focused on the arithmetic mean of individual MSE at each carrier in view of the fact that the optimal filters that minimize all the MSEs, maximize all the signal-to-interference plus noise ratio (SINR) and minimize each of the SERs. In other words, the SINR can be expressed as a function of the MSE by

$$\mathrm{MSE}_n = \frac{1}{\mathrm{SINR}_n} - 1,$$

and the BER can be expressed as a function of the SINR by

$$\mathrm{BER}_n = \alpha Q\left(\beta\sqrt{\mathrm{SINR}_n}\right),$$

where parameters alpha and beta depends on the signalling.

This chapter aims to design the THP filters that minimizes the non-averaged arithmetic mean of the MSE at the output of the linear receiver filter subject to an average power constraint across the carriers and antennas. That is, keeping in mind the structures of the processing filters,

$$\begin{aligned} \{\boldsymbol{B}, \boldsymbol{F}, \boldsymbol{G}\} &= \arg\min \sum_{n=1}^{N} E[|\boldsymbol{d}_n - \hat{\boldsymbol{d}}_n|^2] \\ \text{s. t.} &\quad \sum_{n=1}^{N} \mathrm{tr}(\boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H) \leq \mathcal{E}. \end{aligned}$$

The resolution of this problem is done in a three steps process: first the optimal receive filter is

determined, secondly the feedback filter is calculated and finally the precoding filter is found as the optimal power allocation. Although the formulation of the THP for the MIMO-OFDM channel is not found in the literature, the following results are based on methodologies explained in excellent papers that join transmit-receive optimize MIMO space-time precoders and decoders [SSB+02], multi-carrier MIMO decoders and precoders [PCL03], and non-linear flat-fading MIMO decoders and precoders by [SS04] and [SD08]. As a matter of notation, we define the individual MSEs as $\varepsilon_n^2 = E[|\boldsymbol{d}_n - \hat{\boldsymbol{d}}_n|^2$, the individual error matrix as $\boldsymbol{E}_n = (\boldsymbol{d}_n - \hat{\boldsymbol{d}}_n)(\boldsymbol{d}_n - \hat{\boldsymbol{d}}_n)^H$ for $1 \leq n \leq N$, and the overall MSE and error matrix as $\varepsilon^2 = \sum_n \varepsilon_n^2$ and $\boldsymbol{E} = \text{diag}(\{\boldsymbol{E}_n\}_n)$ respectively. Notice that both are related by the trace operator as $\varepsilon^2 = \text{tr}(\boldsymbol{E})$. In general, the error matrices will have different dimensions (they are $M_n \times M_n$ square matrices), while the overall matrix is $M \times M$ matrix. However, since M_n is given by the rank of the channel frequency response at the n-th carrier, with high probability the uncorrelation of the spatial streams will set this number to the minimum of R and T, having the same transmitted OFDM symbols at each carrier and, thus, $M = NM_n$.

10.1 Optimal Receive Matrix

The optimal receive matrix $\boldsymbol{G}$ is determined by the set of optimal receive matrices $\{\boldsymbol{G}_n\}_n$ by minimizing the individual MSE at each carrier. This is done since the problem is decoupled into N orthogonal individual flat-fading MIMO channels —i.e. $\boldsymbol{G}_n$ does not process at any other band different than the n-th one— and there is no restriction on the structure of $\boldsymbol{G}$. In the sense of minimizing the MSE, the approach is analogous to section 6.1 for the SISO channel. The n-th MSE is written using the n-th error matrix, the expression (9.8) and that $\boldsymbol{d}_n = \boldsymbol{B}_n \boldsymbol{v}_n$:

$$\varepsilon_n^2(\boldsymbol{G}_n) = \text{tr}(\boldsymbol{E}_n) = \text{tr}(\boldsymbol{B}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{B}_n^H - \boldsymbol{B}_n \boldsymbol{R}_{\boldsymbol{v}_n \boldsymbol{y}_n} \boldsymbol{G}_n^H - \boldsymbol{G}_n \boldsymbol{R}_{\boldsymbol{y}_n \boldsymbol{v}_n} \boldsymbol{B}_n^H + \boldsymbol{G}_n \boldsymbol{R}_{\boldsymbol{y}_n} \boldsymbol{G}_n^H).$$

When setting $\nabla_{\boldsymbol{G}_n^H}^T \varepsilon_n^2 = 0$ the receive filter obeys the Wiener filter expression $\boldsymbol{G}_n = \boldsymbol{B}_n \boldsymbol{R}_{\boldsymbol{v}_n \boldsymbol{y}_n} \boldsymbol{R}_{\boldsymbol{y}_n}^{-1}$, which, after mathematical manipulations, is written as

$$\boldsymbol{G}_n = \boldsymbol{B}_n (\boldsymbol{R}_{\boldsymbol{v}_n}^{-1} + \boldsymbol{F}_n^H \boldsymbol{H}_n^H \boldsymbol{R}_{\boldsymbol{n}_n}^{-1} \boldsymbol{H}_n \boldsymbol{F}_n)^{-1} \boldsymbol{F}_n^H \boldsymbol{H}_n^H \boldsymbol{R}_{\boldsymbol{n}_n}^{-1}. \tag{10.1}$$

This expression is analogous to (6.22) for the SISO case, if we define the design matrix at the n-th carrier as

$$\boldsymbol{\Phi}_n \triangleq (\boldsymbol{R}_{\boldsymbol{v}_n}^{-1} + \boldsymbol{F}_n^H \boldsymbol{H}_n^H \boldsymbol{R}_{\boldsymbol{n}_n}^{-1} \boldsymbol{H}_n \boldsymbol{F}_n)^{-1}. \tag{10.2}$$

With this expression of the optimal receive filter and as a result of the orthogonality principle, the individual MSE becomes

$$\varepsilon_n^2(\boldsymbol{B}_n) = \text{tr}(\boldsymbol{B}_n \boldsymbol{\Phi}_n \boldsymbol{B}_n^H), \tag{10.3}$$

which is exclusively in terms of the feedback filter.

10.2 Optimal Feedback Matrix

Equivalently, the optimal matrix $\boldsymbol{B}$ is given by the set of optimal matrices at each carrier. In this case, the individual feedback matrices must accomplish the restriction (6.18) of being lower triangular with one main diagonal to ensure the stability of the feedback system with the modulo operation. In

view of the new expression of the MSE in (10.3), the Lagrangian at the n-th carrier becomes

$$\mathcal{L}_n(\boldsymbol{B}_n) = \mathrm{tr}(\boldsymbol{B}_n\boldsymbol{\Phi}_n\boldsymbol{B}_n^H) - 2\mathrm{Re}\ \mathrm{tr}\boldsymbol{\Lambda}_n(\boldsymbol{B}_n^H \odot \boldsymbol{U}_n^H - \boldsymbol{I}_{M_n}),$$

where $\odot$ stands for the Hadamard product and $\boldsymbol{U}_n$ is an $M_n \times M_n$ upper triangular matrix with all ones. The result of this problem is known from part II and reads

$$\boldsymbol{B}_n = \mathrm{diag}^{-1}(\boldsymbol{\Phi}_n^{-1})(\boldsymbol{U}_n^T \odot \boldsymbol{\Phi}_n^{-1}). \tag{10.4}$$

Up to this point, we may state a very important result: **the error matrix has been diagonalized**. Specifically, given the structure of the feedback filter, when a matrix is left-multiplied by a lower triangular matrix and right-multiplied by an upper-triangular matrix it becomes diagonal. Since the overall error matrix is the block diagonal matrix of all the error matrices, it is then diagonal. This means that the precoding forward filter $\boldsymbol{F}$ is in charge of placing the $M = \sum_n M_n$ symbols over the NT possible bins corresponding to T transmitting antennas at N different carriers.

10.3 Preliminaries for the Optimal Power Allocation

The optimal precoding or power allocation matrix $\boldsymbol{F}$ cannot be determined by minimizing the individual MSE at each carrier. Though the signal processing is carrier-noncooperative —i.e., the precoding matrix is structured as a block diagonal matrix as well—, it is in charge of allocate the power to the whole resources: antennas and carriers. In other words, we might determine the set $\{\boldsymbol{F}_n\}_n$ altogether.

As stated above, this matrix is designed by minimizing the arithmetic (or sum) of the individual MSEs under the power constraint (9.6). Since the error matrix has been diagonalized, the problem is scalarized. Yet, here the minimization of this scalar function is regarded under the convex optimization and majorization theories, which are briefly introduced in the sequel.

10.3.1 Convex Optimization Problems

A general convex optimization problem (e.g. [BV04]) is of the form

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\ \text{s. t.} \quad & f_i(\boldsymbol{x}) \leq 0, \quad 1 \leq i \leq P \\ & h_i(\boldsymbol{x}) \leq 0, \quad 1 \leq i \leq Q, \end{aligned}$$

where $\boldsymbol{x} \in \mathbb{R}^N$ is the optimization variable, the set of f functions are convex functions and the set of h are affine functions. The function f_0 is the objective function or cost function, while the other f functions state the inequality constraints and h the equality constraints. When all the functions involved are affine, the problem is called linear programming and is much simpler to solve. The importance of these types of problems has gained much attention in the last years and they arise from the fact that many engineering problems can be translated into the form of convex optimization and that they can be numerically solved very efficiently. In some cases, the problem can be solved analytically using the Karush-Kuhn-Tucker (KKT) optimality conditions but, in most cases, it has to be solved iteratively.

10.3.2 Majorization Theory: Schur-convex and Schur-concave Functions

This theory (see [MO79]) makes precise the vague notion that the components of a certain vector $\boldsymbol{x}$ are less spread out or more nearly equal that the components of other vector, $\boldsymbol{y}$.

Definition 1 (Additive Majorization)**.** *For a vector $\boldsymbol{x} \in \mathbb{R}^N$, let $x_{[1]}, \ldots, x_{[N]}$ denote the ordering of the components of the vector in a non-increasing order, i.e. $x_{[1]} \geq \ldots \geq x_{[N]}$. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$. The vector $\boldsymbol{y}$ majorizes $\boldsymbol{x}$ (or the second is majorized by the first) if*

$$\sum_{i=1}^{j} x_{[i]} \leq \sum_{i=1}^{j} y_{[i]} \qquad \text{for } 1 \leq j \leq N-1$$
$$\sum_{i=1}^{N} x_{[i]} = \sum_{i=1}^{N} y_{[i]}$$

and it is represented by $\boldsymbol{x} \prec \boldsymbol{y}$.

Definition 2 (Schur-convexity)**.** *A real-valued function f defined on a set $\mathcal{A} \subseteq \mathbb{R}^N$ is said to be Schur-convex on $\mathcal{A}$ if*

$$\boldsymbol{x} \prec \boldsymbol{y} \text{ on } \mathcal{A} \quad \rightarrow \quad f(\boldsymbol{x}) \leq f(\boldsymbol{y}).$$

Similarly, f is said to be Schur-concave on $\mathcal{A}$ if

$$\boldsymbol{x} \prec \boldsymbol{y} \text{ on } \mathcal{A} \quad \rightarrow \quad f(\boldsymbol{x}) \geq f(\boldsymbol{y}).$$

It is important to remark that this classification of functions do not form a partition of the set of all functions: neither of the two sets disjoint, nor cover the entire space of functions.

10.4 Optimal Power Allocation

As a reminder for the reader, the filter $\boldsymbol{F} = \text{diag}(\boldsymbol{F}_1 \ldots \boldsymbol{F}_N) \in \mathbb{C}^{NT \times M}$ is designed to optimize the following scalar function:

$$\varepsilon^2 = \sum_n \varepsilon_n^2 = \text{tr}(\boldsymbol{E}) = \sum_{\forall i} \sum_{\forall n} \varepsilon_{n,i}^2 \tag{10.5}$$

subject to $\sum_n \text{tr}(\boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H) \leq \mathcal{E}$. Here, $\varepsilon_{n,i}^2$ represents the MSE of the i-th symbol at the n-th carrier. That is,

$$\varepsilon_{n,i}^2 = \left[\boldsymbol{B}_n (\boldsymbol{R}_{\boldsymbol{v}_n}^{-1} + \boldsymbol{F}_n^H \boldsymbol{H}_n^H \boldsymbol{R}_{\boldsymbol{n}_n}^{-1} \boldsymbol{H}_n \boldsymbol{F}_n)^{-1} \boldsymbol{B}_n^H \right]_{ii}.$$

Notice that the former problem is convex for a given inner matrix —we refer to inner matrix the enclosed product between the two feedback filters in the former expression— and having the feedback filter as the optimization variable. However, for a given $\boldsymbol{B}_n$, minimize the former problem in terms of the precoding filter is equivalent to minimize the inner Hermitian matrix, which is formed by autocorrelation matrices and quadratic forms, but it is a complicated non-convex problem. Keeping in mind that the output signal of the modulo operator is uncorrelated and white due to the dithering with variance $\sigma_{v_n}^2$, and defining $\boldsymbol{R}_{\boldsymbol{H}_n}$ as the semi-definite Hermitian matrix $\boldsymbol{R}_{\boldsymbol{H}_n} \triangleq \sigma_{v_n}^2 \boldsymbol{H}_n^H \boldsymbol{R}_{\boldsymbol{n}_n}^{-1} \boldsymbol{H}_n$, the optimization is governed by $\left[(\boldsymbol{I} + \boldsymbol{F}_n^H \boldsymbol{R}_{\boldsymbol{H}_n} \boldsymbol{F}_n)^{-1} \right]_{ii}$. Since the problem has been diagonalized, we can vectorize the diagonal of the error matrix into a vector

$$\boldsymbol{e}_n \triangleq \text{diag}((\boldsymbol{I} + \boldsymbol{F}_n^H \boldsymbol{R}_{\boldsymbol{H}_n} \boldsymbol{F}_n)^{-1})$$

and, therefore, the optimization problem can be reformulated as the following Schur problem.

Theorem 3 (Optimal Precoding Matrix for Schur-concave Objective Functions)**.** *Let $f_0 : \mathbb{R}^{M_n} \longrightarrow \mathbb{R}$ be a Schur-concave function. The optimal solution of the following constrained optimization problem:*

$$\begin{aligned} \min_{\boldsymbol{F}_n} \quad & f_0(\boldsymbol{e}_n) \\ s.\ t. \quad & tr(\boldsymbol{F}_n \boldsymbol{F}_n^H) \leq \mathcal{E}_n \ , \ \sum_n \mathcal{E}_n = \mathcal{E}, \end{aligned}$$

where matrix $\boldsymbol{F}_n \in \mathbb{C}^{T \times M_n}$ is the optimization variable, is given by

$$\boldsymbol{F}_n = \boldsymbol{\mathcal{U}}_n \boldsymbol{\Sigma}_n,$$

where $\boldsymbol{\mathcal{U}}_n \in \mathbb{C}^{T \times \breve{M}_n}$ has as columns the eigenvectors of $\boldsymbol{R}_{\boldsymbol{H}_n}$ corresponding to the $\breve{M}_n$ largest eigenvalues in increasing order, and $\boldsymbol{\Sigma}_n = [\boldsymbol{0}\ diag(\{\sigma_{n,i}\}) \in \mathbb{C}^{\breve{M}_n \times M_n}$ has zero elements, except along the rightmost main diagonal, whose elements are $\sigma_{n,i} > 0$, for $1 \leq i \leq M_n$.

The proof can be found in [SSB+02]. The parameter $\breve{M}_n$ is associated to the number of nonzero channel eigenvalues but, by design, we must consider as well the number of symbols that the flat-fading MIMO channels allows to be allocated; so that $\breve{M}_n \triangleq \min(M_n, \text{rank}(\boldsymbol{R}_{\boldsymbol{H}_n}))$. Notice that in the case in which only one symbol per carrier is transmitted each channel use ($M_n = 1$), the former theorem says that this symbol must be transmitted through the eigenmode with the highest gain. On the other hand, when the system tries to allocate M_n symbols on a certain carrier, it may be the case that the channel has lower rank. If this happens, the remainder $M_n^0 = M_n - \breve{M}_n$ remaining symbols are associated with zero eigenvalues (that is why the matrix $\boldsymbol{\Sigma}_n$ may have zeroes).

The EVD of the channel correlation matrix is performed with a descending order of its eigenvalues. That is, $\boldsymbol{R}_{\boldsymbol{H}_n} = \boldsymbol{\mathcal{U}}_n \boldsymbol{\Lambda}_n \boldsymbol{\mathcal{U}}_n^H$, where the inner diagonal matrix contains the $\breve{M}_n$ eigenvalues, $\{\lambda_{n,i}\}_n$, in decreasing order.

Finally, the last step is to find the values of the power allocation parameters, $\sigma_{n,i}$ for all the carriers and all the transmitted symbols with the power constraint (9.6). With the structure of each of the precoding filters, the diagonalized MSE corresponding to the n-th carrier and the i-th symbol turns out to be:

$$e_{n,i}^2 = [(\boldsymbol{I} + \boldsymbol{F}_n^H \boldsymbol{R}_{\boldsymbol{H}_n} \boldsymbol{F}_n)^{-1}]_{ii} = [(\boldsymbol{I} + \boldsymbol{\Sigma}_n^H \boldsymbol{\Lambda}_n \boldsymbol{\Sigma}_n)^{-1}]_{ii} = \frac{1}{1 + \lambda_{n,i} \sigma_{n,i}^2},$$

for which the minimization of the MSE as defined in (10.5) is translated to minimize the following scalar function

$$f_0(\boldsymbol{e}) = \sum_{\forall i} \sum_{\forall n} e_{n,i}.$$

Since f_0 is a Schur-concave function (assuming $e_i^2 \geq e_{i+1}^2$, cf. [PCL03]), the structure of the precoding matrix is given by theorem 3 and, as a result, the power constraint (9.6) can be exclusively written as a function of parameters $\sigma_{n,i}^2$ due to the orthonormality of matrix $\boldsymbol{\mathcal{U}}_n$. Noticing that the same modulo operation is applied at all the carriers, $\sigma_{v_n}^2 = \sigma_v^2$ for all n, the power constraint reads:

$$\sum_{\forall n} \sigma_{v_n}^2 \text{tr}(\boldsymbol{F}_n \boldsymbol{F}_n^H) = \sigma_v^2 \sum_{\forall n} \text{tr}(\boldsymbol{\mathcal{U}}_n \boldsymbol{\Sigma}_n \boldsymbol{\Sigma}_n \boldsymbol{\mathcal{U}}_n^H) = \sigma_v^2 \sum_{\forall i} \sum_{\forall n} \sigma_{n,i}^2 \leq \mathcal{E}.$$

Hence, the problem of minimizing $f_0(\boldsymbol{e}) = \sum_{\forall i} \sum_{\forall n} (1 + \lambda_{n,i} \sigma_{n,i}^2)^{-1}$ subject to $\sum_{\forall i} \sum_{\forall n} \sigma_{n,i}^2 \leq \mathcal{E}/\sigma_v^2$ can be solved very efficiently because the solution has a water-filling interpretation. From the KKT

optimality conditions (e.g. [Gal68]), the power allocation is given by:

$$\sigma_{n,i}^2 = \left(\frac{\mathcal{E}/\sigma_v^2 + \sum_{\forall j} \sum_{\forall m} \lambda_{m,j}^{-1}}{\sum_{\forall j} \sum_{\forall m} \lambda_{m,j}^{-1/2}} \lambda_{n,i}^{-1/2} - \frac{1}{\lambda_{n,i}} \right)^+, \tag{10.6}$$

where m indexes the carriers whereas j indexes the symbol.

We want to highlight the similarities to information theory. Recent approaches (e.g. [SD08]) base the optimization problem on the determinant of the error matrix. If this is done, the objective function becomes Schur-convex and the solution of the power allocation is identical to the one obtained from the maximization of the mutual information on parallel Gaussian channels. Another interesting point is that the use of OFDM splits the task of decomposing an $NT \times NR$ matrix onto N small decompositions of $T \times T$ squared matrices. This is very efficient from the computational point of view. In addition, the diagonalization of the problem has brought into consideration the theory of Schur-convexity and, at its turn, this one has allowed to find a closed-form structure for the precoding filter as well as an efficient way of determining the power allocation over all the carriers and all the transmitting antennas. Finally, the optimization function could be a weighted version of the arithmetic mean of the MSEs. This can be done when the system wants to favor some carriers or eigenmodes.

Nevertheless, expression (10.6) can be easily interpreted as a water-filling process where the indicator of how good is each eigen mode is the associated eigenvalue. Recalling that $\boldsymbol{R}_{\boldsymbol{H}_n} = \sigma_{v_n}^2 \boldsymbol{H}_n^H \boldsymbol{R}_{\boldsymbol{n}_n}^{-1} \boldsymbol{H}_n$, the channel modes with high channel gain and low noise level show a high eigenvalue and, then, since its inverse turns to be small —that is, a mode with low noise—, it is more water-filled; whereas poor modes can be even discarded.

10.5 Remark on Maximum Diversity Model: RAKE Detection

OFDM is in general a bad modulation because its performance is very closed related to the frequency response of the channel at each carrier. In this sense, it is highly probably that, in a block transmission, an error occurs at a given carrier due the attenuation. For this reason, evaluating the system in terms of uncoded SER for maximum throughput is not a realistic case. Instead, most systems use source coding and randomly —i.e. interleave— place the coded symbols at the carriers or, alternatively, a maximum diversity approach is used and the same symbol is placed at all the bands, making the system very robust to channel frequency fadings as the probability of all the carriers being down is very low.

Firstly, the use of THP has resulted in a diagonalization of the error matrix and the theory of majorization could be applied. Secondly, the structure of the precoding matrix has diagonalized the channel correlation matrix $\boldsymbol{R}_{\boldsymbol{H}_n}$, fact that turns the design matrix $\boldsymbol{\Phi}$ and, consequently, the feedback matrix diagonal. With this particular structures and noticing that both output of the modulo operation and noise signals are i.i.d. signals, the full chain of the MIMO-OFDM system is diagonalized as well. We shall define the equivalent overall chain vector as

$$\boldsymbol{q} = [\boldsymbol{q}_1^T \dots \boldsymbol{q}_N^T]^T \triangleq \text{diag}(\boldsymbol{GHFB}^{-1}), \tag{10.7}$$

this vector would contain all ones in a MMSE-ZF optimization process. Since the same information is sent through several parallel channels, the inverse strategy or detector γ^{-1} must be optimal in the

sense of obtaining the constructive signal to be sent to the decider with the optimal conditions for it. Since the decider is maximum likelihood, these conditions are (cf. [Cio02]) maximum signal-to-noise ratio (SNR). Hence, in a maximum diversity strategy, the receiver must obtain the received MIMO symbols and constructively join them under a RAKE [Cio02] receiver criterion in view of this equivalent channel. A RAKE matched filter is a set of parallel matched filters each operating on one of the diversity channels in a diversity transmission, in our case through frequency carriers. Mathematically, the signal sent to the precoder is the weighted sum of the output of the receiver filter signal, after compensating the dithering and the modulo operation, as:

$$\tilde{\boldsymbol{s}} = \sum_{n=1}^{N} \mathrm{diag}(\boldsymbol{q}_n^H)\tilde{\boldsymbol{s}_n}.$$

The RAKE was originally so named by Green and Price in 1958 because of the analogy of the various matched filters being the "fingers" of a garden rake and the sum corresponding to the collection of the fingers at the rake's pole handle, nomenclature thus often being left to the discretion of the inventor, however unfortunate for posterity. The RAKE is sometimes also called a diversity combiner, although the latter term also applies to other lower-performance suboptimal combining methods that do not maximize overall signal to noise strength through matched filter.

Chapter 11

Other Optimization Techniques

With the aim of providing other design criterion and present enough diversity in optimization processes to give the reader a critical point of view, the THP system for the MIMO-OFDM channel with power constraint has been designed under MMSE-ZF,Cholesky factorization and QR decomposition methods.

11.1 Cholesky Factorization

With great similarities to the SISO channel, the THP for the MIMO-OFDM channel is design considering the N flat-fading MIMO channels due to the OFDM modulation, each one using the mathematical results and optimization process stated in section 6.3.

Fistly, the optimal receive filter that minimizes the trace of the error matrix at the n-th carrier is given by the result (10.1), which is a result of the orthogonality principle that holds at the receiver between the error signal and the received signal. Secondly, the error matrix can be written as a quadratic form of the feedback filter with the design matrix $\boldsymbol{\Phi}_n \triangleq (\boldsymbol{R}_{\boldsymbol{v}_n}^{-1} + \boldsymbol{F}_n^H \boldsymbol{H}_n^H \boldsymbol{R}_{\boldsymbol{n}_n}^{-1} \boldsymbol{H}_n \boldsymbol{F}_n)^{-1}$. This result can be obtained by analogy to (6.15), using the SISO$\longrightarrow$MIMO analogies —with $\boldsymbol{F}_n$ fixed—:

$$\begin{aligned} \boldsymbol{G} &\longrightarrow \boldsymbol{G}_n \in \mathbb{C}^{M_n \times R} \\ \boldsymbol{B} &\longrightarrow \boldsymbol{B}_n \in \mathcal{C}^{M_n \times M_n} \\ \boldsymbol{H} &\longrightarrow \boldsymbol{H}_n \boldsymbol{F}_n \in \mathcal{C}^{M_n \times M_n} \\ \boldsymbol{x} &\longrightarrow \boldsymbol{v}_n \in \mathcal{C}^{M_n \times 1}, \end{aligned}$$

or directly using (10.2) and (10.3). In any case, the minimization of the trace of the error matrix is a convex problem, but the feedback filter requires to be lower triangular with one main diagonal. The simple approach is to include this restriction to the Lagrangian function, as it has been done in the MMSE chapter. However, an alternative is to use the Cholesky factorization and the Weyl's inequality [Wey49] to determine the optimal feedback matrix. Specifically, $\boldsymbol{\Phi}_n$ is decomposed as

$$\boldsymbol{\Phi}_n = \boldsymbol{L}_n \boldsymbol{L}_n^H,$$

where $\boldsymbol{L}_n$ is a lower triangular matrix with positive real diagonal elements. The objective function can be rewritten as the Frobenius norm of $\boldsymbol{B}_n \boldsymbol{L}_n$ so that —please refer to subsection 6.3.2— the feedback

filter that attains the lower bound of the MSE, and hence is optimal, is

$$\boldsymbol{B}_n = \mathrm{diag}(\boldsymbol{L}_n)\boldsymbol{L}_n^{-1}.$$

Notice that the error matrix has been diagonalized due to this structure and, more interestingly, the individual MSEs are given by the squared diagonal elements of the Cholesky matrix. That is, $\varepsilon_{n,i}^2 = [\boldsymbol{E}_n]_{ii} = [\boldsymbol{B}_n\boldsymbol{\Phi}_n\boldsymbol{B}_n^H]_{ii} = [\boldsymbol{L}_n^2]_{ii}$.

Finally, the precoding matrix that performs the optimal power allocation is given by the same expression as in theorem 3 with the water-filling parameters in (10.6).

11.2 MMSE-ZF

The optimization of (10.5) under the power constraint (9.6) and the ZF condition $\boldsymbol{GHFB}^{-1} = \boldsymbol{I}$, can be obtained from the MMSE solution by bringing the SNR to infinity or by adding the ZF condition to the Lagrangian optimization, as done in the SISO case. For simplicity the fist approach will be chosen in this single-user MIMO part.

The first step is to obtain the design matrix at the n-th carrier from the expression in (10.2). The precoded signal inverted correlation matrix can be neglected with respect to the quadratic form which involves the inverse of the noise correlation matrix and, as a result, $\boldsymbol{\Phi}_n = (\boldsymbol{F}_n\boldsymbol{H}^H\boldsymbol{R}_{\boldsymbol{n}_n}\boldsymbol{H}_n\boldsymbol{F}_n)^{-1}$. Here we might define as well $\boldsymbol{R}_{\boldsymbol{H}_n} \triangleq \boldsymbol{H}_n^H\boldsymbol{R}_{\boldsymbol{n}_n}^{-1}\boldsymbol{H}_n$, without considering the scalar σ_v^2 as now the design matrix is the inverse of only one term. The feedback matrix is the solution of the optimization problem of minimizing the trace of the error matrix at the n-th carrier, which is of the quadratic form in terms of this new design matrix. The result is known:

$$\boldsymbol{B}_n = \mathrm{diag}^{-1}(\boldsymbol{\Phi}_n^{-1})(\boldsymbol{U}_n^T \odot \boldsymbol{\Phi}_n^{-1}).$$

Hence, the expression of the receive filter —notice that here, though it is a ZF design, we are minimizing the residual individual MSEs— from the result in (10.1) after applying the inversion lemma derives to the ZF condition:

$$\boldsymbol{G}_n = \boldsymbol{B}_n(\boldsymbol{H}_n\boldsymbol{F}_n)^{\dagger},$$

where the superscript † denotes pseudo-inverse. In the SISO channel, the channel matrix is always tall by design, so the pseudo-inverse is given by the number of columns and the invertible product is $\boldsymbol{H}^H\boldsymbol{H}$. Here, the product between the channel matrix and the precoder filter is a $R \times M_n$ matrix, which is always tall —for instance when $R > T$, $M_n = T$— or, at most, square —that is, $R \leq T$, for which $M_n = R$; so here as well, $(\boldsymbol{H}_n\boldsymbol{F}_n)^{\dagger} = (\boldsymbol{F}_n^H\boldsymbol{H}_n^H\boldsymbol{H}_n\boldsymbol{F}_n)^{-1}(\boldsymbol{H}_n\boldsymbol{F}_n)$.

Finally, since the ZF condition is accomplished for any precoder matrix, filter $\boldsymbol{F}_n$ is selected to minimize the Schur-concave function described by the arithmetic sum of the MSEs. The THP diagonalizes the error matrix so the only important terms to consider are the ones at the main diagonal of the design matrix. Vectorizing them onto a vector $\boldsymbol{e}_n \triangleq \mathrm{diag}((\boldsymbol{F}_n^H\boldsymbol{R}_{\boldsymbol{H}_n}\boldsymbol{F}_n)^{-1})$, the optimization problem is solved with the structure of theorem 3, $\boldsymbol{F}_n = \boldsymbol{\mathcal{U}}_n\boldsymbol{\Sigma}_n$. Nevertheless, now the individual vectorized MSEs become simpler:

$$e_{n,i}^2 = \frac{1}{\lambda_{n,i}\sigma_{n,i}^2},$$

and the set of power allocation parameters, $\sigma_{n,i}^2$, are obtained by minimizing $\sum_{\forall n}\sum_{\forall i} e_{n,i}^2$ under the power constraint $\sum_{\forall n}\sum_{\forall i}\sigma_{n,i}^2 \leq \mathcal{E}/\sigma_v^2$. The power allocation, then, consists in placing the portion of

available energy at the (n, i) mode proportional to the inverse squared root of the associated eigenvalue. In other words,

$$\sigma_{n,i}^2 = \frac{\mathcal{E}/\sigma_v^2}{\sum_{\forall m}\sum_{\forall j}\lambda_{m,j}^{-1/2}}\lambda_{n,i}^{-1/2}. \tag{11.1}$$

In this case, the power is allocated directly proportional to the channel bad conditions —as a remainder, a high $1/\lambda_{n,i}$ stands for a noisy channel— to attain the ZF equalization. For this reason, the MMSE-ZF is expected to lag behind with respect to the MMSE.

11.3 QR Decomposition

This last section exploits the QR decomposition of a selected matrix. In view of the power allocation matrix $\boldsymbol{F}_n$, the product $\boldsymbol{H}_n\boldsymbol{F}_n$ is decomposed as

$$\boldsymbol{H}_n\boldsymbol{F}_n = \boldsymbol{Q}_n\boldsymbol{R}_n,$$

where $\boldsymbol{Q}_n$ is an $R\times M_n$ orthonormal matrix and $\boldsymbol{R}_n$ is the right lower $M_n\times M_n$ triangular matrix. This fixes the feedback matrix to the lower triangular matrix obtained through the former decomposition with the main diagonal weighted as follows

$$\boldsymbol{B}_n = \mathrm{diag}(\boldsymbol{R}_n)\boldsymbol{R}_n,$$

and the receiver filter is given by the ZF condition as

$$\boldsymbol{G}_n = \mathrm{diag}(\boldsymbol{R}_n)\boldsymbol{Q}_n^H.$$

It is easy to check that the chain $\boldsymbol{G}_n\boldsymbol{H}_n\boldsymbol{F}_n\boldsymbol{B}_n^{-1}$ has been fully equalized and diagonalized to the identity since, by construction, $\boldsymbol{G}_n\boldsymbol{H}_n\boldsymbol{F}_n\boldsymbol{B}_n^{-1} = \mathrm{diag}(\boldsymbol{R}_n)\boldsymbol{Q}_n^H\boldsymbol{Q}_n\boldsymbol{R}_n\boldsymbol{R}_n^{-1}\mathrm{diag}^{-1}(\boldsymbol{R}_n) = \boldsymbol{I}$. Finally, although a specific decomposition has been used, the ZF condition still holds, so the error matrix can be expressed through the design matrix $\boldsymbol{\Phi}_n$ for the THP ZF as in section 11.2 as, using the inversion lemma:

$$\boldsymbol{E}_n = \boldsymbol{G}_n\boldsymbol{R}_{\boldsymbol{n}_n}\boldsymbol{G}_n^H = \boldsymbol{B}_n^H(\boldsymbol{F}_n\boldsymbol{H}^H\boldsymbol{R}_{\boldsymbol{n}_n}^{-1}\boldsymbol{H}_n\boldsymbol{F}_n)^{-1}\boldsymbol{B}_n,$$

so that the precoding transmit filter that performs the optimal power allocation by minimizing the diagonalized error matrix is given by the MMSE-ZF solution in 11.2. That is, $\boldsymbol{F}_n$ follows the structure in theorem 3 and the power allocation parameters are given by (11.1).

Chapter 12

Performance Analysis

The complete system depicted in the former chapter has been checked under MATLAB Monte-Carlo simulations by transmitting random sequences of QPSK-modulated symbols. The performance indicators are the SER and the normalized mean-square error (NMSE), computed statistically on the $\tilde{\boldsymbol{s}}$ signal. The strategy γ has been selected to be maximum spatial throughput and maximum frequency diversity, for which the same M QPSK symbols are placed on the N available carriers.

The frequency response of the MIMO-OFDM channel is obtained with the Fourier expression (9.3), where each lag $\boldsymbol{H}(\ell)$ is computed following the correlation model in (9.5) with the variance corresponding to the ℓ-th lag of the rich multi-path environment described in table 7.1 of the HIPERLAN radio network standard in [MS98].

The indicator of the system conditions (channel gain/attenuation and noise) is the average SNR. In the SISO channel, all the symbol energy is placed on the unique transmitting antenna and carrier, and the complex noise power is present in all the bandwidth. As a result and by tradition, the SNR is expressed under the E_b/N_o ratio since, as the energy placed at each channel use is $\mathcal{E} = 2E_b$, SNR $= \mathcal{E}/N_o$, being N_o the white noise level. Nevertheless, in a MIMO (and OFDM) channels, the available energy $\mathcal{E}$ at each channel use has to be split onto the N carriers and the T transmitting antennas, whereas the noise is present at the R receiving antennas of the N carriers. In this sense, the power allocation algorithm is in charge of optimally placing this power budged to the NT possible channel inputs, taking into consideration the NR different possible noise levels at each channel output. Hence, the average SNR, for the MIMO-OFDM channel is defined as

$$\text{SNR} \triangleq \frac{\mathcal{E}}{\text{tr}(\boldsymbol{R_n})}.$$

The following sections analyze the four THP solutions (MMSE, MMSE-Cholesky, MMSE-ZF and ZF-QR) under several conditions of spatial channel dimensions, $R \times T$, number of available OFDM carriers N, the spatial correlation parameters ρ_R and ρ_T, and the noise power distribution on frequency and receive antennas. In the sequel, uniform noise stands for having the same white noise variance at the N carriers and R antennas, while other distributions may be considered. Notice that the filter optimization has considered a very general model, as the system can be easily particularized to a SISO OFDM system or a single-carrier MIMO system by doing $R = T = 1$ and $N = 1$, respectively.

12.1 Overall Performance

The first visual result, after transmitting symbols through a 2×2 8-carriers OFDM MIMO channel with uniform noise (refer to figure 12.1) is double: the MMSE solutions outperform with respect to the ZF solutions, and the factorization techniques show the same performance. That is, the MMSE and the MMSE-Cholesky solutions equally perform, as the MMSE-ZF and ZF-QR do.

The MMSE filters work better due to their optimality in the MSE criterion since they do not require the ZF condition. In this line, the feedback filter cancels out the interference provoked by the MIMO channel —that is, the portion of signal coming from the other $M-1$ symbols when detecting a certain symbol— taking into account the expression of the precoding filter $\boldsymbol{F}$. This last filter, optimally places the energy budget on frequency and space under a water-filling algorithm and, finally, the receive filter has another degree of freedom to diminish the residual MSE. The ZF forcing solution, due to the additional constraint, forces the power allocation to be suboptimal and the feedback filter to completely equalize the system chain to the identity. This, at its turn, is reflected with the diversity achieved in frequency, since the MMSE-ZF and ZF-QR make the equivalent channel a few rich channel (refer to section 12.4).

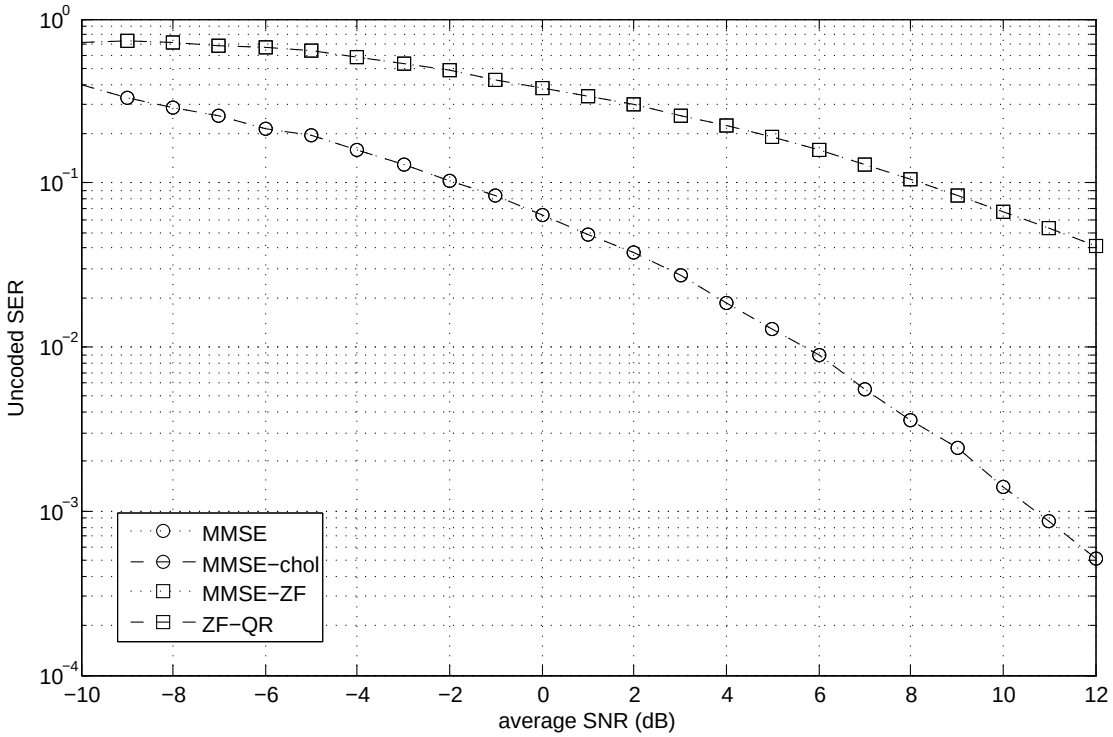
Contrary to the SISO case, here the factorization techniques do not under or out perform with respect to the MMSE and MMSE-ZF solutions. This is due to the characteristics of the channel matrix and the power allocation. With the use of the power allocation, which requires the EVD of the squared form of the channel, the channel regarded between the feedback filter and the receive filter, $\boldsymbol{HF}$ has a particular structure for which the design matrix, when the noise has uniform distribution, is lower triangular. Therefore, the optimal MMSE and MMSE-ZF are optimal as well as the factorization techniques. When the noise is not uniformly distributed on frequency and receiving antennas, there is a little difference in between the MMSE-ZF and the ZF-QR solutions, which is studied in section 12.5.

12.2 Power Allocation

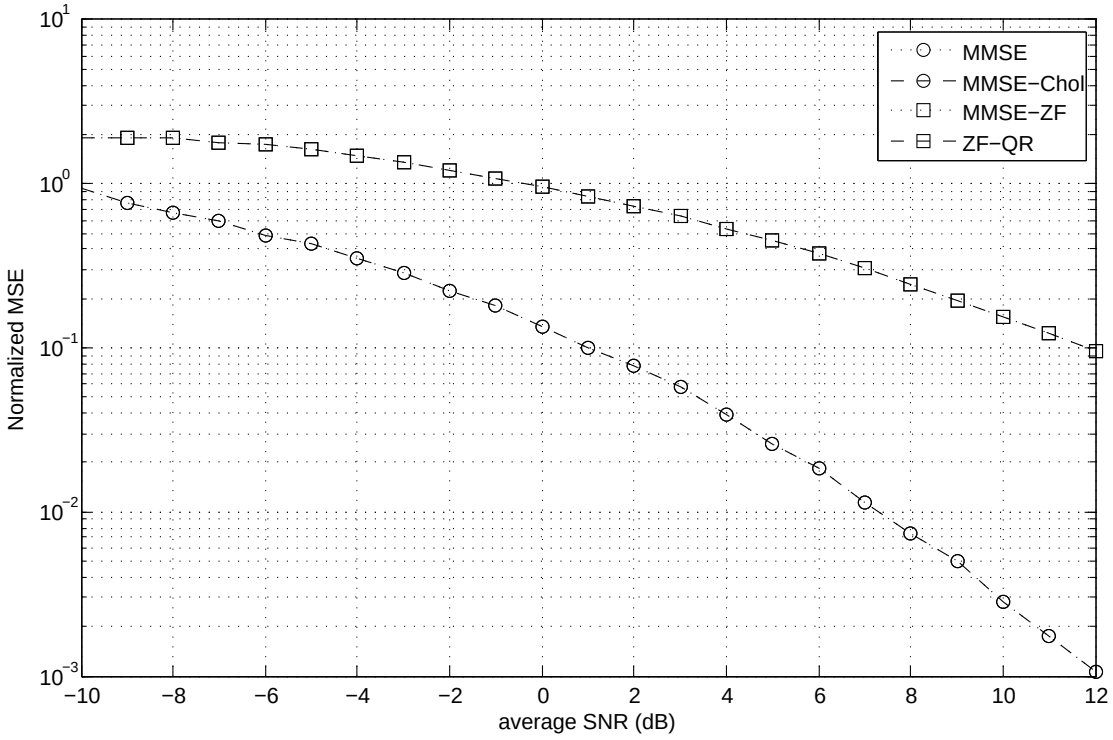
The power allocation has a big impact in the performance of the system regarding the fact that MMSE and MMSE-Cholesky share the same algorithm, as well as MMSE-ZF and ZF-QR. With the aim of understanding the expressions (10.6) and (11.1), the power allocation algorithms have been simulated in a 1×1 channel in a $N = 48$ carriers Wireless LAN typical OFDM where the left bands are better than the right ones, exponentially in descend order. In practice, the quality of the channels depend upon the noise power and the channel gain, so they must not be in any order. However, the power allocation algorithm takes all the possible eigen modes of the MIMO-OFDM channel —at most NT— and performs the power allocation. Here, by simplicity, we consider a SISO channel since it is more intuitive to consider only the frequency axis rather than eigen modes.

Regarding figure 12.2, the two algorithms show a very different behavior.

Firstly, the MMSE and MMSE-Cholesky water-filling process selects only the best bands to place energy while discards the worst modes. Though it is almost like the classical water-filling algorithm from the information theory, here the behavior is different. At high SNR, e.g. 8 dB, the best modes show large eigenvalues and, as a consequence, very low inverses. From the expression (10.6), the squared root of the inverse of the eigenvalue is predominant with respect to the inverse, to the allocated power goes almost proportionally to $\lambda^{-1/2}$, that is, proportionally to the noise. When the band becomes relatively noisy, the terms λ^{-1} and $\lambda^{-1/2}$ are of the same order and the system performs



(a)



(b)

Figure 12.1: SER (a) and NMSE (b) for the $N = 8$ carriers THP 2×2 MIMO channel with $\rho_R = \rho_T = 0.2$ correlation parameters and uniform noise versus average SNR.

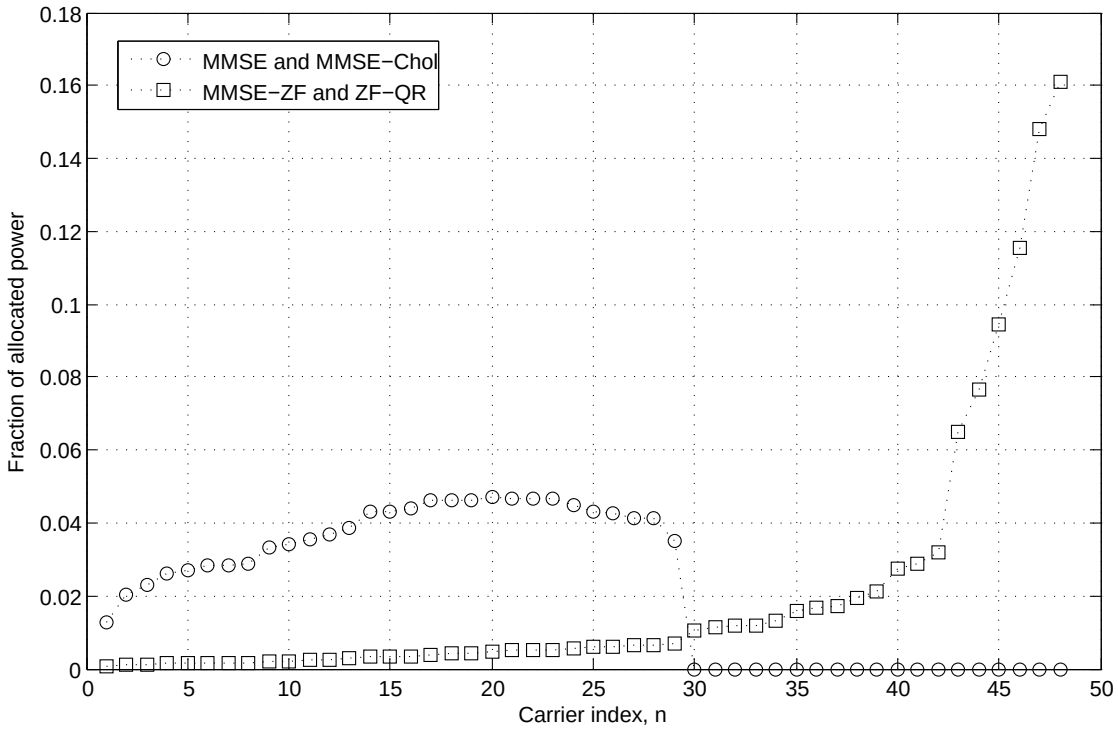


Figure 12.2: Fraction of Power Allocated for the 1×1 THP MIMO channel, with $N = 48$ OFDM carriers, at an average SNR = 8 dB.

very similarly to the classical water-filling: it starts placing less energy to the noisiest bands. Finally, when the budget is run off, it stops and does not use the worst bands.

Secondly, the MMSE-ZF and ZF-QR allocation process uses always all the bands. This can be clearly seen in equation (11.1) —recall that all the eigenvalues are positive since the channel correlation matrix $\boldsymbol{R}_{\boldsymbol{H}_n}$ is Hermitian and, thus, semi-definite positive— where the power allocated at the n-th band is proportional to the squared root of the inverse of its associated eigenvalue. As a result, the power is exponentially allocated and the more noisy the band is, more power is placed. Although it may sound contradictory, this is due to the ZF condition. The condition states that all the chain has to be equalized to the identity and, as a result, the residual error matrix is minimized with the inverse quadratic form $\boldsymbol{F}\boldsymbol{H}\boldsymbol{R}_{\boldsymbol{n}}\boldsymbol{H}^H\boldsymbol{F}^H$, whose result is a known solution in terms of the forward filter, given by (11.1). Interestingly, this power allocation goes in conjunction with the ZF: the left bands are less fed since they do not require as much power as the right bands to equalize them.

The differences in the power allocation are translated in the performance of the MMSE solution versus the ZF solutions, as depicted in figure 12.1. The MMSE power allocation is cleverer that the ZF since it takes profit of the high gain modes rather than spending energy trying to overcome the noisy bands.

12.3 MIMO Parameters

This section is intended to explain how the parameters of the MIMO channel affect the performance of the THP system. Concretely, the system is analyzed by varying the number of transmitting antennas and, finally, the correlation present at the receiving antennas.

12.3.1 Number of Transmitters

The MIMO-OFDM system has been designed to offer maximum diversity in frequency and maximum throughout in space. However, since the spatial throughput is given by the possible amount of information to be transmitted to the MIMO channel —that is, its rank—, some channel characteristics can turn the spatial dimension to achieve diversity as well. This kind of discussion has more sense in a single-user scenario when there is a clear difference between the number of transmitting and receiving antennas.

Figure 12.3 depicts an $1 \times T$ 1-carrier MIMO system, at an average SNR of 0 dB with correlation parameter at the transmitting antennas of $\rho_T = 0.2$ and uniform noise. Since $M = 1$, only one symbol is transmitted per channel use and the precoding filter $\boldsymbol{F}$ places this symbol on the T transmitting antennas. In this way, the receiver obtains T versions of this symbol, attaining spatial diversity. As expected, the larger the number of transmitters, the better the performance for the four solutions, which behave in parallel versus T. This result demonstrates that in a single-carrier MIMO system, though the channel capacity is enhanced, it can be used in a maximum diversity strategy placing the same symbol to the M possible symbols, reducing the throughput but improving the error rate.

12.3.2 Spatial Correlation

Another parameter that is subject to alter the performance of the system is the correlation among the transmitting and receiving antennas. The diversity model chosen in this work is explained in section 9.2, for which the paths between each transmitting and receiving antenna can be more or less similar (correlated) to the others. The use of MIMO channels is justified when the signals obtained at its output are highly uncorrelated, since the conditions of one output are totally distinct to the rest. Following the model (9.5), the correlation parameters that define the correlation matrices determine the correlation level of the MIMO channel.

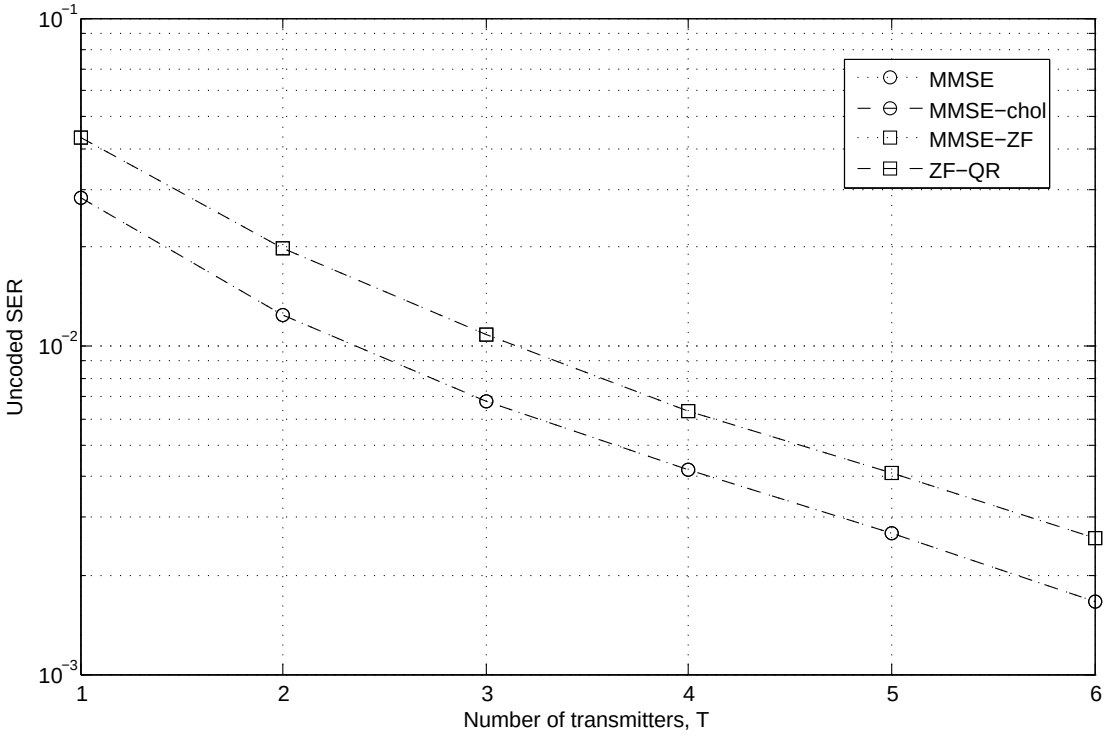
In a very high transmit correlation scenario, $\rho_T = 1$ and matrix $\boldsymbol{R}_T^{1/2}$ becomes all ones and, as a consequence, the resulting channel matrix $\boldsymbol{H}_n$ has identical column vectors and its rank reduces to one. This means that the symbols placed at the transmitting antennas see the same channel conditions to a given receiver.

In a very high receive correlation scenario, it is the receive correlation matrix $\boldsymbol{R}_R$ which becomes all ones and, thus, the channel matrix has identical row vectors and rank equal to one. This is translated to the fact that the channel conditions from a given transmitter are the same to all the receivers.

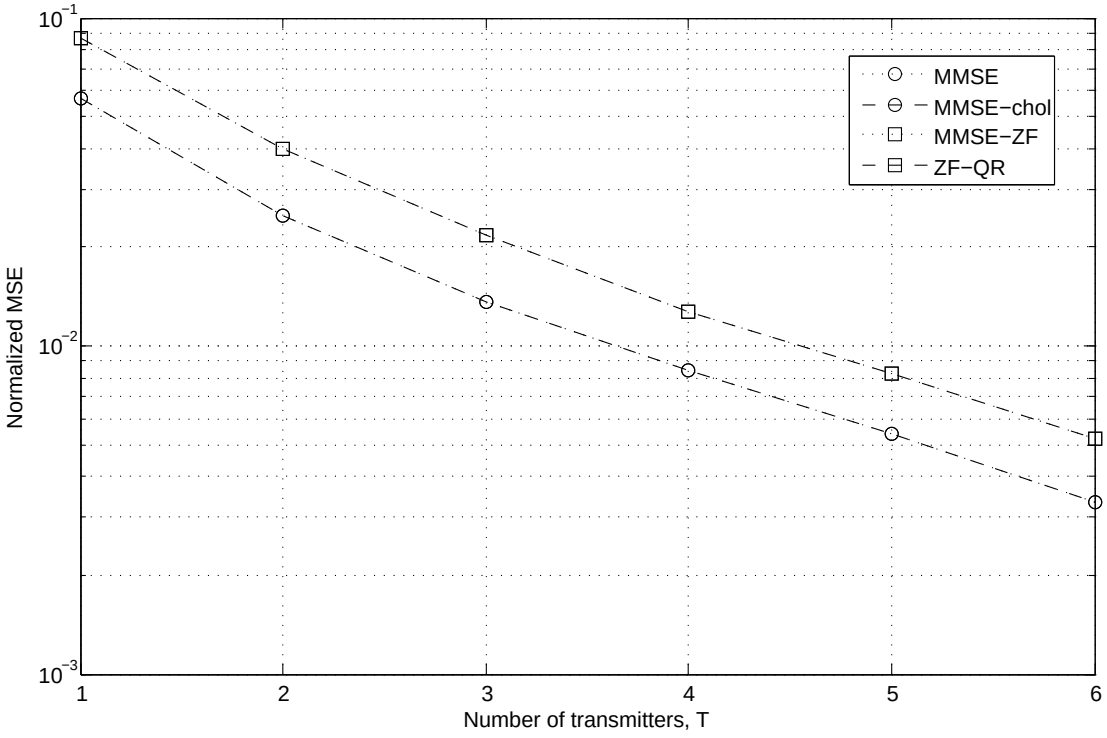
Finally, if both receive and transmit correlation parameters are the unity, the channel matrix has the same value at all its elements and the diversity is reduced to the nullity.

When one of these three last cases happens, the channel matrix is not invertible and the system becomes unstable. However, this situation is very improbable in practical designs as the antennas are uncorrelated if antenna spacing is greater than the half wavelength [Stü01], easily satisfied in systems that use carrier frequencies in the order of GHz.

In this work, the THP system has been tested with an 4×4 8-carriers MIMO-OFDM system at an average SNR of 8 dB, as appreciated in figure 12.4. The transmit correlation parameter has been fixed to 0.2, whereas ρ_R sweeps from 0 to 1. The best performance is found in the truly uncorrelated scenario with $\rho_R = 0$, though it is not realistic since a little coupling among the antennas is unavoidable. On the other extreme, when the correlation increases, both SER and MSE curves increase, leading to a useless system.

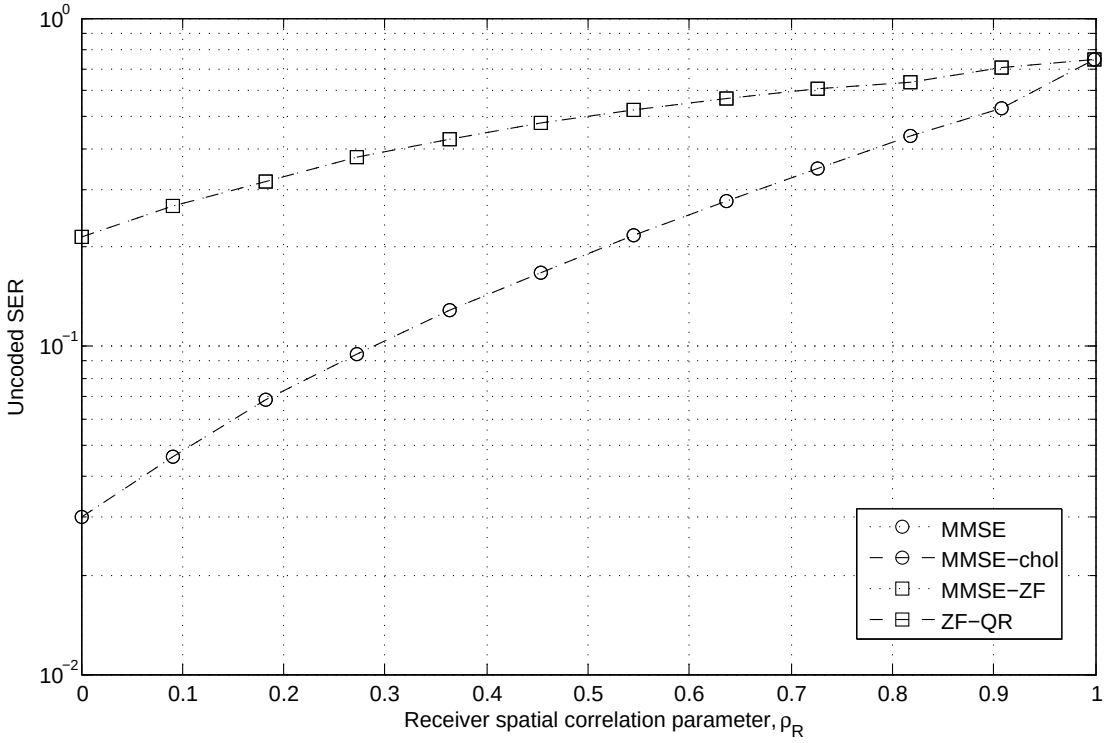


(a)

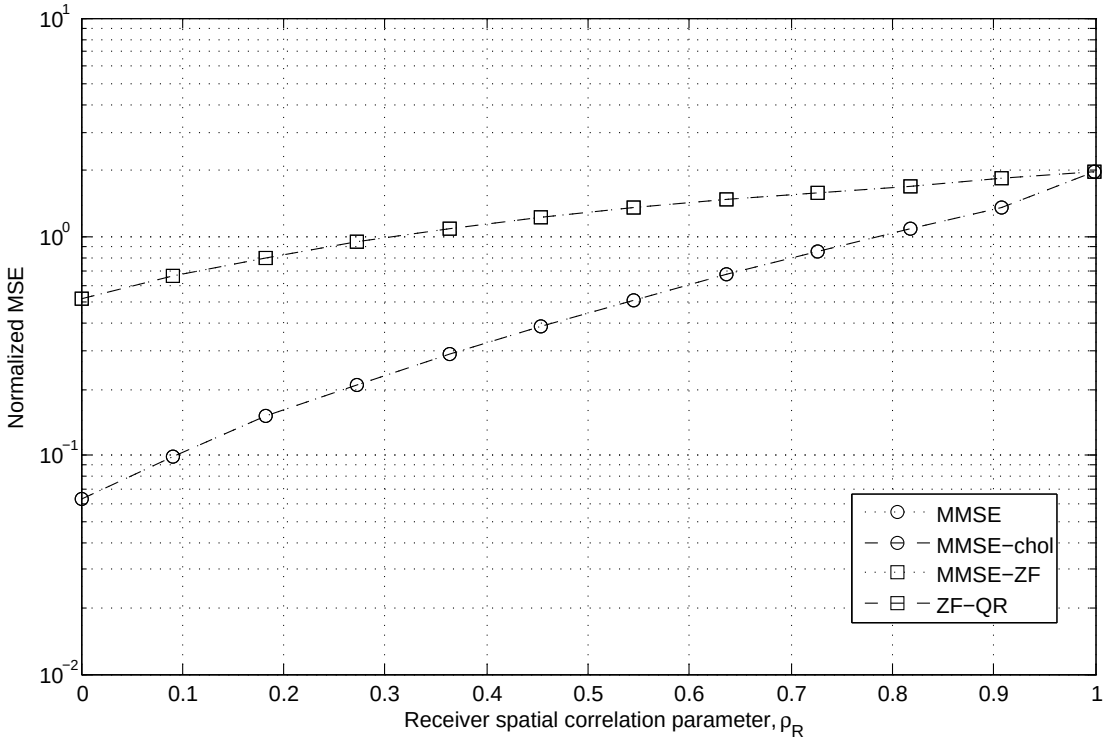


(b)

Figure 12.3: SER (a) and NMSE (b) for the single-carrier THP $1 \times T$ MIMO channel with $\rho_R = \rho_T = 0.2$ correlation parameters and uniform noise versus the number of transmitting antennas, T, at an average SNR = 0 dB.



(a)



(b)

Figure 12.4: SER (a) and NMSE (b) for the $N = 8$ carriers THP 4×4 MIMO channel with $\rho_T = 0.2$ correlation parameter and uniform noise versus the correlation parameter at the receiver antennas, ρ_R, at an average SNR = 8 dB.

12.4 OFDM Parameters

The main reason why OFDM has been used in MIMO systems has been justified in section 10.4 as the complexity of the problem has been separated to N easily attainable matrix problems. This reduction, yet, is done without loss of generality, since the single-carrier system can be obtained when $N = 1$. With the use of OFDM there is an existent trade-off in how the spectrum is used by the system. The period of the OFDM symbol has been lengthened with respect to the single-carrier symbol period, $T_s = 1/W$, as $T_{OFDM} = NT_s + T_{CP}$, being W the system bandwidth, N the number of OFDM carriers and T_{CP} the length of the cycle prefix. For this reason, the OFDM modulation is often used in conjunction with a coding scheme so as N different coded symbols are placed on the carriers. In this sense, at each channel use, the system transmits N symbols of duration approximately N times the duration of the usual symbol, for which there is no loss of throughput.

However, in this work, the OFDM modulation has not been tested under a coding scheme since this is not its scope, and a maximum diversity scenario has been proposed. Under this paradigm, a channel use assumes that the same MIMO symbol is placed on the N carriers, with symbol duration T_{OFDM}. This means that the diversity of the system is enhanced, at the price of reducing the throughput by a factor of N. Nevertheless, there are some similarities in testing a coding scheme or the maximum diversity approach since both base their performance in the low probability in having an error at a large number of carriers at the same time. For this reason, the THP system has been simulated, as depicted in figure 12.5, in a 2×2 N-carriers MIMO-OFDM channel at a SNR of 0 dB with $\rho_R = \rho_T = 0.2$.

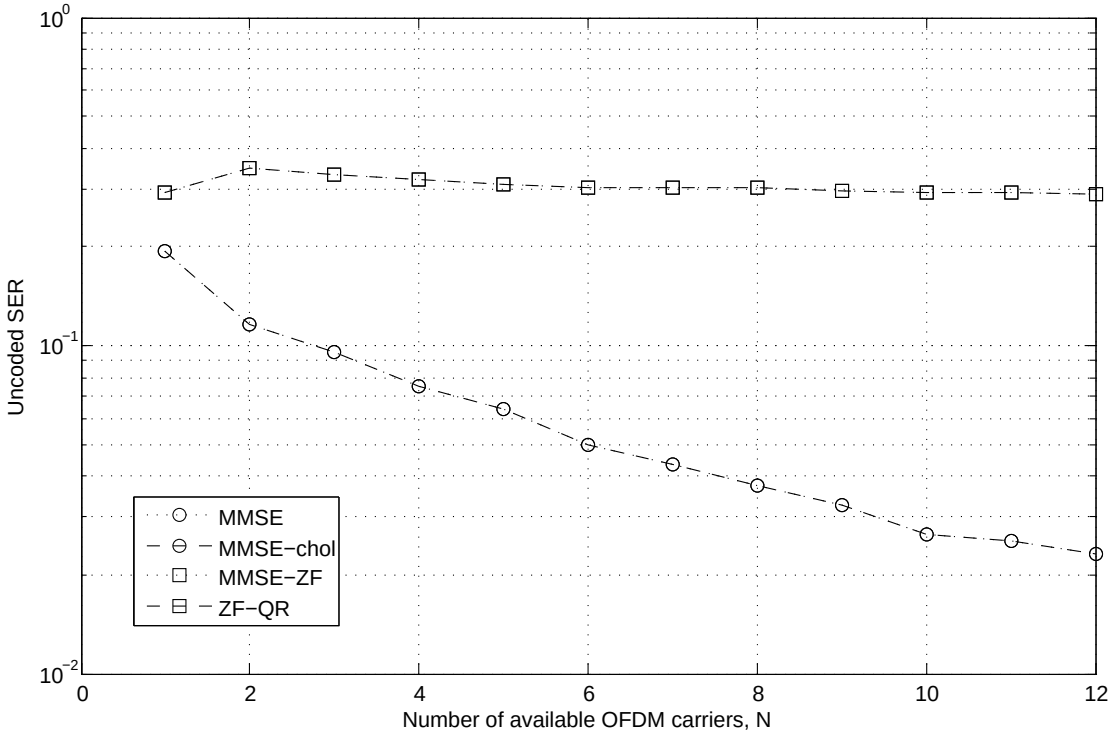
To understand the results in figure 12.5 we shall recall how the frequency diversity is combined at the receiver: by means of the RAKE and the equivalent vectorized channel 10.7.

On the one hand, the MMSE solutions behave as expected: the number of carriers, the degree of diversity and the performance of the system behave proportionally. The explanation behind this result is that the RAKE receiver is able to enhance the level of signal with respect to the level of noise since the detection is done with a bank of adapted filters. That is, after multiplying the output of each carrier by the transpose conjugate of $\boldsymbol{q}_n$, the useful part is correlation-like filtered —a very high value— whereas the noise part is filtered in an unadapted way. This results in an increased SNR at the input of the decider and, therefore, the SER and the MSE are improved.

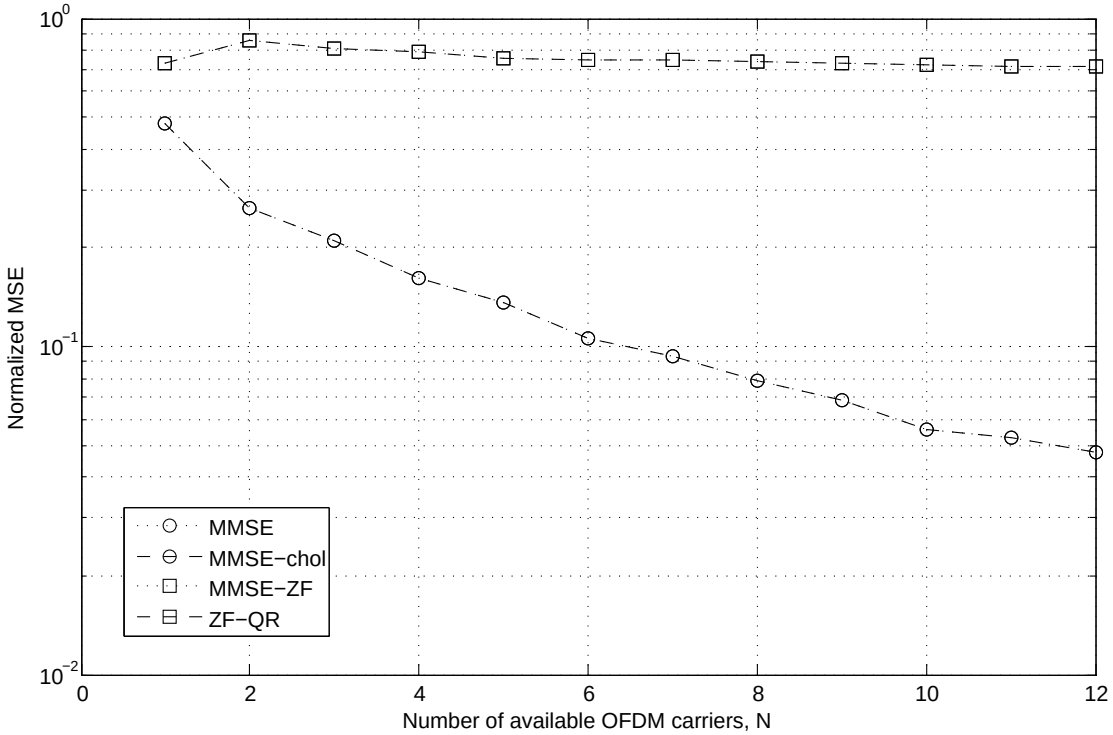
On the other hand, the ZF solutions fully equalize the system chain to the identity for which the equivalent channels $\boldsymbol{q}_n$ are all ones for all n. The persistent difference in performance (i.e. even when $N = 1$) between the ZF and the MMSE solutions is given by the fact that the MMSE receive filter has an extra degree of freedom to truly diminish the MSE, while the ZF receive filter has little to do. Despite of this, when the number of carriers is increased, the MMSE filters show a substantial gain with respect to the ZF, which do not improve with the frequency diversity. The reason comes out with the RAKE combiner, particularizing with $\boldsymbol{q}_n = \mathbf{1}^T$, since:

$$\tilde{\boldsymbol{s}} = \sum_{n=1}^{N} \mathrm{diag}(\boldsymbol{q}_n^H)\tilde{\boldsymbol{s}_n} = \sum_{n=1}^{N} \boldsymbol{I}\tilde{\boldsymbol{s}_n} = \sum_{n=1}^{N} \tilde{\boldsymbol{s}_n}.$$

Taking into account that $\tilde{\boldsymbol{s}_n}$ has the signal term and the noise term, they are both added without any kind of adaptation, and thus, the diversity technique fails. Interestingly, this is analogous to the spatial diversity explained in 12.3.2 above, since in a fully correlated scenario —that is, all the channel paths are the same—, the channel does not provide any kind of diversity.



(a)



(b)

Figure 12.5: SER (a) and NMSE (b) for the THP 2×2 MIMO channel with $\rho_R = \rho_T = 0.2$ correlation parameters and uniform noise versus the number of available OFDM carriers, N, at an average SNR = 0 dB.

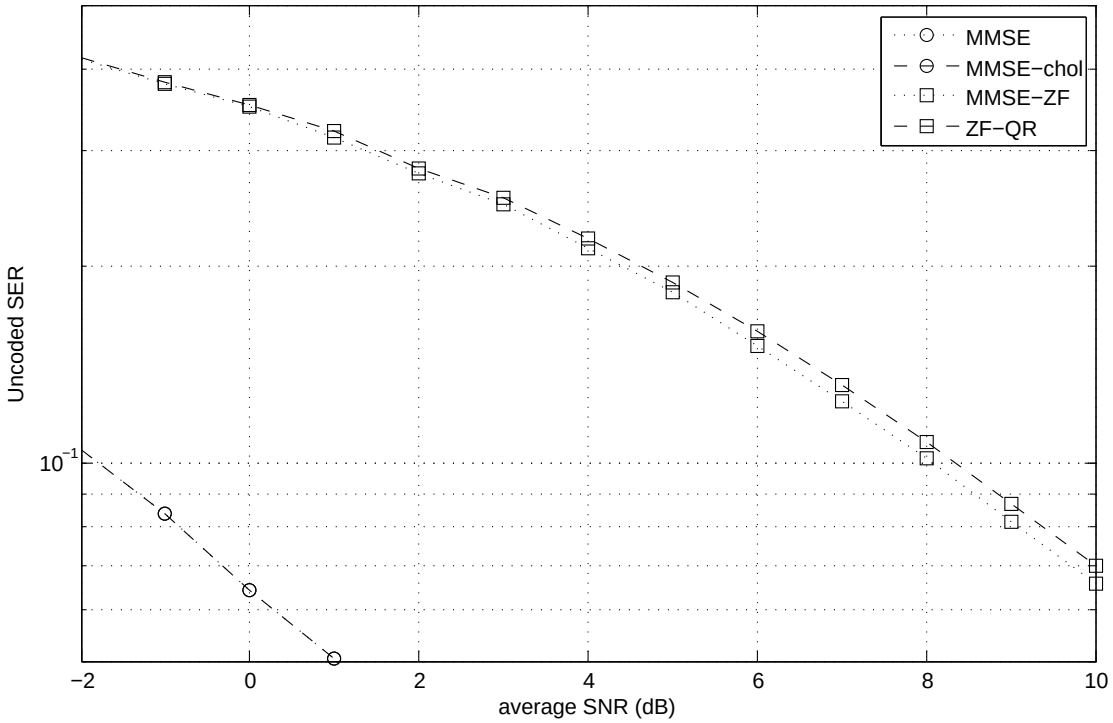
12.5 Noise Distribution

This last section of the THP MIMO-OFDM part considers the case where the distribution of the noise over the R receiving antennas and the N carriers is not uniformly distributed. This might be the general case, mainly in frequency rather than space, of real wide band systems since the noise existing at a certain carrier comes from several sources such as inherent thermal noise and other services and systems' interference.

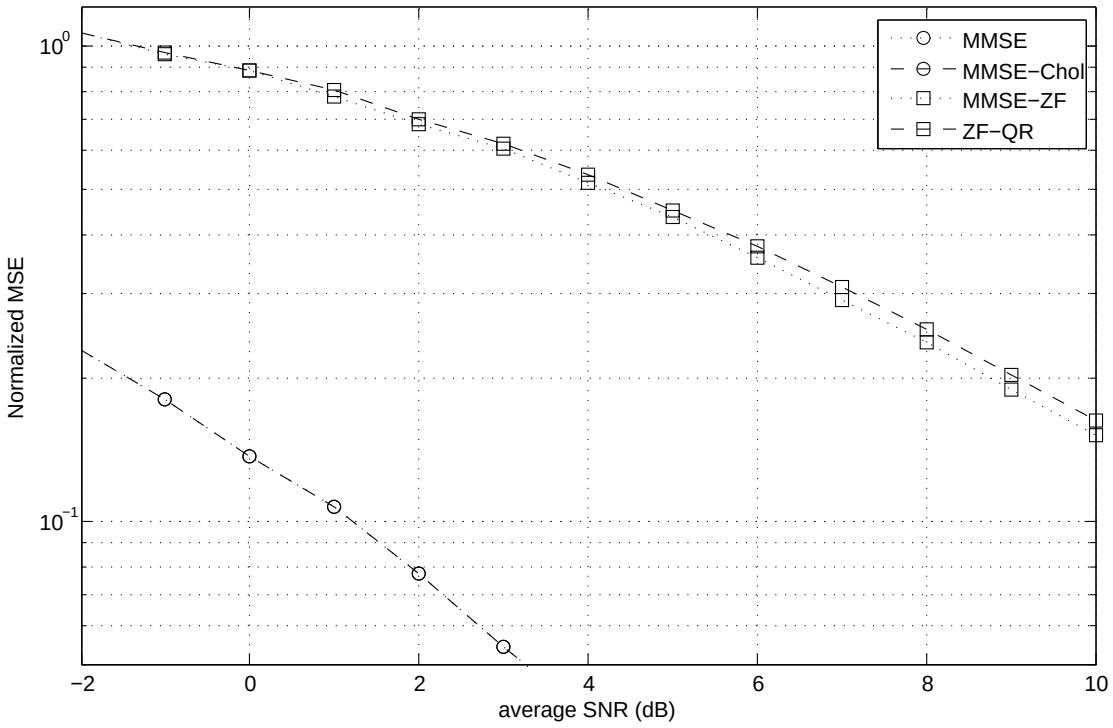
The simulation checks the system over randomly distributed noise and, as depicted in figure 12.6, only the ZF solutions differ from each other. The reason is that, though MMSE-ZF obeys the ZF condition, the receiver filter tries to minimize the residual MSE and takes into account the noise distribution in space —remember that OFDM splits the problem, so each of the n-th filters only see a flat-fading MIMO channel—, whereas the ZF-QR solution simply follows from the QR decomposition.

Yet the difference is laughable, both MMSE-ZF and ZF-QR coincide when the noise is uniformly distributed in space, for which the n-th noise correlation matrix becomes $\sigma^2_{n_n}\boldsymbol{I}$ and acts as an scaling factor in the optimization process, leading the two solutions be the same.

Finally, when the noise distribution is different in frequency, this affects in the power allocation process. However, the algorithm considers the noise and the channel gain/attenuation as a pack, whose overall resulting condition is reflected in the (i, n)-th eigenvalue.



(a)



(b)

Figure 12.6: SER (a) and NMSE (b) for the single-carrier THP 2×2 MIMO channel with $\rho_R = \rho_T = 0.2$ correlation parameters and colored noise versus average SNR.

Part IV

Spatial THP for the Multi-User MISO Frequency-Selective Broadcast Channel

Chapter 13

Introduction to Multi-User Communication

13.1 The Broadcast Channel, the Multiple-Access Channel and their Interferences

This final part attempts to study the Tomlinson-Harashima precoding (THP) technique for the scenario where one source simultaneously communicates to several receivers. Introduced in 1972 by Thomas M. Cover [Cov72], this so-called broadcast channel (BC) models many situations of past and recent wireless and wire line telecommunication systems such as broadcasting terrestrial digital television information from a transmitter to multiple receivers in a given local area, the transmission of user information through a bundle of digital subscriber lines (DSLs), the transmission of voice and data on the cellular wireless network, the data stream downlink channel of a WLAN or the broadcast information sent by the home automation top box to several home machines through powerline communication (PLC). In addition, the same channel seen in the uplink point of view models the so-called multiple-access channel (MAC) where a set of individual users, which act as transmitters, aim to access a single resource. What is more interesting is the recent apparent relationship between the BC and the MAC channels, which seem to claim duality properties from the information point of view.

The first evident question that arises in multi-user communication is how to separate the users, a concept that refers to multi-user multiplexing. The resource or dimension used to attain this problem might be an interesting parameter of design. Though traditional systems made use of time and frequency to overcome the multi-user multiplex —for instance the GSM cellular network—, modern systems employ more sophisticated techniques such as spread spectrum techniques —like code domain multiplexing (CDM) in UMTS or EDGE, or frequency hopping in Bluetooth— and orthogonal multitone techniques —orthogonal frequency division multiple access (OFDMA) in the Wimax standard. However, the need of signal processing is clearly justified when the separation is questioned due to channel conditions or when the transmission is done simultaneously to a set of users, who might share time, frequency and code.

With the last wave of multiple-input multiple-output (MIMO) techniques, the MIMO BC channel has very interesting promises as the capacity of the channel can be enhanced. It is the case when multiple data streams are transmitted simultaneously that a new type of interference appear in the

system: the multi-user interference (MUI) for the downlink BC and the multiple-access interference (MAI) for the uplink MAC. Though the BC and the MAC channels might be observed with very similitudes to the single-user MIMO channel, the simple fact that the receivers cannot cooperate turns the problem out to be a challenge, from both signal processing and information point of views.

A large body of knowledge is available on algorithms to face BC and MAC systems. First examples in the literature regarding cancellation of these types of interferences are, e.g., cross-talk cancellation for the DSL by Cioffi in [GC02] and MAI and MUI reduction for the wireless channel by Barbarossa and Scaglione in [BS00]. As seen in the previous part III, when single-user data streams are transmitted in a system with multiple transmitting antennas and multiple receiving antennas, the spatial multiplexing aims at maximizing the throughput of the system while diminishing the inter channel interference caused by these multiple streams. However, parallel transmission of independent data streams, inherent on BC and MAC channels, introduces such interference that cannot be abated with point-to-point MIMO techniques such as linear zero-forcing (ZF) or minimum mean-square error (MMSE) decision-feedback detection, developed for V-BLAST systems (e.g. [WFGV98]). Unfortunately, this is not feasible for the downlink multi-user MIMO channel, since conventional point-to-point signal processing techniques cannot work because of the absence of coordination among receivers of different independent mobile or decentralized users. An alternative solution is to place the preprocessing technique at the transmitter to eliminate the MUI at the receivers.

The key resides in the use of CSI, only fully known at the transmitter side, which makes conjunction with the concept of asymmetric signal processing, motivated by [GS99]. Though linear processing (both MMSE and ZF) can be designed for the multi-user downlink problem, since no receiver processing is needed, non-linear processing achieves better performance. Indeed, non-linear Tomlinson-Harashima Precoding is an attractive solution for a scenario where the transmission system employs multiple antennas at the transmitter and multiple users as receivers, in addition to the fact that it is one of the important techniques to achieve near MIMO channel capacity with reasonable complexity. In [WFH04], they show the out performance of the THP ZF system versus the corresponding linear ZF.

In this work, the THP system is optimized under the MMSE and ZF criteria for the multi-user downlink system aimed to diminish the MUI.

13.2 Preliminary Motivation to the MAC-BC Duality

A very powerful tool in multi-user systems is the apparent similarities between the multi-user downlink and uplink channels, namely the BC and MAC channels. This duality is seen from an information theory perspective in [VJG03]. The duality MAC-BC refers to the equivalence of channel conditions in both down and up links and, thanks to it, the resulting optimization signal processing done in one direction can be easily translated to the other, without the need of optimizing again, and vice versa. The key resides in that if a set of signal-to-noise ratios (SNRs) is achievable in the uplink, it is also achievable on the downlink under the same total power constraint for any linear and non-linear precoding systems.

This idea will be exploited along the present part, for solving concepts such as the multi-user power allocation and user ordering. Here we are interested in the downlink system but, in terms of system capacity, the uplink side is much easier to deal with and most of the optimization problems admit convex formulation, which therefore allow many already developed computationally efficient algorithms apply in this study. What is more, the MAC-BC duality will help us to obtain critical

interpretation of the results obtained for the BC.

13.3 Challenges of Multi-User Precoding

It should be noted that designing the THP for the BC is a different and much more challenging problem than the design of the MMSE or ZF non-linear joint transmitter-receiver processing for point-to-point MIMO system, because in multi-user MIMO channels, the mobiles are decentralized and uncoordinated, and hence additional constraints must be added to the problem. This section tries to summarize the accessory topics inherent to multi-user systems.

13.3.1 Asymmetric Signal Processing

The first consequence of the uncooperative scheme is the limited processing capability of the receivers. In the simple case where the mobile users own a single receive antenna, the receive filter is downed to a scalar because it cannot process the signals received by the others users, as they are decentralized and do not cooperate. The scalar condition is translated to a diagonal receive matrix in the vector transmission notation and this new restriction has to be taken into consideration in the optimization process. The first apparent consequence is that traditional decomposition techniques such as the singular value decomposition (SVD) are no more valid. This is the reason why recent literature regarding multi-user communication has welcomed more sophisticated techniques such as the introduced by the former parts of this work: Cholesky factorization (refer for the wireless channel to [KJUB05]).

Interestingly, this scheme makes conjunction with the asymmetric signal processing philosophy, motivated by [GS99]: while the most signal processing is moved to the transmitter —which has true complete CSI—, the mobile terminals are kept as simplest as possible. First introduced by [JU05], all the receivers perform the same scalar automatic gain control (AGC), $g \in \mathbb{R}_+$, obtaining a **solution where the receiver has very low complexity and, thus, low cost and low consumption**; extended in [HJU05], users might have different gain factor $g_k \in \mathbb{R}_+$ though there is no apparent gain in performance; as well in [MS07] the receiver acts as accessing technique with a single gaining factor. In general, this gaining factor used in literature is well adopted and, in fact irrelevant, because it conserves the individual SNRs, and it preserves the noise statistical whiteness.

13.3.2 Multi-User Ethics

In both single-input single-output (SISO) and MIMO single-user systems, all the resources where managed to offer a certain quality of service (QoS) to the same user. In the MIMO case it has been seen that the spatial dimensions can be used as diversity technique to improve the symbol error rate (SER) or SNR, or as throughput technique, enhancing the system final rate. Now, the multi-user scenario requires the system to manage all the resources and do its best to overcome the users' needs. As a consequence, we shall make use of the theory of rates and capacities from the information theory to measure the use of the system by each user. This quantification is done with the achievable rates for each user, $\{R_k\}_{1\leq k\leq K}$, at which the users are able to decode the information from the transmitter for a given transmitted power constraint at the receiver. As motivated in part I of this thesis, the THP is considered as an implementation of dirty paper coding with successive user encoding. For a given encoding order, $\mathcal{O} = \{\pi(1), \ldots, \pi(K)\}$ —i.e., π is a mapping function $j \mapsto \pi(k) = k$ so that user k is precoded at the j-th position—, the achievable rate of the $\pi(k)$-th user is given by the achievable MIMO rate considering that the already precoded users do not interfere the given user by the dirty

paper coding (DPC) theorem [Cos83] ([CS03, YC04] are very good recent references):

$$R_{\pi(k)}^{BC} = \log \frac{\left| \sum_{j \geq k} \boldsymbol{H}_{\pi(k)} \boldsymbol{\Sigma}_{\pi(j)} \boldsymbol{H}_{\pi(k)}^{H} + \boldsymbol{R}_{\boldsymbol{n}_{\pi(k)}} \right|}{\left| \sum_{j > k} \boldsymbol{H}_{\pi(k)} \boldsymbol{\Sigma}_{\pi(j)} \boldsymbol{H}_{\pi(k)}^{H} + \boldsymbol{R}_{\boldsymbol{n}_{\pi(k)}} \right|}, \tag{13.1}$$

for $1 \leq k \leq K$, and where $\boldsymbol{H}_{\pi(k)}$ denotes the downlink channel of the user ordered at the k-th position, $\boldsymbol{\Sigma}_{\pi(k)}$ is the covariance matrix of the power allocated for user $\pi(k)$ and $\boldsymbol{R}_{\boldsymbol{n}_{\pi(k)}}$ is the noise autocorrelation matrix of the same user. For the Gaussian K-users channel, the capacity region is a convex region in the K-dimensional space and it reflects the trade-off between the individual data rates of the different users in the system competing for the limited resources. Of course, expression (13.1) holds if all channel inputs and all channel outputs are Gaussian. This can be seen from two perspectives: with a large number of users the interference is also Gaussian thanks to the central limit theorem, or the expression is seen as a bound and, since we are interested in optimizing (maximize), it applies. However, as motivated in part I, both DPC and THP user coset coding methods, i.e., each information symbol is represented by a large set of codewords (DPC) or signals (THP). The codeword or signal actually transmitted depends on the interference encountered as well as the symbol to be transmitted. Nonetheless, THP incurs a finite loss —studied in this work for the SISO case in section 7.2— which, together with the use of uncoded QPSK —not a capacity-achieving modulation— are modeled through the so-called *gap* [Cio02]. This factor depends on the modulation and the target SER and it has been studied for several modulation and communication schemes (e.g. [GA06]). However, it can be shown (e.g. [FYL07]) that this gap, which is in signal-to-interference plus noise ratio (SINR) terms, can be translated to a rate gap so that the true achievable rate due to the use of THP is the capacity theoretical limit minus a fixed term, i.e., $R_{\pi(k)}^{BC} - \log \Gamma$, with Γ the gap. This factor is relevant when designing the system for a given set of requirements, but in our work, the maximization process is unaltered and, therefore, neglected.

Let $\{\alpha_k\}_{1 \leq k \leq K}$ be the weights associated to users which reflect the priority of each user, with $\sum_k \alpha_k = 1$. The boundary of the capacity is given by the user rates that maximize the linear combination according to the well-defined weights:

$$R = \sum_{k=1}^{K} \alpha_k R_k.$$

It is worthy to highlight that the distribution of the weights does not bring any information on the distribution of the achievable rates. For instance, if the same priority is assigned to all the users, i.e. $\alpha_k = 1/K$ for all k, a given user with a poor channel conditions could obtain a lower data rate than the other users, or even a zero data rate. Practical characterizations are:

- The *single-user rates*, R_k^1 are the maximum rates in the single-user communication case, that is, setting only one weight different from zero. In this case, all the transmitted power is allocated to the transmitting user and, as the others do not transmit, there is no interference at all. It corresponds to time domain multiplexing (TDM) and the key point would be in designing the time slot sharing in the system. Nevertheless, the single-user rates are an indicator of the achievable rates of each user and, therefore, of each user's channel capacity. In other words, if the system ensures a certain rate R_k to the k-th user, the ratio R_k/R_k^1 represents the relative cost that the system pays due to other users sharing the resources, a parameter which is more representative than the absolute rates.

- The *maximum sum-rate* corresponds to setting the weights so the linear combination is maximum for a given user rates. This is easily obtained from optimization theory as we maximize $\sum_k \alpha_k R_k$ subject to $\sum_k \alpha_k = 1$. Differentiating with respect to α_j and using the Lagrange duality, it becomes $\alpha_k = 1/K \ \forall k$; so the system is designed to maximize $1/K \sum_k R_k$ or, equivalently, $\sum_k R_k$. It has been widely used in literature —it has been conceptually and deeply explored by [CS03] and [YC04] in general; and exploited by [TUBN05] for the multi-user MIMO orthogonal frequency division multiplexing (MIMO-OFDM) channel and by [MW05] for the MIMO-OFDM-multiplexed users BC, both based on THP—, but might result in unfair situations where the users with best channel conditions have a much higher rate than the others. However, this scheme is very powerful as the maximization of the arithmetic sum of the user rates is, by definition, the BC capacity (e.g. [CS03, YC04, VJG03]):

$$\mathcal{C}^{BC}(\boldsymbol{H}_1, \ldots, \boldsymbol{H}_K, \mathcal{E}) = \max_{\boldsymbol{\Sigma}_k : \sum_k \mathrm{tr}(\boldsymbol{\Sigma}_k) \leq \mathcal{E}} \sum_k R_k \tag{13.2}$$

  and, hence, it allows to make use of the MAC-BC duality as an interesting tool. Here we introduced the parameter $\mathcal{E}$ as the transmitted energy constraint or budget. The maximization, as it is appreciated, is done over the set of downlink covariance matrices, each of one is an $T \times T$ positive semi-definite matrix.

- The *maximum common rate*, which is obtained by setting $R_1 = \ldots = R_K$. This approach under performs when the single-user rates are very diverse, as the common rate is a waist of resources because if forces the users with the best channels to lower their rate dramatically to reach the level of the weakest channels.

- When establishing a *set of rate requirements*, denoted by $\{\varphi_1, \ldots, \varphi_K\}$, the system is optimized to achieve $R_k = \varphi_k$ for each user. This structure assume that there is enough available power at the transmitter side and that the optimization process equals to find the minimum power to achieve the rate requirement (e.g. [FYL07]). With sum power constraint, this scheme might not be possible —maybe only a small number of users are served— and some other strategies are considered, like

- the *maximum balanced rate*, first explored in [SVL05], is such that

$$\frac{R_1}{\varphi_1} = \ldots = \frac{R_K}{\varphi_K}, \tag{13.3}$$

  where φ_k represents the target rate for user k, and it is usually set to $\varphi_k = R_k^1$. This means that all the users transmit at a rate proportional to the single-user rate offered by their own channel. As a result, the relative cost implied by the coexistence with other users in the same system is the same for all users. Though it seems to be the most fair situation, it has only been analyzed in [SVL05] for the scalar memoryless BC providing water-filling algorithm, and [KLV07] for multi-tone systems, providing loading for the two users case. Obviously, the balanced capacity concept should be applied in systems where individual rate requirements are crucial.

- The *product of rates* considers the maximization of $\Pi_k R_k^{\tilde{\alpha}_k}$ where now the priorities are normalized to $\Pi_k \tilde{\alpha}_k = 1$. This case ensures that if a user is unfairly treated, it is reflected with more weight in the target function and, as a result, the system will tend to treat all the users more equally, taking into account their priorities.

The key point when optimizing the system filters is determining the objective function that best reflects the user requirements. Firstly, the QoS of each user is a qualitative parameter that depends on the SER, the data rate, the SNR, and the type of service that the user is using at a certain moment, which is indeed a very difficult task. However, quantitatively speaking, the mean-square error (MSE) or the SNR are eligible parameters to be optimized as they are very nice related to the channel capacity through the result presented by Shamai and Verdú in [GSV05]:

$$\frac{\partial}{\partial \mathrm{SNR}} I(\mathrm{SNR}) = \mathrm{MSE}(\mathrm{SNR}), \tag{13.4}$$

where I stands for the mutual information between the input and the output of the channel. On the one hand, since I is a measure of the amount of information that can be obtained from the output of the channel —notice that the maximum rate is defined as the maximum mutual information I— and the MSE is a monotonically decreasing function of the SNR, the maximization of the former and the minimization of the later are equivalent problems. On the other hand, the study can be focused on the MSE as its minimization is also the minimization of the SER and maximization of the SNR. Referring to THP literature for multi-user system, authors adopt several approaches:

- Addressing the non-linear precoding for MIMO multi-user downlink transmission with different QoS requirements in nowadays telecommunication applications by modeling the user requirement as an individual SNR, SNR_k, is very desirable in multimedia applications, where several types of transmission with different reliability constraints must be supported. On this line, Morelli [SM07], Lim [FYL07] and Tellambura [FTK07a] recently have designed the Tomlinson-Harashima optimal filters by maximizing the individual SNRs and rates with an overall transmitted power constraint.

- Another approach is to optimize the system to minimize the SER for each user, SER_k (e.g. [BF08]).

- Finally, the most typical approach —followed mainly by Utschick in, e.g., [JBU04] and Liu [LK05]— is to consider the weighted sum of MSE as target function to minimize. The set of weights $\alpha_1 \dots \alpha_k$ try model the QoS requirement of each user:

$$\mathrm{MSE} = \sum_{k=1}^{K} \alpha_k \mathrm{MSE}_k,$$

  where the k-th alpha reflects the priority given to the k-th user, as in the information theory results.

13.3.3 Multiplexing Technique

Another issue of interest is how the multiplex for BC and multiple access for MAC is done. Interestingly, we can exploit many communication-like dimensions to separate the users. In systems where the orthogonality is ensured —to be cited TDM, frequency domain multiplexing (FDM), OFDM, CDM with multi-carrier or single-carrier approaches,...— the use of a precoding technique is not relevant. However, user precoding might be needed if the orthogonality is not ensured, usually due to inexactitudes when estimating the channel or, in addition, when one or more of those dimensions is used to serve user in a simultaneously manner.

Referring to THP, in [MW05] an OFDM modulation is used so as each user transmits at one single carrier; German writers in [KJUB05] and [JUN05] for spatial domain multiplexing (SDM); Joham and others in [JBU04] for spatio-temporal precoding; Morelli at [MS07] exploits a unified framework for OFDMA —that is, each user is assigned a certain group of OFDM carriers, not only one— and MC-CDMA —users' data are spread over all the OFDM carriers using orthogonal codes— multiplex for the downlink transmission; Letaief [ZL07] perform THP precoding at the frequency domain for single-carrier spatially multiplexed users; [FTK07a] use orthogonal space-time block coding OFDM —that is, diversity is attained by transmitting a given number of symbols over some time slots, applied to OFDM at sub carrier level...—; and so, a direct-sequence CDMA is considered by [WFVH04].

13.3.4 User Scheduling, Power Allocation and User Ordering

The power allocation in the multi-user systems is a much more complicated issue than the single-user case. In fact, it involves three very closed topics, which are detailed after the following itemize.

- User scheduling is the management of serving users in time, last resort when there are not enough resources,
- the power allocation refers to how the transmitted energy is placed on the available resources to serve the users simultaneously, and
- user ordering deals with how users are ordered in the precoding process.

Firstly, on the one hand, as highlighted in [UL07], if the number of users exceeds the number of transmitting antennas, such systems requires user scheduling, a part from user ordering. If $K \gg T$ and users are allowed to be carefully chosen such they are approximately orthogonal to each other, a simple channel inversion would then suffice to achieve good performance —for instance, using linear precoding. However, if not $K \gg T$ or the users are served randomly, such as round-robin scheduling, a more sophisticated techniques are needed to ensure the fairness among the users, such as VBLAST (e.g. [WFH04]). The last case in which $K \leq T$, the issue of user scheduling does not apply since the users are served simultaneously at all time.

Secondly, the **power allocation for multi-user systems is a complex issue and it has a very different connotation compared to the single-user (MIMO) systems**. The single-user MIMO system allocates the power to obtain the best overall performance as there is only one user. This means that some symbols may be treated with more care by the system at the penalty of other symbols; situation that is not accepted in the multi-user system as the penalty of other symbols can be the penalty of other users. For this reason, several water-filling algorithms has been encountered and studied under an information theory point of view by Cioffi and Goldsmith for different cases to solve the problem of power allocation in multi-user systems. In all of them, power is allocated so as to maximize the so-called sum capacity or sum of data rates of the users. It deserves attention to highlight [GC02] for the multi-user water-filling allocation algorithms by maximizing the weighted sum of the data rates of all the users in the DSL channel, and [YC04], [JRV$^+$05] for multi-user water-filling allocation by maximizing the sum capacity of the multi-user MIMO channels with power constraint. In [YC04], Yu and Cioffi treat the iterative water-filling algorithm for the conventional MIMO MAC problem, with individual power constraints for each user. This algorithm can also be applied to the sum power MIMO MAC problem, but is however, inefficient, since it requires a search for covariances over all power splits among the K users in the system. In [JRV$^+$05], A. Goldsmith finds the optimal iterative power allocation algorithm for the BC making use of the results in [YC04] and the proper

transformations thanks to the duality property between the MAC and the BC to obtain the downlink power allocation.

Thirdly, German authors in [JU05] and sequels propose to order the users before the precoding loop with a permutation matrix. In classical MIMO systems, the topic of ordering the signal before the processing is done by considering the branches with higher SNR first (e.g. [FGVW99]). However, when dealing with feedback techniques, such as decision feedback equalization (DFE) or THP, it is not as simpler. On the one hand, for the uplink DFE or V-BLAST techniques, the user ordering in the detection process is optimal under the *best-first* paradigm. Let us assume that the first user, u_1 is detected first and then u_2. Clearly, there is one interference for the detection of the first user as the second user has not been detected yet and it is an unknown interference; while the second detection is interference free since u_1 has been detected. With absence of error propagation, the first detected user needs to have more robust channel conditions than the others. As a consequence, in systems where the detection is done iteratively, a best-first techniques is employed. On the other hand, since THP performs the precoding at the transmitter side, by means of the MAC-BC duality, the encoding order in the downlink should be the inverse of the decoding order in the uplink; that is, *best-last*. More specifically, we assume that u_1 is precoded first, and then u_2. When precoding the first user, the second user is treated as a potential interferer to u_1 because we do not know the coded version of u_2 yet. However, when the first user is already precoded, it is fully known to the transmitter and it is feedback subtracted to u_2, which is interference free. As this is the case of interest in this work, generally, the user ordered at position k, $u_{\pi(k)}$, sees the set of users $\{u_{\pi(k+1)} \dots u_{\pi(K)}\}$ as interference, while the already precoded set of users, $\{u_{\pi(1)} \dots \{u_{\pi(k-1)}\}$ do not interfere. Under this considerations, it is shown (e.g. [KJUB07], [FYL07], [LK05]) that the optimal user ordering is achieved with best-last paradigm. The main motivation, more extended in the next sections, is that the user with worst channel conditions deserves more power so that, accordingly, the one who deserves more energy is the one that interferes the less, i.e., the first one.

The job behind all the explained above is which parameter best expresses the channel conditions and user priority. In the point-to-point communication parts II and III, the performance parameter user in the design of the system has been the parameter reflecting the channel gain/attenuation and the noise power:

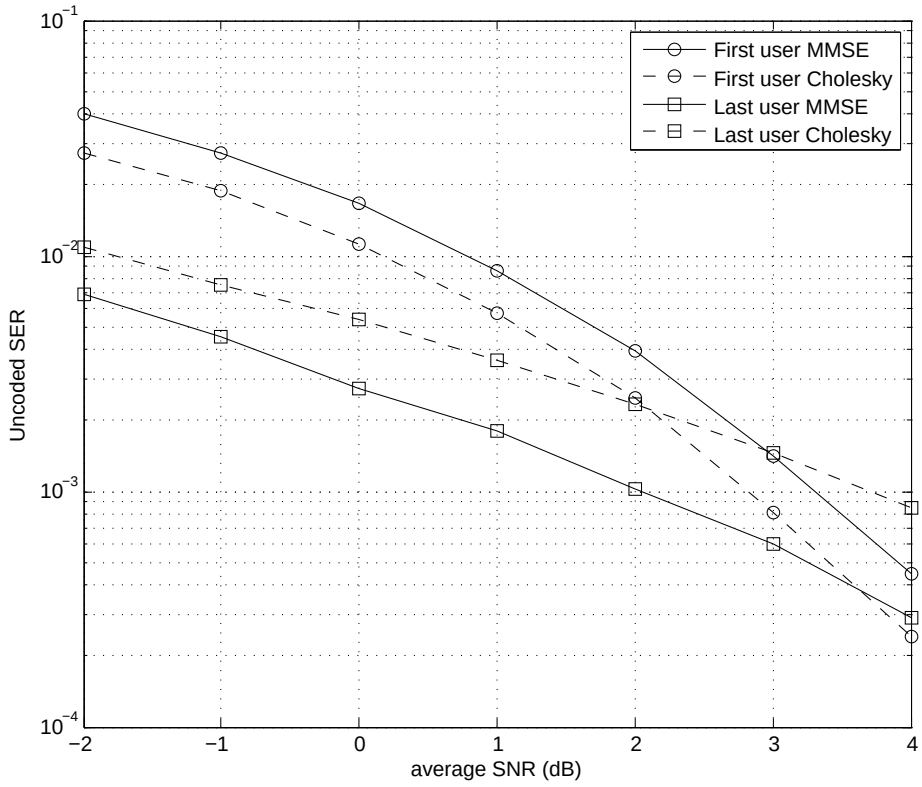
$$\boldsymbol{H}\boldsymbol{R}_{\boldsymbol{n}}\boldsymbol{H}^{H},$$

so in the multi-user system something similar must appear, with the corresponding user priority, if applies.

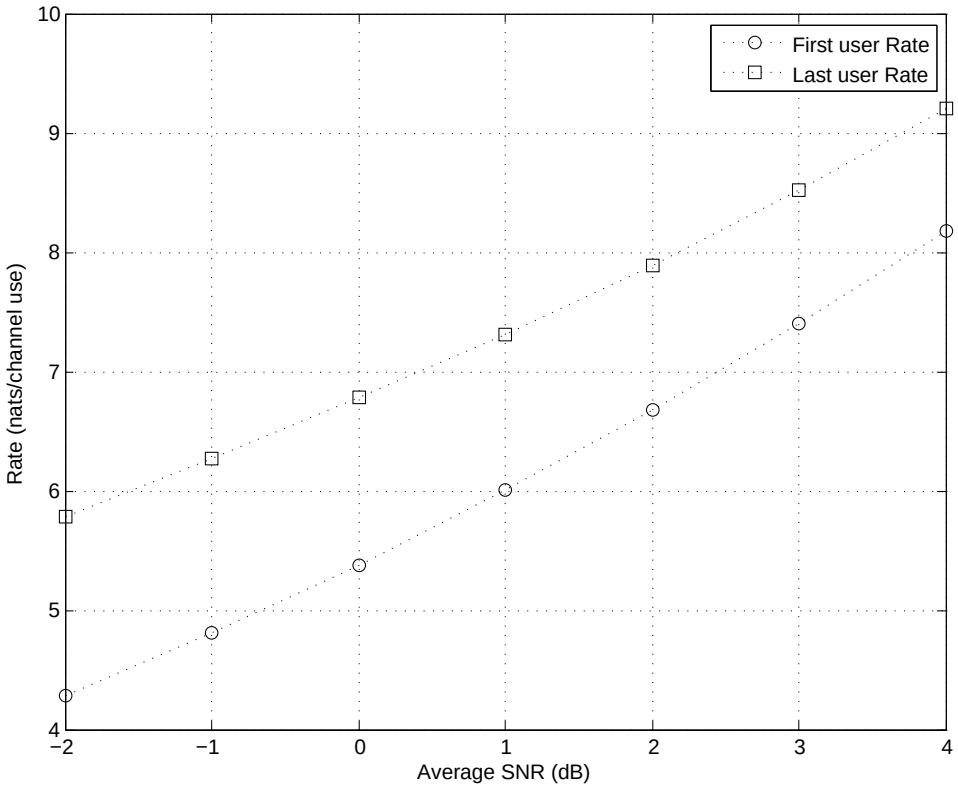
13.4 Objectives Outline

In this work, as a continuation to the single-user MIMO-OFDM part III, a multi-user system is proposed where users are spatially multiplexed onto an OFDM where each user receives data streams though a set of carriers, which might be used to receive coded information or, analogous to the point-to-point MIMO-OFDM, the same symbol is transmitted at all the carriers with the aim of enhance the system's diversity. At its turn, the power allocation must be done from an information theory point of view by means of the user rates, while the optimal user ordering is employed as explained in the former chapter under a best-last paradigm.

The optimal transmit filters are optimally obtained under the MMSE criterion without and with the use of Cholesky factorization with a transmit power constraint, whereas it is shown that the



(a)



(b)

Figure 13.1: SER (a) and user rates (b) for the $K = 2$ users $N = 4$ carriers THP 2×1 MISO BC versus average SNR, for each user, with diverse channel conditions $2\xi_{n,1} = \xi_{n,2}$.

MMSE-ZF is equivalent to the pure MMSE and that the QR decomposition is not feasible in a receive uncooperative system. The receive filters will be downed to a single scalar factor at each carrier which can be regarded as an AGC. The performance of the THP for the multiple-input single-output (MISO) OFDM frequency-selective BC with $K = 2$ active users sharing $N = 4$ carriers is depicted in figure 13.1 in terms of the achieved SER and user rates. On the one hand, the structure of the feedback filter is very similar to the single-user SISO vector and point-to-point MIMO channels, but, on the other hand, the transmit forward filter $\boldsymbol{F}$, which performs power allocation —notice that we want to transmit information from K users through a transmitter equipped with $T \geq K$ transmitting antennas—, deserves much attention to be cited in this introductory chapter. Specifically, the k-th column of the power allocation filter at the n-carrier is nothing more that the set of weights in which the precoded information from user $\pi_n(k)$ –that is, the user precoded in the k-th position at the n-th carrier— is placed to the T transmitting antennas. It is obtained that

$$\boldsymbol{f}_{n,\pi_n(k)} = \frac{A_{n,\pi_n(k)} q_{n,\pi_n(k)}}{\sigma_v^2} \boldsymbol{B}_{n,\pi_n(k)}^{-1/2} \boldsymbol{U}_{n,\pi_n(k)}.$$

Though the meaning of each term can be found in detail in section 15.2, it is worthy to say that the power that is allocated to a given user is proportional to the interference it will receive —this is given by the parameter $A_{n,\pi_n(k)}$— and inversely proportional to the interference it will cause —reflected by $q_{n,\pi_n(k)}$ in the dual problem—, while the information is properly rotated —by matrices $\boldsymbol{B}_{n,\pi_n(k)}$ and $\boldsymbol{U}_{n,\pi_n(k)}$— to direct the signal in a way to maximize the use of its channel but also to minimize the effect it will have to the rest of users. All this considerations result in the fact that the first precoded user is the one that deserves more power because it is the one that receives more interferences and, additionally, the one that interferes the less. As a result, as seen in figure 13.1, the first precoded user has much better performance in terms of SER and MSE at high SNR compared to the last precoded user. An interesting aspect to highlight is the fact that the SER and MSE curves of the users cross each other at a given SNR point. This creates, somehow, two working regions, whose separation point if function of many parameters of design.

Chapter 14

The MISO-OFDM Frequency-Selective BC

14.1 Preliminaries and General Assumptions

In general, a BC or MAC system with K users is formed by K $T \times R_k$ MIMO channels, with N orthogonal carriers, C codes and S available time-slots, which are assigned to the K users. This is a very general model, and it can be diminished by considering a general MIMO matrix, which can be obtained from a spatio-temporal multiplex, OFDM multiplex or spread spectrum techniques or any combination of them. With this purpose, in this work we consider spatially multiplexed data streams, without spreading codes, with one receiving antenna for each user so that they time-share N orthogonal frequencies. As a result, in the sequel, we refer to the multi-user MISO-OFDM BC.

When setting $R_k = 1$ for each user, the MIMO inter-channel interference (ICI) disappears because the transmitter will only send one symbol per user. As a consequence, the THP is in charge of reducing only the MUI or MAI interferences. On the one hand, setting the number of receiving antennas to one for all the receivers has considerable practice importance, since many portable devices have room for only one antenna. As well, many devices have a matched filter to combine multiple diversity signals so that the received signal is always reduced to a scalar decision, whose scenario is mathematically equivalent to the single receive antenna one. On the other hand, motivated above by the capacity-region duality in one-antenna receivers, it can be shown that the achievable rates are monotonically related to the SINR, so makes the information theory approach when determining the user ordering a nicely justified solution, since the monotonically relation to the capacity region is achieved in parallel with the MSE and the BER in most modulations.

14.2 Multi-User Multi-Carrier Channel Notation

The multi-user MISO channel is the means though which the information is sent from the transmitter to the users. Each one of the topic involving resources (user multiplexing, the trade-off between throughput and diversity) are, therefore, intrinsically linked. Luckily, the vectored transmission and the matrix notation used along this work for both single-user and multi-user scenarios comes out to be a very unified and general methodology for treating the channel, whatever the multiple-inputs and outputs mean, or which the multiplexing technique is. Mathematically speaking, the amount of available resources that the channel offers is given by the number of eigen channels, which is the rank

of the channel matrix $\boldsymbol{H}$.

In our particular system ($K \leq T$), K one-receiving antenna equipped users communicate to a transmitter equipped with T transmitting antennas along N orthogonal carriers. This means that the channel matrix is a block diagonal matrix formed by N $K_n \times T$ MIMO channels,

$$\mathcal{H} = \text{diag}(\boldsymbol{H}_1, \ldots, \boldsymbol{H}_N)\ 1 \leq n \leq N, \tag{14.1}$$

where $K_n \leq K$ is the number of users allocated in the n-th band, and each $K_n \times T$ MIMO channel represent the channel response from the transmitter to the K_n one-receiving antenna users. That is,

$$\boldsymbol{H}_n \triangleq \begin{bmatrix} \boldsymbol{h}_{n,1} \\ \vdots \\ \boldsymbol{h}_{n,K_n} \end{bmatrix}, \tag{14.2}$$

where the $T \times 1$ MISO (vector) channel $\boldsymbol{h}_{n,k}$ is formed by the T channel frequency responses from each antenna to the user k at the n-th carrier. The rank of $\mathcal{H}$ is, at most, NM; where $M = \sum_n M_n$ and, as in the single-user case, $M_n \triangleq \text{rank}(\boldsymbol{H}_n)$. The rank of the n-th channel matrix will be the number of active users, as $\text{rank}(\boldsymbol{H}_n) = \min\{T, K_n\}$ and, since they do not cooperate, there is no correlation among them for which with very high probability $M_n = K\ \forall n$. On the other hand, however, if the channel shows a large level of attenuation at a certain frequency, all the users will probably see this bad channel conditions, independently on the spatial correlation. For this reason, each user transmits over the N bands, reducing notably the likeliness of detecting errors with the use of codes or frequency diversity.

This last paragraph is intended to argue the issue of user ordering related to notation. From (14.2), we assume that user k receives a linear combination of the signals transmitted by the T antennas through the k-th row of the channel matrix. This information is irrelevant for the receivers, as they only see their own channel. From the transmitter point of view, it needs to know which row corresponds to the k-th user to place the proper information on the transmitting antennas. When the system performs user ordering, one might think that the channel matrix needs to be reordered, but this is not true, as it is adrift to the users and to the information theory capacity expressions, as they are given by the encoding order, which is exclusively known to the transmitter. In other words, we assume that users are listed from 1 to K and, before the precoding loop, are ordered through a permutation matrix at each carrier, $\boldsymbol{P}_n$, which performs the permutation π_n of order $\{\pi_n(1) \ldots \pi_n(K_n)\}$. After the precoding, the order should be restored to the original list to perform the power allocation, transmission over the channel and reception. In this sense, though a control channel is needed to inform the users about the dynamic resource management due to the user allocation —this control channel can be a different common channel or a logical channel inside the data—, this scheme does not require the transmitter to communicate the precoder order to the receivers, as they will receive a linear combination of the transmitted precoded signal according to their MISO channel response.

Figure 14.1 shows the system model of the MISO BC at the n-th carrier. The OFDM aspect of the system is not treated here as it is completely equivalent to the point-to-point MIMO system (cf. 9.1). We denote $\boldsymbol{x}_n \in \mathbb{C}^{T \times 1}$ as the signal containing the T symbols to be placed at the transmitting antennas and at the n-th carrier, $x_{n,i}\ 1 \leq i \leq T$, $\boldsymbol{H}_n$ as defined in (14.2), $n_{n,k}$ is the scalar white Gaussian noise with variance $\sigma^2_{n_{n,k}}$ and $y_{n,k}$ is the received signal. For convenience, the noise and received signals may be wrapped onto two $K_n \times 1$ vectors at the n-th carrier, denoted by $\boldsymbol{n}_n$ and $\boldsymbol{y}_n$

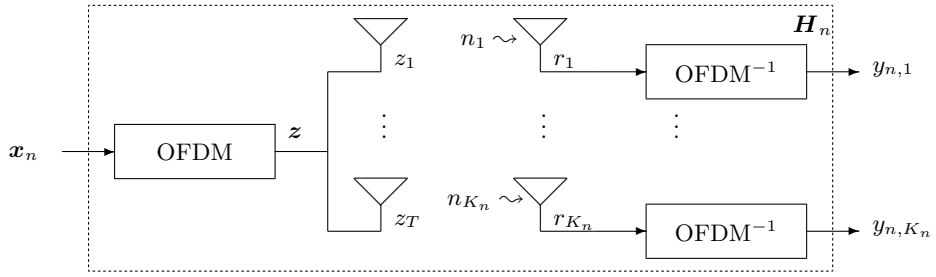


Figure 14.1: MISO-OFDM BC channel

respectively. As a result, the MISO-OFDM BC model reads

$$\boldsymbol{y}_n = \boldsymbol{H}_n \boldsymbol{x}_n + \boldsymbol{n}_n,$$

or, the signal received by the k-th user is, from (14.2):

$$y_{n,k} = \boldsymbol{h}_{n,k} \boldsymbol{x}_n + n_{n,k}.$$

As it will be seen, (14.2) is more akin rather than the wrapped version as it will help to better understand signal processing topics such as the power allocation and the filter design. A first consideration that justifies this asseveration is that the diversity model explained in section 9.2 has to applied individually to the K_n MISO channels. Concretely, we recall the expression (9.5) for the MISO case in which the receive correlation is an scalar —without loss of generality is set to one—:

$$\boldsymbol{h}_{n,k} = \boldsymbol{h}_{\omega,k} \boldsymbol{R}_T^{1/2},$$

where now $\boldsymbol{h}_{\omega,k}$ is a vector of i.i.d. complex Gaussian variables whose variance represent the gain or attenuation at the n-th band for the k-th user. Here, $\boldsymbol{R}_T^{1/2}$ is obtained following an AR model of order one as well and the transmitting antennas, with the proper transmitting antennas placement [Stü01], offer enough spatial diversity to separate the users.

14.3 Signal Processing

Without loss of generality, we concentrate the study at a given carrier as the **THP structure aims to reduce the interference caused amongst K_n users that actively and simultaneously use the downlink channel**, with power allocating along all the carriers. With the use of OFDM, the issues discussed for the single-user MIMO channel in section 9.3 apply also here, with the additional topic of *user capability*. The matrix channel model allows us to consider $\boldsymbol{H}$ as a set of resources through which the system can use to transmit information. In normal conditions, the rank of $\boldsymbol{H}$ is KN and therefore, keeping in mind that by design $K \leq T$, its maximum lays on TN, which means that our system is capable to cope with TN "informations" per channel use. The management of these resources —let them call bins— is up to the system depending on what is required. In other words, the system can be configured with one of theses strategies:

Maximum Diversity It corresponds to place the same symbol onto the KN bins, with K active users. If $K = 1$ (single-user case), the multiple-output nature of the channel acts as spatial diversity for the single-user, while with $K = T$ (maximum allowed) the spatial diversity is

required to separate the users.

Maximum Throughput This configuration places as much different symbols onto the bins. It would correspond to place N different symbols at the N carriers, as the users only receive one symbol and the spatial dimension cannot be used to increase the rate (but so the diversity does).

Maximum User Capability This last perspective is introduced to highlight the fact that the system can cope with TN different users. However, this approach reduces the diversity and the throughput to one symbol (no diversity) per channel use.

In this work, as the THP is put under study to analyze its capacity in reducing the MUI, the best configuration that reflects this with real results is having $K \leq T$ users in the system, all of them using all the N carriers to place the same symbol in order to enhance the diversity.

The signal model is analogous, again, to the single-user SISO and MIMO channels in parts II and III. The processing is done separately on the N carriers, each one with K_n active users. The full system proposed in this chapter is detailed in figures 14.2 and 14.3, which depict the transmitter and the k-th receiver respectively.

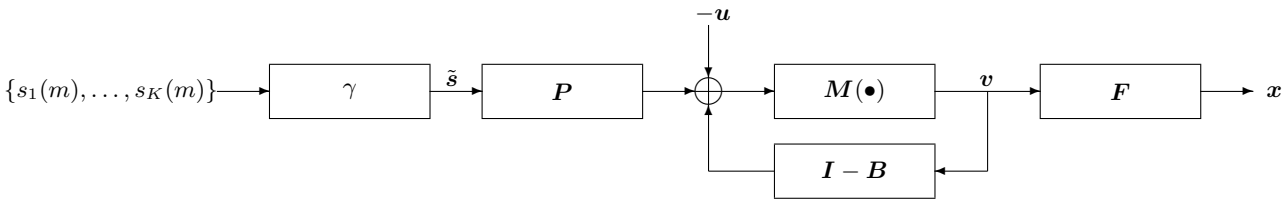


Figure 14.2: MISO-OFDM BC Tomlinson-Harashima precoder

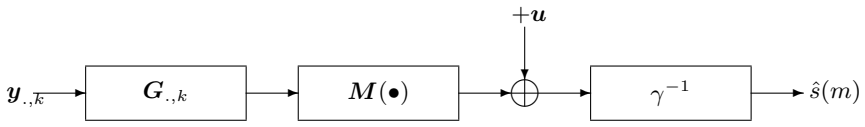


Figure 14.3: MISO-OFDM BC Tomlinson-Harashima k-th receiver

The modulo operation can be treated linearly using the auxiliary signal at each carrier $\boldsymbol{a} = [\boldsymbol{a}_1^T \ldots \boldsymbol{a}_N^T]^T \in \mathbb{C}^{\sum_n K_n}$ and the proper dithering signal, $\boldsymbol{u} = [\boldsymbol{u}_1^T \ldots \boldsymbol{u}_N^T]^T \in \mathbb{C}^{\sum_n K_n}$, will be used at both sides as justified in 5.2.3. Therefore, defining $\boldsymbol{d} \triangleq \tilde{\boldsymbol{s}} - \boldsymbol{a} - \boldsymbol{u}$, the overall equation reads

$$\hat{\boldsymbol{d}} = \boldsymbol{G}\mathcal{H}\boldsymbol{F}\boldsymbol{B}^{-1}\boldsymbol{P}\boldsymbol{d} + \boldsymbol{G}\boldsymbol{n}, \tag{14.3}$$

where $\boldsymbol{d}$ contains the equivalent signal of the K users along the N carriers. As the receivers have only one antenna, they simply perform an scalar weight, which we allow to be the same among the users but different at each carrier. As a result, from figure 14.3, $\boldsymbol{G}_{.,k} = \text{diag}(g_1, \ldots, g_N)$ for all k and, by construction of the formal expression (14.3):

$$\begin{aligned}
\boldsymbol{B} &= \text{diag}(\boldsymbol{B}_1 \ldots \boldsymbol{B}_N) \in \mathbb{C}^{\sum_n K_n \times \sum_n K_n} \\
\boldsymbol{F} &= \text{diag}(\boldsymbol{F}_1 \ldots \boldsymbol{F}_N) \in \mathbb{C}^{NT \times \sum_n K_n} \\
\boldsymbol{G} &= \text{diag}(g_1 \boldsymbol{I}_{K_1} \ldots g_N \boldsymbol{I}_{K_N}) \in \mathbb{C}^{\sum_n K_n \times \sum_n K_n},
\end{aligned}$$

while the channel matrix is defined in (14.1). Following this result, the input-output relation in (14.3) is rewritten as a set of equations for each carrier. That is,

$$\hat{\boldsymbol{d}}_n = g_n \boldsymbol{H}_n \boldsymbol{F}_n \boldsymbol{B}_n^{-1} \boldsymbol{P}_n \boldsymbol{d}_n + g_n \boldsymbol{n}_n \tag{14.4}$$

Of course, this is a non-cooperative model. At each carrier, K_n users are precoded so that the feedback filter is a $K_n \times K_n$ matrix and the forward matrix is $T \times K_n$. Before starting the optimization problem, we shall discuss some properties and constraints of the THP filter.

Firstly, the receiver is kept as simplest as possible and performs an AGC. In this sense, since the SINR is preserved —signal, noise and interference are equally amplified —, this parameter should be irrelevant in the optimization. However, it might be needed mathematically in the process.

Secondly, the feedback filter $\boldsymbol{B}$ performs the precoding at each carrier as, assuming complete orthonormality in frequency, it is in charge of iteratively canceling out the interference that users cause among them. The iterative nature makes these filters to be lower triangular with one main diagonal as the encoder can only use information of already precoded users to precode a given user signal.

Thirdly, the permutation matrix is added into the equations in order to consider the user ordering. Since the optimization will try to make $\hat{\boldsymbol{d}}$ and $\boldsymbol{d}$ very similar, its effect will be counteracted.

Fourthly, the modulo operation is treated as in the single-user case as individual scalar arithmetic modulo operation to each symbol. The size of the modulo set, given by τ_I and τ_Q in bi-dimensional constellations, is adapted to the modulation. This makes sense in third and forth generation of telecommunication services, where the system might serve several users who may be using very distinct services.

Finally, the power allocation matrix deserves attention to the multi-user case. Let us temporally omit the user ordering and the carrier notation so that v_k is the precoded version of user k. The system must allocate the K v_k signals to the T antennas, namely the x_1 to x_T signals, and it is a controversial issue since one might think that each user will be placed at one transmitting antenna. The fact is that the filter $\boldsymbol{F}$ merges the K signals along the transmitting antennas following some distribution. This means that x_1 will be obtained from a portion of v_1, v_2,...,v_K following a linear combination according to the first row of $\boldsymbol{F}$, $\boldsymbol{f}_1$; and successively. Determining $\boldsymbol{F}$ is equivalent to find the K $T \times T$ covariance matrices —which will be given by the corresponding columns $\boldsymbol{f}_k \in \mathbb{C}^{1\times T}$— that optimally allocate the users' power. That is,

$$\boldsymbol{F} = \left[\begin{array}{ccc} \boldsymbol{f}_1 & \dots & \boldsymbol{f}_K \end{array} \right]. \tag{14.5}$$

This consideration on the columns of the forward filter is very important in the multi-user systems as, though we can use a compact vectorial notation, the information of user k —which is spitted along the T antennas through $\boldsymbol{f}_k$— is captured by user j —through its MISO channel $\boldsymbol{h}_j$— and if $j \neq k$, all the power from user k is a potential interferer to user j. This has a very important consequence in the determination of the user rates and capacities, as it will be seen in the sequel. The forward filter, hence, must accomplish the sum power constraint over all bandwidth and transmitting antennas:

$$\sum_n \mathrm{tr}(\boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H) \leq \mathcal{E}, \tag{14.6}$$

where with the use of dithering and the modulo operation, $\boldsymbol{R}_{\boldsymbol{v}_n}$ is fixed to $\tau_I \tau_Q / 6 = \sigma_v^2$ (cf. 5.2.3). As a result, the transmitted power constraint can be rewritten as

$$\sigma_v^2 \sum_n \sum_{k=1}^{K_n} \boldsymbol{f}_{n,\pi_n(k)}^H \boldsymbol{f}_{n,\pi_n(k)} \leq \mathcal{E}. \tag{14.7}$$

Chapter 15

Optimal MMSE Filters

This chapter provides the optimization of the filters described in the previous chapter for the proposed THP structure. The system assumes no priorities among users so that $\alpha_k = 1/K$ for all k and, without loss of generality, the maximization and minimization processes will be held with $\alpha_k = 1$. This applies to the BC sum capacity, $\mathcal{C}_{BC} = \max \sum_k R_k$ and the minimization of the MSE, $\min \sum_n \sum_k \varepsilon^2_{n,k}$.

15.1 A Priori User Ordering

The user ordering is performed through the carrier-dependent permutation matrix, $\boldsymbol{P}_n$, which is defined by the ordering function π_n that orders the K_n users as $\mathcal{O}_n = \{\pi_n(1) \dots \pi_n(K_n)\}$. This is done by setting to zero all the elements of the permutation matrix except for the elements located at the $\pi_n(k)$-th column and k-th row, which are ones. In other words, if $\pi_n(k) = j$, it means that user j is placed at the k-th position. This can be mathematically seen using the unitary vector $\boldsymbol{i}_j$ which is the j-th column of the identity matrix. Therefore,

$$\boldsymbol{P}_n \triangleq \sum_{k=1}^{K_n} \boldsymbol{i}_k \boldsymbol{i}^T_{\pi_n(k)}. \tag{15.1}$$

This matrix is orthonormal so $\boldsymbol{P}_n^T \boldsymbol{P}_n = \boldsymbol{P}_n \boldsymbol{P}_n^T = \boldsymbol{I}_{K_n}$. The order that minimizes the residual MSE after the complete optimization process should be found by finding the $K!$ possible permutation and then pick up the bets one. However, as motivated in subsection 13.3.4, side information can be considered here prior to the process. As a result of the MISO structure of the BC, expression (13.1) is reduced to

$$R^{BC}_{n,\pi_n(k)} = \log \left(1 + \frac{\boldsymbol{h}_{n,\pi_n(k)} \boldsymbol{\Sigma}_{n,\pi_n(k)} \boldsymbol{h}^H_{n,\pi_n(k)}}{\sigma^2_{n_{n,\pi_n(k)}} + \boldsymbol{h}_{n,\pi_n(k)} \left(\sum_{j=k+1}^{K_n} \boldsymbol{\Sigma}_{n,\pi_n(j)} \right) \boldsymbol{h}^H_{n,\pi_n(k)}} \right), \tag{15.2}$$

where $\boldsymbol{h}_{n,\pi_n(k)}$ is the channel frequency response $T \times 1$ vector for user $\pi_n(k)$ at carrier n, $\boldsymbol{\Sigma}_{n,\pi_n(k)} \triangleq E[\boldsymbol{x}_{n,\pi_n k} \boldsymbol{x}^H_{n,\pi_n k}]$ is the transmit covariance $T \times T$ matrix for user $\pi_n(k)$ at carrier n and $\sigma^2_{n_{n,\pi_n(k)}}$ is the corresponding noise term. Clearly, equation (15.2) is a non-convex problem and the power allocation process requires sophisticated concepts. However, it is easy to deduce that the more power we assign to one user, the more it will interfere to the users who have been already precoded. In this sense, finding the optimal user ordering deals with minimizing the interference that they cause among them. From (15.2), the last user, namely $\pi_n(K_n)$, is interference free because the precoder knows all the

other precoded signals, but it will interfere all the users so it would be adequate to place a small amount energy to it. On the other hand, the first user does not interfere to any user and it is the user that needs to cope with more interference, so it is reasonable to let it place as much power as desired. This simple discussion suggests that, at each carrier, $\mathcal{E}_{n,\pi_n(1)} \geq \ldots \geq \mathcal{E}_{n,\pi_n(K_n)}$. Therefore, to **treat the users fairly** —that is, we assume that all the users deserve the same service quality—, it makes sense to place the user with best channel conditions last as it will receive a small portion of energy and, consequently, place the user with the worst channel condition to the first position. The question that arises is how to measure the channel conditions before knowing the user ordering. A simple way —without having defined any error signal— is to compute the single-user rate at the n-th carrier setting $\alpha_j = 0\ \forall j \neq k$ so that from (15.2) it derives

$$R^1_{n,\pi_n(k)} = \log\left(1 + \frac{\boldsymbol{h}_{n,\pi_n(k)}\boldsymbol{\Sigma}_{n,\pi_n(k)}\boldsymbol{h}^H_{n,\pi_n(k)}}{\sigma^2_{n_{n,\pi_n(k)}}}\right).$$

From it, we can establish the scalar parameter $\xi_{n,\pi_n(k)}$ as the measurement of channel conditions as

$$\xi_{n,\pi_n(k)} \triangleq \frac{\boldsymbol{h}_{n,\pi_n(k)}\boldsymbol{h}^H_{n,\pi_n(k)}}{\sigma^2_{n_{n,\pi_n(k)}}}.$$

Notice that this parameter is the squared value $\pi_n(k)$-th eigenvalue of the equivalent channel $\boldsymbol{R_n}_n^{-1/2}\boldsymbol{H}_n$ —this equivalent channel is usually employed in information theory to reduce equations to known expressions and it can be regarded as a whitening process from the communications point of view, as it is done in the next subsections. Finally, the user ordering can be given with the following condition, for each carrier, under **best-last** approach:

$$\pi_n(k) = \underset{j \notin \{\pi_n(k+1),\ldots,\pi_n(K_n)\}}{\arg\max}\ \xi_{n,j}, \tag{15.3}$$

for k from K_n to 1. Having, $\mathcal{O}_n$, it is automatic to determine $\boldsymbol{P}_n$ from (15.1). Though the user ordering has been determined with a priori but experienced point of view, after the whole optimization design this issue will be reviewed in section 15.4 in order to check whether the best-worst approach is ethically optimal.

15.2 Optimal Power Allocation

The power allocation consists in determining the set of matrices $\{\boldsymbol{F}_n\}_{1\leq n\leq N}$ that optimally places the information of the K users in the system along the T transmitting antennas and the N available OFDM carriers. From the expression of the MISO BC achievable rate (15.2), this is equivalent to determine the set of downlink transmit variances $\{\boldsymbol{\Sigma}_{n,\pi_n(k)}\}_{1\leq n\leq N, 1\leq k\leq K_n}$ that maximize the sum rate. The relation between sigmas and f's is trivial as it is defined as the expected value of the transmitted signal of the $\pi_n(k)$-th user at the n-th frequency along the transmitted antennas. Since it only must consider the selected user, it is obtaining by setting all the other users' precoded signals to zero:

$$\boldsymbol{\Sigma}_{n,\pi_n(k)} \triangleq E[\boldsymbol{x}_n\boldsymbol{x}_n^H]|_{v_{n,\pi_n(j)}=0\ \forall j\neq k}.$$

From the structure of $\boldsymbol{F}_n$ in (14.5), it follows that

$$\boldsymbol{\Sigma}_{n,\pi_n(k)} = \sigma^2_v \boldsymbol{f}_{n,\pi_n(k)}\boldsymbol{f}^H_{n,\pi_n(k)}. \tag{15.4}$$

With the order of precoding given by the former section, these matrices are designed to minimize the interference caused amongst the users, minimization that is equivalent to maximize the rate, maximize the SINR and minimize the MSE. Under a critical view, the forward filter must give direction to the information of a given user to not only considering its channel vector, but also taking into account how it interferes the other users, as the interference is the scalar value resulting from the product $\boldsymbol{h}_{n,\pi_n(k)} \left(\sum_{j=k+1}^{K_n} \boldsymbol{\Sigma}_{n,\pi_n(j)} \right) \boldsymbol{h}_{n,\pi_n(k)}^H$. This simple observation justifies that we cannot solve the BC multi-user power allocation as if it was a point-to-point communication, neither a single antenna BC (there is no direction at all). Mathematically speaking, the problem written by

$$\{\boldsymbol{\Sigma}_{n,\pi_n(k)}\}_{1\leq n\leq N, 1\leq k\leq K_n} = \underset{\{\boldsymbol{\Sigma}_{n,\pi_n(k)}\}:\sum_{n=1}^{N}\sum_{k=1}^{K_n}\mathrm{tr}(\boldsymbol{\Sigma}_{n,\pi_n(k)})\leq\mathcal{E}}{\arg\max} \sum_n \sum_k R_{n,\pi_n(k)}^{BC}$$

is non-convex (e.g. [MO79]). Here, the power constraint (14.6) has been rewritten from (14.7) using (15.4). The good news are that using the MAC-BC duality [VJG03], the corresponding MAC optimization problem turns out to be convex.

The following two subsections provide the optimal power allocation for the MISO-OFDM BC THP in two steps: first the power is partitioned along the N carriers and, then, users are loaded at each carrier. Notice that in systems with full granularity —that is, the system does not distinguish in between frequency or space eigen channels to place the users— the power allocation problem can be done in one step. In our case, however, we are interested in placing $K \leq T$ spatially multiplexed users, each of one making use of as many OFDM carriers as possible; fact that requires a two steps allocation.

15.2.1 Frequency Power Allocation

This problem consists of finding the set of $\{\mathcal{E}_{n,\cdot}\}_{1\leq n\leq N}$ that maximizes the sum of rates along frequency: $R_{1,\cdot}^{BC} + \ldots R_{N,\cdot}^{BC}$, where $R_{n,\cdot}^{BC} \triangleq \sum_{k=1}^{K_n} R_{n,\pi_n k}$. The important point is that there is no interference between $R_{n,\cdot}^{BC}$ and $R_{m,\cdot}^{BC}$, $m \neq n$, as we assume complete orthogonality between carriers. As a consequence, we can regard this maximization problem as a point-to-point MIMO channel as the resulting channel matrix $\mathcal{H}$ is block diagonal. Therefore, the optimal power allocation would be to choose the input covariance matrices along the eigenvectors of the channel matrix —remember, there are a total of $\sum_n K_n$— by water-filling on its eigenvectors. However, this approach does not make any distinction between frequency and space —like the single-user MIMO part— and it could result in placing zero energy to a certain user along the N carriers. For this reason, we are not interested in finding the corresponding covariance matrices, only the portion of power directed to each carrier bin. To do so, the system will water-fills each carrier according to the user that deserves more energy at that given carrier. This is very well-known (e.g. [Gal68] or [Tel99] for MIMO) to be

$$\mathcal{E}_n \triangleq \mathcal{E}_{n,\cdot} = \max_{k\in\{1,\ldots,K_n\}} \left(\frac{1}{\mu} - \frac{1}{\xi_{n,\pi_n(k)}} \right)^+, \tag{15.5}$$

where μ is selected to fulfill the power requirement $\sum_n \mathcal{E}_{n,\cdot} = \mathcal{E}$ with equality. Notice that inside each carrier the users are ordered trough the parameter xi, but not along the carriers.

15.2.2 Spatial Power Allocation

Given the set $\{\mathcal{E}_n\}_{1\leq n\leq N}$, the system must calculate the set of transmit covariances of all the users at all the carriers, $\boldsymbol{\Sigma}_{n,\pi_n(k)}$. Once this is done, it is automatic to determine the forward filters through

(15.4) taking into consideration that the covariance matrix is positive semi-definite and admits the corresponding SVD.

As motivated above, solving (15.2) is a non-convex problem, therefore obtaining the optimal rates and transmission policy is a very difficult task. The duality technique presented in [VJG03] transforms the non-convex downlink problem into a convex sum power uplink MAC problem, which is much easier to solve, from which the optimal downlink variance matrices can be found. The key point is, thus, determining efficient algorithms to solve optimization in the uplink sense. However, in the sum power MAC problem, the users in the system have a joint power constraint instead of individual constraints as in the conventional MAC. As in the case of the conventional MAC, there exist standard interior point convex optimization algorithms (cf. [BV04]) that solve the problem. An interior point algorithm, however, is considerably more complex that other algorithms and do not scale when there are a large number of users. Other algorithms, such as based on minimax techniques, suffer also from high complexity. In this work, an iterative algorithm inspired by the water-filling algorithm discussed by Cioffi in [YRBC04] will be used. This algorithm has been discussed by A. Goldsmith in [JRV$^+$05] for the general MIMO BC channel, here we particularize it for the THP and MISO BC case.

The set of N MISO BC channels with rate expressed in (15.2) read the following optimization problem:

$$\mathcal{C}_n^{BC} = \max_{\{\boldsymbol{\Sigma}_{n,\pi_n(k)}\}:\sum_{k=1}^{K_n}\mathrm{tr}(\boldsymbol{\Sigma}_{n,\pi_n(k)})\leq\mathcal{E}_n} \sum_{k=1}^{K_n} \log\left(1+\frac{\boldsymbol{h}_{n,\pi_n(k)}\boldsymbol{\Sigma}_{n,\pi_n(k)}\boldsymbol{h}_{n,\pi_n(k)}^H}{\sigma^2_{n_{n,\pi_n(k)}}+\boldsymbol{h}_{n,\pi_n(k)}\left(\sum_{j=k+1}^{K_n}\boldsymbol{\Sigma}_{n,\pi_n(j)}\right)\boldsymbol{h}_{n,\pi_n(k)}^H}\right),$$

where the sum-rate capacity is function of $\mathcal{C}_n^{BC}(\boldsymbol{h}_{n,1},\ldots,\boldsymbol{h}_{n,K_n},\mathcal{E}_n)$, to highlight that the channel vectors are settled in the downlink. The power allocation consists in determining the $T\times T$ variance matrices to achieve this maximum. The duality MAC-BC states that the dirty paper region for the MIMO BC is equal to the capacity region of the dual MIMO MAC. This implies that the sum capacity of the MISO BC is equal to the sum capacity of the single-input multiple-output (SIMO) MAC whose uplink channels are the trasposed conjugated of the corresponding downlink channels. In other words, at the n-th carrier:

$$\mathcal{C}_n^{BC}(\boldsymbol{h}_{n,1},\ldots,\boldsymbol{h}_{n,K_n},\mathcal{E}_n) = \mathcal{C}_n^{MAC}(\boldsymbol{h}_{n,1}^H,\ldots,\boldsymbol{h}_{n,K_n}^H,\mathcal{E}_n),$$

where the SIMO MAC sum-rate capacity is given by (e.g. [YRBC04] particularized for each user with one transmitting antennas and the receiver with T receiving antennas) maximizing the sum of rates achieved in the uplink. In this case, we consider that the user detection order is the inverse of the precoding —that is, user $\pi_n(K_n)$ is detected first, through $\pi_n(1)$— for which users already decoded, as are known to the receiver, are not considered interference. As a result, the achievable rate of user $\pi_n(k)$ is dual to expressions (13.1) and (15.2):

$$R_{n,\pi_n(k)}^{MAC} = \log\frac{\left|\sum_{j\leq k}\boldsymbol{h}_{n,\pi_n(j)}^H q_{n,\pi_n(j)}\boldsymbol{h}_{n,\pi_n(j)}+\boldsymbol{R}_{\boldsymbol{n}_n}\right|}{\left|\sum_{j<k}\boldsymbol{h}_{n,\pi_n(j)}^H q_{n,\pi_n(j)}\boldsymbol{h}_{n,\pi_n(j)}+\boldsymbol{R}_{\boldsymbol{n}_n}\right|}, \tag{15.6}$$

where now $q_{n,\pi_n(k)}$ represents the (scalar) transmit variance and $\boldsymbol{R}_{\boldsymbol{n}_n}$ is the noise covariance $T\times T$ matrix of the MAC channel. and the sum-rate capacity reads

$$\mathcal{C}_n^{MAC} = \max_{\{q_{n,\pi_n(k)}\}:\sum_{k=1}^{K_n}q_{n,\pi_n(k)})\leq\mathcal{E}_n} \sum_{k=1}^{K_n} R_{n,\pi_n(k)}^{MAC}$$

Comparing the rate expressions for the MISO BC (15.2) and the SIMO MAC case (15.6), we notice a subtle difference which is the essence of the existence of the duality. While in the BC the interference that one user causes due to its power allocated in the coexisting system is received by the uncooperative receivers by their own channel, which makes the problem of finding the downlink covariance matrices a very difficult problem; the interference in the MAC channel, though it is caused by the not yet decoded users, they come in different channels and the problem can be solved with the concept of **effective channel**. Let us take expression (15.6) and multiply both numerator and denominator by the inverse of the denominator. Since there are common parts, it becomes:

$$R^{MAC}_{n,\pi_n(k)} = \log \left| \boldsymbol{I} + \left(\sum_{j<k} \boldsymbol{h}^H_{n,\pi_n(j)} q_{n,\pi_n(j)} \boldsymbol{h}_{n,\pi_n(j)} + \boldsymbol{R}_{\boldsymbol{n}_n} \right)^{-1} \boldsymbol{h}^H_{n,\pi_n(k)} q_{n,\pi_n(k)} \boldsymbol{h}_{n,\pi_n(k)} \right| .$$

We define the new term as the Hermitian matrix $\boldsymbol{B}_{n,\pi_n(k)} \triangleq \left(\sum_{j<k} \boldsymbol{h}^H_{n,\pi_n(j)} q_{n,\pi_n(j)} \boldsymbol{h}_{n,\pi_n(j)} + \boldsymbol{R}_{\boldsymbol{n}_n} \right)$. This is a term that takes into account the noise and the potential interferer users to $\pi_n(k)$. To simplify, we take the square root of matrix $\boldsymbol{B}^{-1}_{n,\pi_n(k)}$ and, using the matrix property $|\boldsymbol{I} + \boldsymbol{AB}| = |\boldsymbol{I} + \boldsymbol{BA}|$, where both $\boldsymbol{A}$ and $\boldsymbol{B}$ matrices are Hermitian, we get

$$R^{MAC}_{n,\pi_n(k)} = \log \left| \boldsymbol{I} + \boldsymbol{c}^H_{n,\pi_n(k)} q_{n,\pi_n(k)} \boldsymbol{c}_{n,\pi_n(k)} \right| , \tag{15.7}$$

where $\boldsymbol{c}_{n,\pi_n(k)} \triangleq \boldsymbol{h}_{n,\pi_n(k)} \boldsymbol{B}^{-1/2}_{n,\pi_n(k)}$ is the uplink effective channel of user $\pi_n(k)$. This effective channel is nothing more than a whitened and adapted channel, since transform the complicated problem of multiple interferences among users to a set of point-to-point channels. Notice that this development is the axiom of the MAC-BC duality in [VJG03]. As it will be seen, the same structure can be held in the downlink to obtain another effective channel. The difficult task, done by Goldsmith, is to find the relation between both effective channels, relationship that gives the tool to the duality. With analogous procedure, the MISO BC rate expression can be reduced to

$$R^{BC}_{n,\pi_n(k)} = \log \left(1 + \boldsymbol{c'}_{n,\pi_n(k)} \boldsymbol{\Sigma}_{\pi_n(k)} \boldsymbol{c'}^H_{n,\pi_n(k)} \right) ,$$

where now the downlink effective channel is $\boldsymbol{c'}_{n,\pi_n(k)} = A^{-1/2}_{n,\pi_n(k)} \boldsymbol{h}_{n,\pi_n(k)}$, where the scalar $A_{n,\pi_n(k)}$ —which is a $R \times R$ matrix if the receivers have R antennas— is defined similarly by $A_{n,\pi_n(k)} \triangleq \left(\boldsymbol{h}_{n,\pi_n(k)} \left(\sum_{j>k} \boldsymbol{\Sigma}_{\pi_n(j)} \right) \boldsymbol{h}_{n,\pi_n(k)} + \sigma^2_{n_{n,\pi_n(k)}} \right)$ and reflects again the interference experienced by the user (the proof could be an exercise for the reader). We would like to insist, again, that the capacity region for the BC is a non-convex problem, so determine the set of transmit covariance matrices should be done exhaustively.

Continuing our development with (15.7), we know can state that the power allocation problem in the uplink reads

$$\{q_{n,\pi_n(k)}\}_{1 \le k \le K_n} = \underset{\{q_{n,\pi_n(k)}\} : \sum_{k=1}^{K_n} q_{n,\pi_n(k)} \le \mathcal{E}_n}{\operatorname{argmax}} \sum_{k=1}^{K_n} \log \left| \boldsymbol{I} + \boldsymbol{c}^H_{n,\pi_n(k)} q_{n,\pi_n(k)} \boldsymbol{c}_{n,\pi_n(k)} \right| . \tag{15.8}$$

The objective is now clear: first this set is determined and, afterward translated to the covariance matrices for the downlink dual. The use of the effective channel is very clever because it transforms the MAC convex problem into a known problem, as expression (15.7) is a point-to-point SIMO channel with channel $\boldsymbol{c}_{n,\pi_n(k)}$, thus it is well known that the power must be allocated according to the

eigenvectors of the channel matrix —which takes into consideration its channel, the noise and the interference! But it is worthy to highlight that once the rate is optimized, the input variances change, so another optimization has not be done. For that reason, it is an iterative algorithm as no closed form is possible since the parameters to optimize depend on themselves optimization. It has been proved that this algorithm converges (e.g. [YRBC04] or [JRV+05]).

The input variances $q_{n,\pi_n(k)}$ must therefore satisfy the water-filling condition, according to the eigenvalues of the effective channels. However, since there is a sum power constraint —remember that it is the dual problem of our BC—, the water level of all users in the carrier must be equal. This is akin to saying that no advantage will be gained by transferring power from one user to another with a higher water-filling level to another with lower level. Note that his is different from the conventional MAC problem, where there are individual power constraints.

Finally, the leitmotiv of iterative algorithm is to treat the effective channels as parallel and non interfering channels —as it is from (15.8). Therefore, the variances at the step are the solution of the maximization of the sum rate MAC capacity, which is a convex problem whose solution is given by the classical water-filling with total power $\mathcal{E}_n$ [Tel99] as the MAC sum capacity is a Schur-convex function (e.g. [PCL03]) —we omit the development here, which is obtained from Lagrange duality [MN99], a very well-known result in information and communication theories. The i-th iteration of the algorithm is:

1. Generate the set of K_n effective channels using the variances obtained in the last step, $i-1$:

$$\boldsymbol{c}^i_{n,\pi_n(k)} \triangleq \boldsymbol{h}_{n,\pi_n(k)} \left(\sum_{j<k} \boldsymbol{h}^H_{n,\pi_n(j)} q^{(i-1)}_{n,\pi_n(j)} \boldsymbol{h}_{n,\pi_n(j)} + \boldsymbol{R}_{\boldsymbol{n}_n} \right)^{-1/2}.$$

2. Obtain the water-filling variances for this step:

$$q^i_{n,\pi_n(k)} = \left(\frac{1}{\mu_n} - \frac{1}{\boldsymbol{c}^i_{n,\pi_n(k)} (\boldsymbol{c}^i_{n,\pi_n(k)})^H} \right)^+, \tag{15.9}$$

where the water-filling level μ_n is chosen such that $\sum_{k=1}^{K_n} q_{n,\pi_n(k)} = \mathcal{E}_n$.

Once the algorithm converges —that is, the obtained sum-rate is stabilized: $|\mathcal{C}^{MAC,i}_n - \mathcal{C}^{MAC,i-1}_n| < \epsilon$— the downlink transmit variances are obtained through the MAC-BC duality [VJG03] as:

$$\boldsymbol{\Sigma}_{n,\pi_n(k)} = \boldsymbol{B}^{-1/2}_{n,\pi_n(k)} \boldsymbol{U}_{n,\pi_n(k)} \boldsymbol{V}^H_{n,\pi_n(k)} A^{1/2}_{n,\pi_n(k)} q_{n,\pi_n(k)} A^{1/2}_{n,\pi_n(k)} \boldsymbol{V}_{n,\pi_n(k)} \boldsymbol{U}^H_{n,\pi_n(k)} \boldsymbol{B}^{-1/2}_{n,\pi_n(k)}, \tag{15.10}$$

where we made use of the SVD decomposition $\boldsymbol{B}^{-1/2}_{n,\pi_n(k)} \boldsymbol{h}^H_{n,\pi_n(k)} A^{-1/2}_{n,\pi_n(k)} = \boldsymbol{U}_{n,\pi_n(k)} \boldsymbol{D}_{n,\pi_n(k)} \boldsymbol{V}^H_{n,\pi_n(k)}$, with $\boldsymbol{D}_{n,\pi_n(k)}$ a $T \times T$ square diagonal matrix. To obtain the covariance matrices, from the definition of the parameter A, we shall start by the last user, $\pi_n(K_n)$. The main reason it that in the MAC-BC duality, Goldsmith (see [VJG03], section IV.B in pages 2661 and 2662 for a complete development) makes use of another effective channel that takes into consideration the interference due to the MAC channel —which is represented by left-multiplying the channel by $\boldsymbol{B}^{-1/2}_{n,\pi_n(k)}$, corresponding to the not yet decoded dual users; i.e. the already precoded users— and the BC channel —which is represented by right-multiplying the channel by $A^{-1/2}_{n,\pi_n(k)}$, corresponding to not yet precoded users; i.e. the already decoded dual users. In other words, the overall effective channel is $\boldsymbol{B}^{-1/2}_{n,\pi_n(k)} \boldsymbol{h}^H_{n,\pi_n(k)} A^{-1/2}_{n,\pi_n(k)}$. In this sense, the covariance matrices take into consideration the channel direction of all the users: both already precoded users because it is a potential interferer to them but also the not yet precoded users,

as they are actually interference to it. As a result, starting by the last precoded user, as it does not see any interferer it selects the covariance matrix according to its effective channel, which is only formed by the dual MAC users, with already known variance. The next user is $\pi_n(K_n-1)$ who computes its variance according to its effective channel, formed by the dual MAC users and his potential interferer: the last user whose covariance matrix has been computed. This is done successively to the first user, who will direct its information —that is why we deal with covariance matrices in the downlink thanks to the multiple-inputs— accordingly to all the already known users, which are all of them interferer.

We shall highlight that since we are in a MISO BC channel, the effective channel is a $T \times 1$ vector, for which the SVD returns a single eigenvalue in matrix $\boldsymbol{D}_{n,\pi_n(k)}$ and $\boldsymbol{V}_{n,\pi_n(k)}$ is a unit-norm scalar. As a result, the former expression can (15.10) be rearranged to

$$\boldsymbol{\Sigma}_{n,\pi_n(k)} = A_{n,\pi_n(k)} q_{n,\pi_n(k)} \boldsymbol{B}_{n,\pi_n(k)}^{-1/2} \boldsymbol{U}_{n,\pi_n(k)} \boldsymbol{U}_{n,\pi_n(k)}^{H} \boldsymbol{B}_{n,\pi_n(k)}^{-1/2}, \tag{15.11}$$

noticing that $A_{n,\pi_n(k)}$ is also scalar. Finally, as we are interested in computing the filter that generates these covariances, we recall (15.4) to obtain:

$$\boldsymbol{f}_{n,\pi_n(k)} = \frac{A_{n,\pi_n(k)} q_{n,\pi_n(k)}}{\sigma_v^2} \boldsymbol{B}_{n,\pi_n(k)}^{-1/2} \boldsymbol{U}_{n,\pi_n(k)}, \tag{15.12}$$

with appropriately placing these columns to the corresponding $\boldsymbol{F}_n$ filter (14.5). Equation (15.12) gives a very intuitive interpretation. Regarding the scalar value, it is deduced that the energy placed for this user is proportional to the interference it receives —that is, $A_{n,\pi_n(k)}$—, and also proportional to the power that it would be allocated in a dual MAC channel to that user —that is, $q_{n,\pi_n(k)}$. From (15.9), it is seen that in the dual MAC problem, more power is allocated to users with higher effective channel, which correspond to high individual channel gain and/or small level of noise and dual interference — simply recall how $\boldsymbol{c}_{n,\pi_n(k)}$ is determined. The dual MAC interference is nothing more than a measure of this certain user being a potential interferer to the other users. Hence, a user who is considered a potential interferer —that is, it has a lot of dual MAC interference and therefore little amount of MAC energy, $q_{n,\pi_n(k)}$, allocated— in the BC, a little amount of energy will be placed, as our intuition said. On the other hand, a user who is not an interferer but copes with interference—the first user in the BC— will be assigned a big amount of energy thanks to the dual MAC channel. Finally, the vectorial part of (15.12) states that the information of user $\pi_n(k)$ will be placed according to the direction of its effective channel, i.e., in the direction of its own channel rotated by means of the inverse of $\boldsymbol{B}_{n,\pi_n(k)}$, which again reflects the potential dual interferer's of the MAC. Therefore, the information of a given user is placed to maximize its own channel use but also minimize the effect that it can produce to others.

All this interpretation and results could not be possible to obtain without the theory of DPC form the information theory point of view. The use of the MSE as a target function to optimize does not give such clarity of real multi-user systems.

15.3 Optimal Feedback Matrix

Given the optimal power allocation filters, $\{\boldsymbol{F}_n\}_{1\leq n\leq N}$, as well as a given user ordering at each carrier, $\mathcal{O}_n$ determined through the set of permutation matrices $\{\boldsymbol{P}_n\}_{1\leq n\leq N}$, the feedback filter is in charge of canceling out the MUI present at each carrier. In the single-user scenarios, this filter diminished the inter-symbol interference (ISI) and the ICI inherent in frequency-selective SISO and MIMO channels.

However, both temporal and spatial interferences where somehow used by the system to enhance the performance using diversity, as the information transmitted in the multi-path and multi-antennas was **from the same user**. This fact is immediately justified with the results depicted in figures 7.3 and 12.3 for which the higher the number of multi-paths, i.e. ISI, and transmitting antennas, i.e. ICI, the more richness is given to the system's diversity. In this sense, on the one hand, the feedback filter in the single-user systems ensures the recoverability of the information because, together with the receive filter, it diagonalizes the problem —as seen in the MIMO case— by means of a subset of the signal space, defined by the modulo operation. On the other hand, in multi-user systems the interference caused amongst them is not used by the receivers as they do not cooperate so it is truly interference to them. In this sense, the feedback filter has the very important paper, together with the modulo operation, to successively encode each user regarding the interference that it must cope with. In other words, a user cannot take advantage of the portion of his or her information that is received through another user's channel, as he or she has no access to it. As a consequence, the feedback filter plays a crucial part as it is the essence of the precoding which must cancel the interference of the already precoded users. The intuition springs to mind and recommends the system to diminish the non-diagonal entries of the transmission chain in a MMSE point of view or force it to zero (ZF).

The classical MMSE approach states that the feedback matrix and the optimal receive parameter g_n should be obtained through

$$\{\boldsymbol{B}_n, g_n\} = \arg\min_{\boldsymbol{B}_n, g_n} \mathrm{tr}(\boldsymbol{E}_n),$$

subject to the feedback filter being lower triangular with one main diagonal and the receive parameter non-zero. Here, the $K_n \times K_n$ error matrix is defined as usual as $\boldsymbol{E}_n \triangleq (\boldsymbol{d}_n - \hat{\boldsymbol{d}}_n)(\boldsymbol{d}_n - \hat{\boldsymbol{d}}_n)^H$. From (14.4) and noticing that $\boldsymbol{d}_n = \boldsymbol{P}_n^T \boldsymbol{B}_n \boldsymbol{v}_n$, the MSE becomes

$$\begin{aligned} \varepsilon^2(\boldsymbol{B}_n, g_n) &= \mathrm{tr}(\boldsymbol{P}_n^T \boldsymbol{B}_n \boldsymbol{R}_{\boldsymbol{v}} \boldsymbol{B}_n^H \boldsymbol{P}_n + g_n^2 \boldsymbol{H}_n \boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}} \boldsymbol{F}_n^H \boldsymbol{H}_n^H + & (15.13)\\ & \quad g^2 \boldsymbol{R}_{\boldsymbol{n}_n} - g_n \boldsymbol{P}_n^T \boldsymbol{B}_n \boldsymbol{R}_{\boldsymbol{v}} \boldsymbol{F}_n^H \boldsymbol{H}_n^H - g_n \boldsymbol{H}_n \boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}} \boldsymbol{B}_n^H \boldsymbol{P}_n). & (15.14) \end{aligned}$$

The derivation of the solution to this problem is analogous to parts II and III. Firstly, the Lagrangian associated to the feedback filter reads, by letting g_n be a constant,

$$\mathcal{L}(\boldsymbol{B}_n) = \varepsilon^2(\boldsymbol{B}_n) - 2\mathrm{Retr}\boldsymbol{\Lambda}(\boldsymbol{B}_n^H \odot \boldsymbol{U}^T - \boldsymbol{I}_{K_n}),$$

where the $K_n \times K_n$ square matrix $\boldsymbol{U}_n$ is defined as in (6.17) and $\odot$ stands for the Hadamard product. Secondly, we differentiate the former expression with respect to the feedback filter transpose conjugated and taking the transpose of the resulting expressions, i.e. we do $\nabla^T_{\boldsymbol{B}_n^H} \mathcal{L}(\boldsymbol{B}_n) = \boldsymbol{0}_{K_n}$. Finally, taking into account that g_n is an scalar and that the autocorrelation of the precoded signal $\boldsymbol{v}_n$ is diagonal —the information among users is uncorrelated— and with the same variance —given by the modulo operation and the use of dithering—, i.e. $\boldsymbol{R}_{\boldsymbol{v}} = \sigma_v^2 \boldsymbol{I}_{K_n}$ at all the carriers, it is arrived to the following feedback filter expression:

$$\boldsymbol{B}_n = \mathrm{diag}^{-1}(\boldsymbol{P}_n \boldsymbol{H}_n \boldsymbol{F}_n)(\boldsymbol{P}_n \boldsymbol{H}_n \boldsymbol{F}_n \odot \boldsymbol{U}_n^T). \tag{15.15}$$

Notice that the parameter g_n disappears from the expression of all the filters in the transmitter side —assuming, evidently, that it is a non-zero value. Expression (15.15) gives a very nice interpretation of the feedback precoding. First of all it is worthy to highlight that the permutation matrix $\boldsymbol{P}_n$ appears straight on the expression. This results —as expected!— in the fact that the precoding order

is given by this matrix, which implements the order $\mathcal{O}_n$ as defined in (15.1). Secondly, we note that the feedback filter, the essence of THP, performs interference pre-subtraction. From figure (14.2), the received signal of the k-th user at the n-th carrier may be written as $y_{n,k} = \boldsymbol{h}_{n,k}\boldsymbol{F}_n\boldsymbol{v}_n + n_{n,k}$ —remember that $\boldsymbol{h}_{n,k}$ is a row vector—, and taking into consideration the internal structure of $\boldsymbol{F}_n$ in (14.5) it reads

$$y_{n,k} = \boldsymbol{h}_{n,k}\boldsymbol{f}_{n,k}v_{n,k} + \boldsymbol{h}_{n,k}\sum_{j\neq k}\boldsymbol{f}_{n,j}v_{n,j} + n_{n,k} \tag{15.16}$$

where $v_{n,j}$ is the precoded signal for the j-th user. For complete interference subtraction, the second term of the former equation must be eliminated. The pre-subtraction is done through the feedback filter which, as $\boldsymbol{B}_n^{-1}\boldsymbol{P}_n\boldsymbol{d}_n = \boldsymbol{v}_n$ and given the lower triangular structure of the feedback filter, becomes $v_{n,k} = d_{n,\pi_n(k)} - \sum_{j=1}^{k-1} b_{n,k,j}v_{n,j}$, where the term $b_{n,k,j}$ has been introduced as $[\boldsymbol{B}_n]_{kj}$. In other words, $b_{n,k,j}$ is the term in charge of canceling the known interference caused by user j to user k, always $j < k$. From (15.16), it is deduced that the known interference from user j to user k is nothing more than $\boldsymbol{h}_{n,k}\boldsymbol{f}_{n,j}v_{n,j}$. That is why the feedback filter obtained in (15.15) has the inner product $\boldsymbol{H}_n\boldsymbol{F}_n$. In other words, each contributing element of the feedback filter —that is, $b_{n,k,j}$ with $j < k$ and $b_{n,k,j} = 1$ for $j = k$— is proportional to the canceling factor $\boldsymbol{h}_{n,k}\boldsymbol{f}_{n,j}$ with the order $\mathcal{O}_n$:

$$b_{n,k,j} = \frac{\boldsymbol{h}_{n,\pi_n(k)}\boldsymbol{f}_{n,\pi_n(j)}}{\boldsymbol{h}_{n,\pi_n(k)}\boldsymbol{f}_{n,\pi_n(k)}},$$

which concords to (15.15).

Finally, since the receive parameter g_n is irrelevant in the design, we allow it to be any value as an AGC at each receiver. That is,

$$\forall g_n > 0$$

Similarly to the single-user MIMO channel in part III, we shall define the useful communication chain —that is, the channel input-output pairs that transmit useful information— of the k-th user as

$$\boldsymbol{q}_{.,k} = [\boldsymbol{q}_{1,k}^T \dots \boldsymbol{q}_{N,k}^T]^T \triangleq \operatorname{diag}(\boldsymbol{G}_{.,k}\boldsymbol{H}_{.,k}\boldsymbol{F}_{.,k}\boldsymbol{B}_{.,k}^{-1}),$$

where $\boldsymbol{H}_{.,k}$ is a matrix containing the N channel responses of user k and similarly defined for the feedback and power allocation filters. Analog to MIMO-OFDM, the multi-user MISO-OFDM receiver shall combine the symbol obtained through the N bands under a maximum diversity strategy by a RAKE receiver [Cio02]. This can be done assuming that the k-th receiver has complete CSI—notice that complete refers to the knowledge of the channel realization— of its N channels. This knowledge is obtained through sufficient training in the downlink channel, as well as information on the precoding set of filters $\{\boldsymbol{f}_{n,k}\}_{1\leq n\leq N}$ sent by the transmitter to user k. Mathematically, the signal sent to the precoder is the weighted sum of the output of the receiver filter signal, after compensating the dithering and the modulo operation, as:

$$\tilde{\boldsymbol{s}}_k = \sum_{n=1}^{N}\operatorname{diag}(\boldsymbol{q}_{n,k}^H)\tilde{\boldsymbol{s}}_{n,k}.$$

Notice that we can define the equivalent channel matrix at the n-th carrier taking into consideration all the signal processing filters that appear from signal $\boldsymbol{d}$ to its estimated value. That is, from (14.4):

$$\boldsymbol{Q}_n = g_n\boldsymbol{H}_n\boldsymbol{F}_n\boldsymbol{B}_n^{-1}\boldsymbol{P}_n^T. \tag{15.17}$$

In the single-user cases, since both receive and transmit filters cooperate, the complete diagonalization of the problem was easy to accomplish due to the structure of the power allocation filter and the corresponding expression of the receive filter. In a non-cooperative scenario, the lower triangular structure of the feedback matrix is a restriction to fully diagonalize the transmission chain. In other words, the precoding scheme can only force zeros in the elements below the main diagonal, for which matrix $\boldsymbol{Q}_n$ is upper triangular. This reflects the fact that the first user sees $K-1$ interferences, while the last user is interference free.

15.4 Review of User Ordering

As pronounced at the beginning of this chapter, the user ordering is an issue that completely affects the performance of the multi-user system, as it defines the level of interference that each user will perceive due to the precoding technique structure. Let us analyze, without having any result, how the order may change the behavior of the system in terms of power allocation and sum capacity. To do so, we assume that two users, who simultaneously receive information through one carrier from a transmitter equipped with two transmitting antennas, have very different channel conditions —if the channel conditions are very similar, which could be the case, the user ordering has little impact, as expected— so that user 1 has better channel conditions that user 2, i.e. $\xi_{n,1} > \xi_{n,2}$. Denote the best-last user ordering, depicted in (15.3), as $\mathcal{O}_n^L$, which in this case is $\mathcal{O}_n^L = \{2,1\}$; and the best-first user ordering as $\mathcal{O}_n^F = \{1,2\}$. Notice that these are the precoding orders in the downlink, whereas the uplink dual order is the inverse. The theory necessary to deal with user ordering scope is given by the effective up and downlink channels, namely $\boldsymbol{c}_{n,k}$ and $\boldsymbol{c}'_{n,k}$ respectively. The downlink effective channel of the first precoded user is its channel plus the penalization of the second user which acts as interferer —i.e. matrix $\boldsymbol{B}_{n,1}$—, whereas the downlink effective channel of the second (last) precoded user, as it does not see interference, is itself.

Under this consideration, the best-first strategy places the user with better channel conditions to be precoded first. This means that its effective channel will be better than if this user was placed in another order —he or she is in a privileged position—, since all the interference show worst channel. Hence, since user 1 is not a potential interferer to others, more power will be allocated to 1 and, as it has the best channel conditions, its performance is outstanding. However, the other users are considered potential interferer and have little interference —lower $A_{n,k}$— and, as a consequence, less energy is placed to these users that, coupled with their bad channel conditions, the performance is penalized.

On the other hand, the best-last strategy is more conservative and the performances are less dispersive. This is justified highlighting the fact that now, user 1 is a potential interferer to 2, and therefore, less energy is placed to 1 compared to the best-first strategy. Similarly, user 2 is precoded first so it is not an interferer and more enery is placed to 2 compared to the best-first strategy, improving its performance.

These asseverations can be seen in figure 15.1 for the dotted lines —best-first— and straight lines —best-last. As it can be appreciated, user 1 (circles) has always better performance than user 2 because it has much better channel condition but a best-first strategy is better for user 1 and worst for user 2, which makes the performances of both users more opposite. In this sense, the selected best-last strategy at the beginning has been a good election. An interesting comment is that the sum-rate capacity, i.e. the maximization of sum of the rates of all the users, is the same regardless the ordering strategy. This has sense as the capacity is a parameter inherent in the channel, independently on the

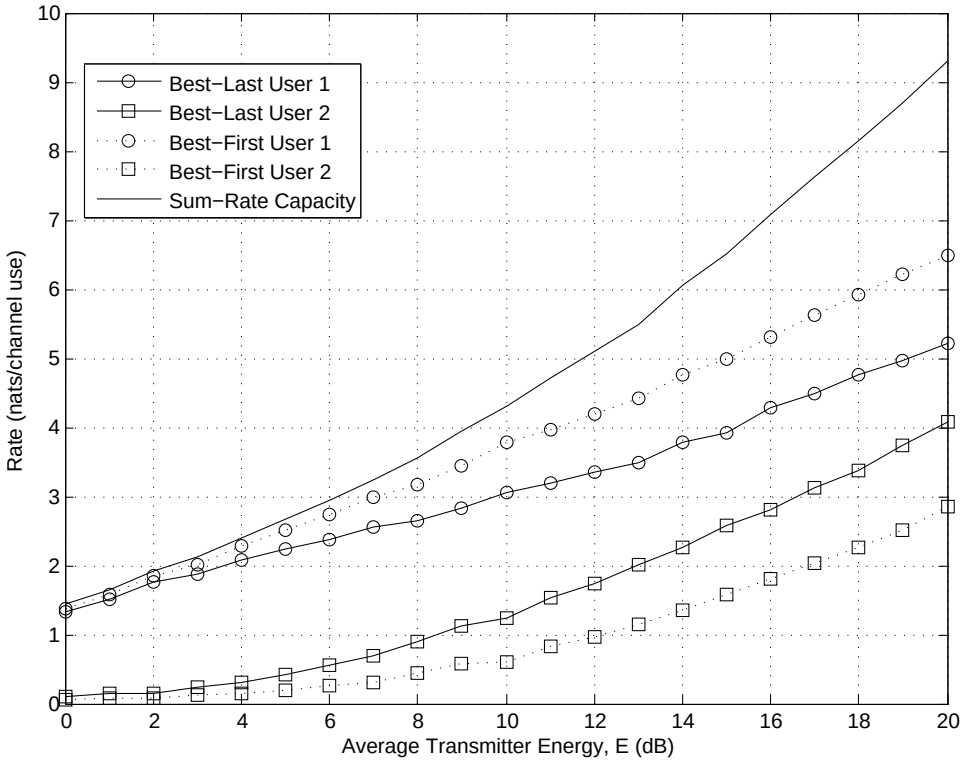


Figure 15.1: Rate capacity for users 1 and 2, with $\xi_{n,1} > \xi_{n,2}$ at a given carrier, versus the average transmitted energy per channel use, $\mathcal{E}$, in a 2×1 MISO BC for two ordering strategies: best-last $\mathcal{O}_n^L = \{2, 1\}$ and best-first $\mathcal{O}_n^L = \{1, 2\}$.

signal processing or precoding techniques.

Chapter 16

Other Optimization Techniques

16.1 Cholesky Factorization

Recently, the use of Cholesky factorization for the MIMO channel reduces the complexity when the complete channel information is assumed to be at the transmitter side (e.g. [KJUB05] for the flat-fading MIMO channel and [JU05] for the BC). Morelli in [MS07] optimizes the Tomlinson-Harashima system for OFDMA-MIMO channels and [KJUB07] gives an excellent comparison between V-BLAST, DFE and THP techniques. The use of Cholesky factorization is well justified under the development explained in this thesis, in subsection 6.3.2. Though the key in Cholesky resides in the fact that the factorization technique provides lower triangular matrices, a structure that is forced to the feedback matrix, some authors apply the factorization as a receipt rather than justifying it. This is not the case of [KJUB07], an excellent paper where Cholesky is used to Tomlinson-Harashima precode spatially multiplexed user data streams. It is worthy to remark that these German authors do not make any consideration on the information theory provided in this work, which is in fact the key point in multi-user systems. However, they present the Cholesky factorization as a very interesting technique to simplify the complexity of the precoding system.

Motivated in 6.3.2, MSE should read in terms of the Frobenius norm of the feedback filter and the Cholesky decomposed matrix. This is easily obtained —as it has been done for single-user channels— making use of the orthogonality principle that holds at the receiver. Though we now deal with multiple users, the same principle holds under the MMSE point of view. The reader is invited to determine the orthogonality principle and the resulting MSE in terms of the feedback filter following the classical development by taking the MSE in equation (15.14), finding the orthogonality principle $E[(\boldsymbol{d}_n - \hat{\boldsymbol{d}}_n)\boldsymbol{y}_n^H] = \mathbf{0}_{K_n}$ which is a result of the optimality at the receivers, i.e. differentiating the MSE with respect to g_n an equaling to zero and, finally, applying this result to the MSE with mathematical manipulations to come by an expression that computes the norm of the feedback filter. Here, nonetheless, we will take on the results of the SISO channel in part II to quickly obtain our objective. To do so, we take the MSE expression (6.16) with the matrix $\boldsymbol{\Phi}$ in (6.15) making the following SISO⟶BC analogies —obtained from comparing (14.4) to (5.8) with $\boldsymbol{F}_n$ and $\boldsymbol{P}_n$ fixed—:

$$\begin{aligned}
\boldsymbol{G} &\longrightarrow g_n \boldsymbol{I}_{K_n} \in \mathbb{R}_+^{K_n \times K_n} \\
\boldsymbol{B} &\longrightarrow \boldsymbol{P}_n^T \boldsymbol{B}_n \in \mathcal{C}^{K_n \times K_n} \\
\boldsymbol{H} &\longrightarrow \boldsymbol{H}_n \boldsymbol{F}_n \in \mathcal{C}^{K_n \times K_n} \\
\boldsymbol{x} &\longrightarrow \boldsymbol{v}_n \in \mathcal{C}^{K_n \times 1}.
\end{aligned}$$

As a result, the orthogonality principle reads

$$\boldsymbol{P}_n^T \boldsymbol{B}_n \boldsymbol{R}_{\boldsymbol{v}_n} (\boldsymbol{H}_n \boldsymbol{F}_n)^H = g_n \left(\boldsymbol{H}_n \boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}_n} (\boldsymbol{H}_n \boldsymbol{F}_n)^H + \boldsymbol{R}_{\boldsymbol{n}_n} \right); \tag{16.1}$$

while the MSE can be expressed through

$$\varepsilon^2(\boldsymbol{B}_n) = \mathrm{tr}(\boldsymbol{P}_n^T \boldsymbol{B}_n \boldsymbol{\Phi}_n \boldsymbol{B}_n^H \boldsymbol{P}_n) \tag{16.2}$$

where the design matrix is now $\boldsymbol{\Phi}_n = \left(\boldsymbol{R}_{\boldsymbol{v}_n}^{-1} + \boldsymbol{F}_n^H \boldsymbol{H}_n \boldsymbol{R}_{\boldsymbol{n}_n}^{-1} \boldsymbol{H}_n \boldsymbol{F}_n \right)^{-1}$. The expression of the design matrix $\boldsymbol{\Phi}_n$ is analogous to (10.2) for the single-user MIMO channel. However, noticing that the orthogonality principle relates two matrices, notice that to apply this result to the expression of the MSE and obtain (16.2) we must consider the matrix $g_n \boldsymbol{I}_{K_n}$, which is Hermitian, and not simply the scalar factor g_n.

It is important to give heed to the fact that (16.2) adverts of the MSE being computed in the order $\mathcal{O}_n$. However, since the trace operator is linear with respect to the diagonal it allows us to write $\mathrm{tr}(\boldsymbol{P}_n^T \boldsymbol{B}_n \boldsymbol{\Phi}_n \boldsymbol{B}_n^H \boldsymbol{P}_n) = \mathrm{tr}(\boldsymbol{B}_n \boldsymbol{\Phi}_n \boldsymbol{B}_n^H \boldsymbol{P}_n \boldsymbol{P}_n^T)$ and, since $\boldsymbol{P}_n \boldsymbol{P}_n^T$ is the identity, it reads $\varepsilon^2(\boldsymbol{B}_n) = \mathrm{tr}(\boldsymbol{B}_n \boldsymbol{\Phi}_n \boldsymbol{B}_n^H)$. This result is not surprising, because the MSE is computed as the sum of MSEs. Despite of this, matrix $\boldsymbol{\Phi}$ has information on the channel conditions of the K_n users —noise and channel gain/attenuation— ordered from 1 to K_n. For this reason we cannot neglect the permutation matrix, as the feedback matrix must precode in the order established by the permutation matrix, $\mathcal{O}_n$. In this line, the conventional Cholesky factorization is not used, but the Cholesky factorization with symmetric permutation at its place (e.g. [KJUB07]). Notice that we cannot employ first the Cholesky factorization to obtain a lower triangular matrix and, afterward, apply the permutation matrix to it to obtain the correct order; because the resulting matrix will not be lower triangular, unless $\mathcal{O}_n = \{1, \ldots, K_n\}$. The hint resides in applying correctly the permutation matrix to $\boldsymbol{\Phi}_n$ with the aim to correctly order both its rows and columns —as it is a Hermitian matrix formed by correlation matrices. To do so, we recall how $\boldsymbol{\Phi}_n$ is defined at the very commencement (6.13) for the single-user SISO channel —with the corresponding SISO⟶BC analogies—:

$$\boldsymbol{\Phi}_n = \boldsymbol{R}_{\boldsymbol{v}_n} - \boldsymbol{R}_{\boldsymbol{v}_n \boldsymbol{y}_n} \boldsymbol{R}_{\boldsymbol{y}_n}^{-1} \boldsymbol{R}_{\boldsymbol{y}_n \boldsymbol{v}_n}.$$

This means that, regarding that a correlation matrix is obtained from the expected value of a column signal multiplied to a transposed conjugated row signal we must multiply the design matrix by $\boldsymbol{P}$ on the left side —to order $\boldsymbol{v}_n$— and $\boldsymbol{P}_n^T$ on the right-side —to order $\boldsymbol{v}_n^H$. Therefore, the matrix is Cholesky factorized as

$$\boldsymbol{P} \boldsymbol{\Phi}_n \boldsymbol{P}_n^T = \boldsymbol{L}_n \boldsymbol{L}_n^H,$$

so the feedback matrix in the Cholesky factorization technique reads:

$$\boldsymbol{B}_n = \mathrm{diag}(\boldsymbol{L}_n) \boldsymbol{L}_n^{-1}. \tag{16.3}$$

The receive filter, i.e. $g_n \boldsymbol{I}_{K_n}$, can be obtained from the orthogonality principle from a formal point of view. In practice, as g_n does not appear in any transmitter filter, from the MMSE criterion it is irrelevant to consider its value at the receiver. What is more, it is simply a linear amplifier that performs AGC because it preserves the SNR at its input and the noise statistical properties.

16.2 MMSE-ZF

The filter optimization for the MMSE-ZF criterion in systems with multiple users which do not cooperate is often referred interference cancellation, regarding the fact that ZF means to completely cancel —i.e. make their value zero— all the known signals that are not considered useful, namely noise and interferences. In single-user channels, the known interference is the useful signal coming from different time delays —frequency-selective SISO channel, which provokes ISI— or different paths at the same time —frequency-selective MIMO channel, which provokes ICI— and this knowledge makes feasible to design the transmitter and the receiver jointly to diminish (MMSE) or fully cancel (ZF) these interferences and the use of the SVD and other factorization techniques. In multi-user channels, though, the transmitter has only knowledge of the interference that the already precoded users will cause, while the receiver has no access to others' received signals. Immediate consequences are that neither SVD nor QR on the channel matrix are handy and that the differences between MMSE and ZF are minimum.

Firstly, the power allocation filter has the same expression for any signal processing technique since it has been obtained through information theory consideration which are independent on the policy applied in the system. As a result, the power allocation filter does not suppose an additional degree of freedom in minimizing the MSE or achieving the ZF condition, as it is fixed. Recalling its formula, (15.12) and the expression of the equivalent channel $\boldsymbol{c}_{n,k}$, the power allocation filter allocates the information taking into account the interferences that one user might receive and cause. In this sense, as all the users receive interference and/or cause interference, the product $\boldsymbol{H}_n\boldsymbol{F}_n$ is neither diagonal nor lower or upper triangular.

Secondly, the feedback filter is lower triangular by construction, as the precoder can only precode a user by means of the already precoded users. This, together with the former discussion on the power allocation filter, makes impossible to fully invert the channel and, as a consequence, the ZF condition cannot hold. This fact has been observed with the definition of $\boldsymbol{Q}_n$ in (15.17), which becomes upper triangular. In this milieu, the first precoded user, $\pi_n(1)$, will see $K-1$ interferences because its precoded signal is itself as the others' precoded signals are unknown; whereas the last precoded user will not see any interference so it will see a ZF channel. However, the power allocation technique is expected to place more energy to the first user rather than the last because the latter is a big interferer. This means that, though user $\pi_n(k)$ sees more interferences than user $\pi_n(k+1)$, more power will be allocated to the former so that the effect is balanced. The consequence of this dissertation is that the feedback filter expression for the MMSE-ZF is the same as the MMSE: from (14.4) we see that the feedback filter resulting from the MSE minimization (15.15) is already a ZF condition since it tries to invert the channel seen from the precoded signal to the received signal, but it cannot fully inverted due to its structure.

Finally, the receive parameter g_n is the only that plays a part in this discussion. From the mathematically point of view, we see that g_n could only weight the resulting upper triangular matrix, neither make all the values be the same. However, since neither $\boldsymbol{F}_n$ nor $\boldsymbol{B}_n$ depend on it and its value does not affect the performance, MMSE and MMSE-ZF are equivalent problems. Nonetheless, g_n has a different value in the MMSE and in the MMSE-ZF optimizations, which we recall here with an illustrative aim. The MMSE g_n is obtained through the orthogonality principle by differentiating

ε^2 with respect to g_n. That is, by taking the trace operator in (16.1):

$$g_n^{MMSE} = \frac{\mathrm{tr}\left(\boldsymbol{P}_n^T \boldsymbol{B}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H \boldsymbol{H}_n^H\right)}{\mathrm{tr}\left(\boldsymbol{H}_n \boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H \boldsymbol{H}_n^H + \boldsymbol{R}_{\boldsymbol{n}_n}\right)};$$

notice that it has form of typical MMSE receive gain since, recalling the expression of the feedback filter, the numerator expresses the useful signal energy while the denominator reflects the sum of signal energy plus noise. That is analog to $S/(S+N) = \mathrm{SNR}/(\mathrm{SNR}+1)$. This helps us to deduce this parameter for the ZF, making $\mathrm{SNR} \to \infty$, to neglect the noise correlation part in the former expression to get:

$$g_n^{ZF} = \frac{\mathrm{tr}\left(\boldsymbol{P}_n^T \boldsymbol{B}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H \boldsymbol{H}_n^H\right)}{\mathrm{tr}\left(\boldsymbol{H}_n \boldsymbol{F}_n \boldsymbol{R}_{\boldsymbol{v}_n} \boldsymbol{F}_n^H \boldsymbol{H}_n^H\right)}.$$

16.3 QR Decomposition

Lastly, as the QR decomposition has been exploited for the single-user SISO and MIMO channels, it deserves some lines for the multi-user scenario though, as it will be proofed, it has no practicality. The QR decomposition decomposes the channel matrix into a unitary matrix and a lower or upper triangular matrix. This approach, however, as well as the SVD, requires the channel to be attacked by both transmitter and receiver sides with equal capability. Assume that $\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{R}$, where $\boldsymbol{Q}$ is the unitary matrix and $\boldsymbol{R}$ the triangular matrix.

For the single-user uplink and downlink channels, both transmit and receive filters are completely cooperative so the QR decomposition is a good tool to be used in iterative signal processing techniques due to the lower triangularity or matrix $\boldsymbol{R}$. For example, a single-user link with DFE and V-BLAST at the receiver, the transmitter —both up and down links— will prepare (rotate) the information according to $\boldsymbol{Q}$ and the receiver, which performs the iterative detection through DFE or V-BLAST, will follow the coefficients of the lower triangular matrix $\boldsymbol{R}$. For a single-user link with THP at the transmitter, the papers are switcher and, as done in sections 6.6 and 11.3, the feedback THP filter follows the lower triangular matrix while the receiver rotates the received signal according to the unitary matrix, preserving the noise statistical properties as it is unitary.

This means that, in general, the QR decomposition is only applicable in point-to-point channels with the structures defined in the former paragraph. For the multi-user links channels, there is only one side with full processing capabilities: the transmitter in the downlink and the receiver in the uplink. As a consequence, QR has been used in multi-user subsystems where joint signal processing is done. For example, in DSL, the use of central offices (COs) and optical network units (ONUs) with the aim of shorten the loop gives the opportunity of perform signal processing at the CO/ONU under QR-like scheme for both up and down links, as reported in [GC02]. Another example would be the use of relay stations in cellular networks (e.g. [CTHC08]), in which the link between the base station and the relay station is a point-to-point MIMO link, for which both station can employ joint signal processing to prepare, transmit and detect the signals of all the users. However, the channel between the relay station and the mobile terminals is, again, a BC, for which the use of QR has no sense.

Chapter 17

Performance Analysis

This last chapter is intended to show the performance of the THP system proposed in this thesis for the multi-user BC. MATLAB simulations have been done by randonly transmitting QPSK-modulated symbols through the $T \times K$ BC channel with $\rho_T = 0.2$, the channel characteristics defined by the rich multi-path environment described in table 7.1 of the HIPERLAN radio network standard in [MS98] using the diversity model (9.5) for each user's MISO channel. Each user receives the same symbol in the donwlink repeated onto the N OFDM carriers, with the aim of achieving maximum diversity in frequency —which would be analogous to being able to constructuvely receive the same symbol through different paths in time.

The performance indicators are the normalized mean-square error (NMSE) and the SER, as well as the user rate, R_k defined by (15.2), in terms of channel condition, number of active users or other system parameters; the three of them statistically obtained using Monte-Carlo simulations. The optimization techniques studied are the MMSE and the MMSE-Cholesky, since the MMSE-ZF is equivalent as justified in 16.2 and QR is not possible in an uncooperative scenario.

In multi-user systems we define the average SNR as

$$\text{SNR} \triangleq \frac{\mathcal{E}}{\text{tr}(\boldsymbol{R}_n)},$$

that is the energy allocated along the T transmitting antennas and the N carriers versus the noise power present at the K one-receiving antennas users at each one of the N carriers. The SNR at each user would be defined similarly taking into consideration only the fraction of allocated power to that user divided by both the interference that it receive due to the other users and the noise present at its receiver, i.e., a SINR.

17.1 Power Allocation Algorithm

Seeing that the distribution of the available energy among the user plays an extremely essential paper in multi-user system, this first section focuses on the behavior of the power allocation described in section 15.2. This analysis is divided into three parts. Firstly, the convergence of the proposed iterative algorithm is exemplified for the case with two users and one carrier in terms of allocated power for each user at each iteration as well as the resulting rate for each user. This will allows us to define the number of needed iterations to ensure convergence in the two users case (cf. to [VJG03]). Secondly, the three user iteration algorithm is studied to check its convergence as well to provide an example

of user selection. Finally, the resulting power distribution will be illustrated in a system with three users sharing two carriers with different channel conditions.

17.1.1 Convergence of the Iterative Water-Filling Algorithm

The water-filling algorithm depicted in 15.2 is reasoned in the case with two users, namely $K = T = 2$, and one single-carrier $N = 1$. We assume that the users are correctly ordered under the best-last paradigm so that $\boldsymbol{O}_n = \{\pi_n(1) \dots \pi_n(K)\}$ with

$$\xi_{n,\pi_n(1)} \leq \dots \leq \xi_{n,\pi_n(K)};$$

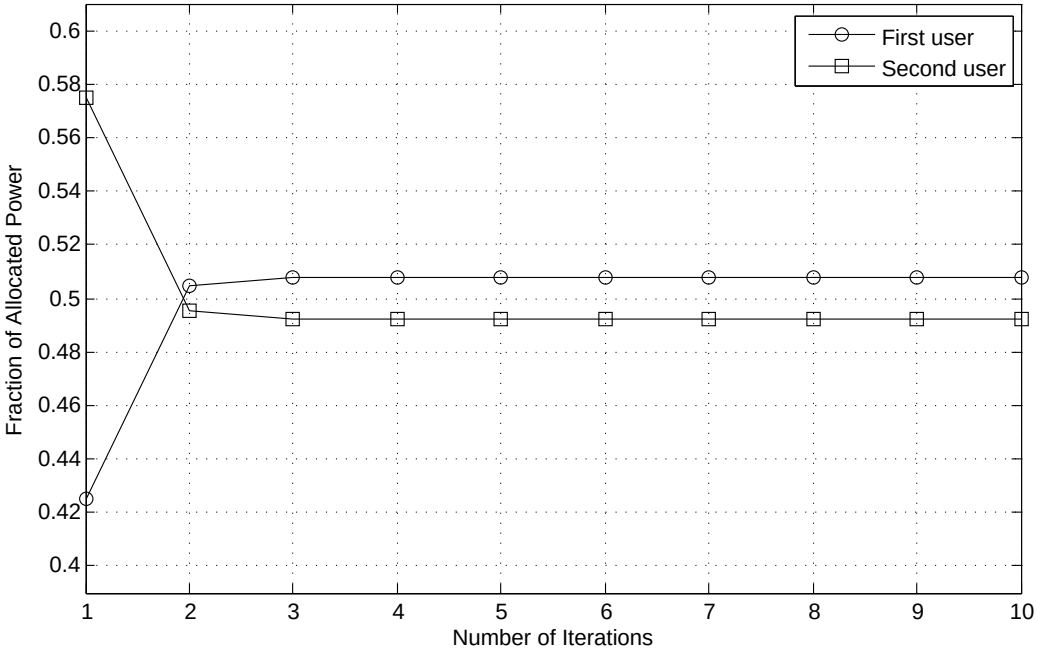
which in this case we refer to $\pi_n(1)$ as the first and $\pi_n(2)$ as the last user.

Figure 17.1 shows the convergence of the algorithm for two users in terms of allocated power and associated rate. In the first iteration, the algorithm assumes that each user has no interference from others because the dual uplink variances are initialized at zero. This can be seen in the figure in two senses: first, more power is allocated to the last user as it has better channel condition and, second, the rate achieved by the second user is apparently much better than the first user. Amazingly, the algorithm puts things in order with the second iteration as the calculation of the effective channels for the both users makes the system realize that they are not alone. As a consequence, the algorithm detects the last user as a potential interferer to the first user and, despite its better channel conditions, penalizes its allocated power; whereas this portion is allocated to the first user who must cope with more interference. Since there are only two users, the algorithm converges in three or four steps. The second graphic in figure 17.1 reflects similar results: on the one hand, the second user at the end of the algorithm has less rate capacity because it has been penalized by the fact that it is a potential interferer to all other users and, on the other hand, the first user, despite having worst channel conditions, improves its rate capacity because more power is allocated to it. An important point is that the user with better channel conditions usually obtains the best channel rate, as it was also obtained in figure 15.1. It is the case where both user have very similar conditions that the fact that the first user has more power allocated arises to obtain better rate capacity than the last user. Another aspect to remark is the one of the user ordering. Since the users have been ordered through best-last strategy, the rate capacity of the users with best channel conditions tend to cease while at its turn the users with worst channel conditions, as they are first precoded, perceive a rate capacity gain. This is seen in figures 15.1 and 17.1 as the rate capacities of the two users approximate to each other. In a best-first strategy, the effect would be the contrary and the rate capacities would separate from each other.

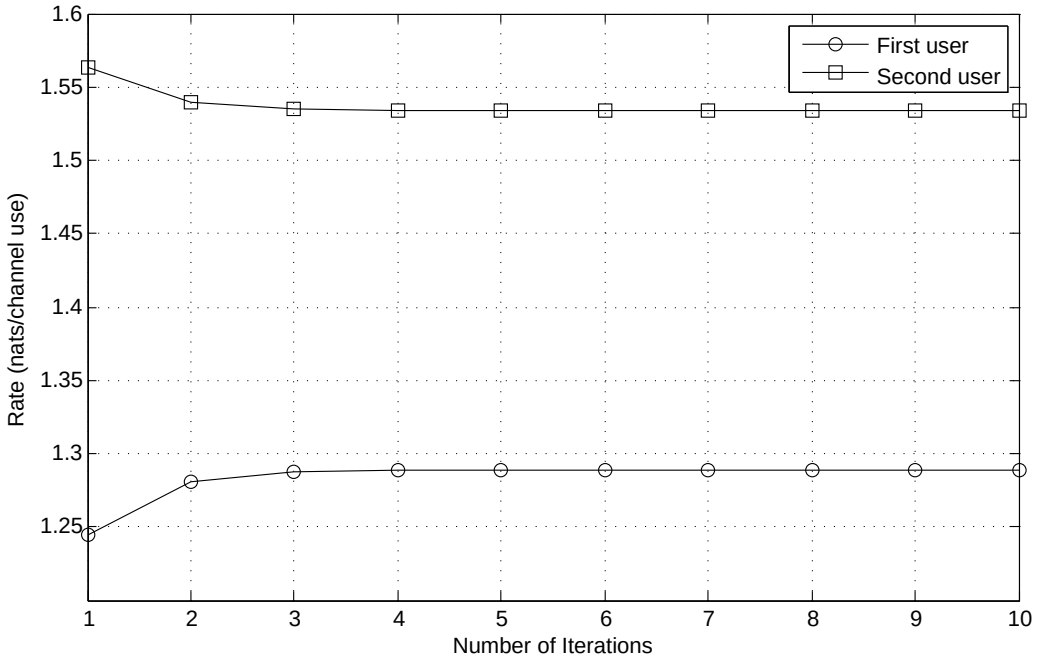
17.1.2 User Selection

The concept of user selection reflects the fact that the power allocation algorithm might select only a subset of $K_n \leq K$ users at the n-th carrier to be simultaneously transmitting. This is analogous to the MIMO channel in part III under the MMSE power allocation policy in which the system might omit some carriers —the worst— in the transmission of information.

It is then interesting to discuss is how the algorithm selects users. In order to illustrate it, two simulations have been done with a system with $T = K = 3$ users with the same overall SNR — this means that now we have more noise, so the transmitted power is accordingly higher to be fair in comparison— depicted in figure 17.2. From it, we deduce several behaviors. Firstly, comparing



(a)



(b)

Figure 17.1: Fraction of allocated power (a) and associated rate (b) for first and last users, with $\xi_{n,\pi_n(2)} = 2\xi_{n,\pi_n(1)}$ at a given carrier, versus the number of iterations in the water-filling algorithm, at an average SNR = 0 dB, in a 2×1 MISO BC.

$\mathcal{E}_{n,k}/\mathcal{E}$ — Rate	$n=1$	$n=2$	Both carriers
User 1	0.23 — 0.79	0.13 — 0.78	0.36 — 1.57
User 2	0.32 — 1.84	0.07 — 0.46	0.39 —2.30
User 3	0.12 — 1.10	0.13 — 1.50	0.25 — 2.6

Table 17.1: Fraction of power allocated, $\mathcal{E}_{n,k}/\mathcal{E}$, and user rate in nats per channel use for $K=3$ users in a system with $N=2$ available carriers, at an average SNR = 0 dB in a 3×1 MISO BC.

graphics (a) and (b), we deduce that the algorithm works better when the user ordering is established by very different values of channel conditions, namely $\xi_{n,k}$. This means that though the best user is penalized and the worst user is improved, if they are separated enough in the first iteration, they do not cross. In this sense, the algorithm converges quicker. Secondly, if the users' channel conditions are very similar, as they start in very closed positions, the algorithm might be slower to converge. Even, if some user approximates to the 0 of fraction of allocated power and this user comes from an upper position —which is the case, for instance, of the last user in graphic (b) of figure 17.2—, the algorithm could allocate no power to that user and, as a result, this user would not use this band.

17.1.3 Power Allocation in Space and Frequency

A simple and representative simulation on the power allocation algorithm has been done for the case of $K=3$ active transmitting users along $N=2$ available OFDM carriers. Notice that the general strategy in this two-dimensional problem is to firstly perform frequency water-filling to find the optimal fraction of power allocated to each band, namely $\mathcal{E}_n$ for $1\leq n\leq N$; and, after, distribute this power into the user axis. As the user ordering is given by the channel conditions of each user and since we assume complete orthogonality among carriers, the users transmitting at one carrier no more interfere other users at the rest of the carriers so the user ordering is complete defined with the set $\{\xi_{n,k}\}_k$ at a given carrier. As a consequence, the user ordering at one carrier might differ from the user ordering at another carrier. This is translated to having different permutation matrices $\boldsymbol{P}_n$, which, as stated in the former chapter, are exclusively used and defined by the transmitter.

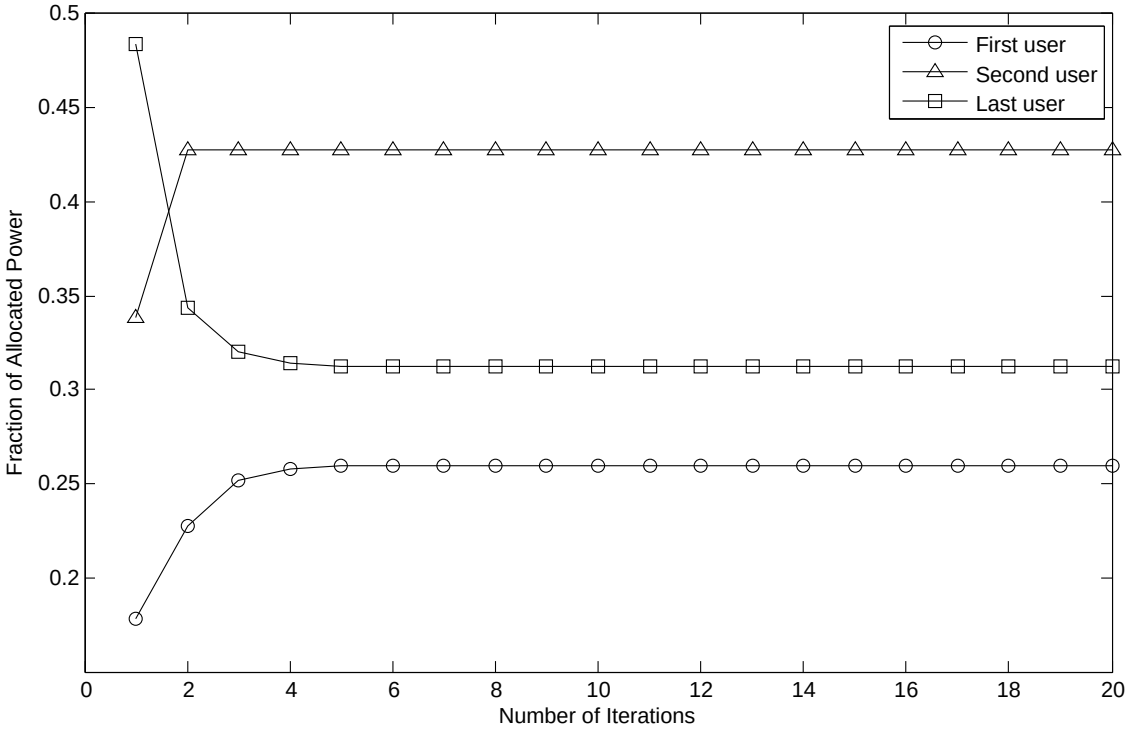
The scenario proposed in this subsection is the one with best frequency response for the first carrier. This means that, with high probability, $\xi_{1,k}\leq\xi_{2,k}$; but it does not implies any order. In general, at each carrier, the algorithm's performance is similar to the one studied in subsection 17.1.1, in which the users with best channel conditions (the last ones) will tend to cease power to other users, while the users with worst channel conditions (the first ones) will tend to receive more energy than the allocated in the first iteration. However, the set of ordered fractions of allocated power might vary from the user ordering. In other words, the user order stated by the set of performance parameters $\{\xi_{n,k}\}_k$ at each carrier defines the order for all the iterations, as the leitmotiv of the algorithm is how users interfere each other, so the order cannot be changed during the algorithm. At the first carrier, the set of performances read

$$\xi_{1,1}>\xi_{1,2}>\xi_{1,3},$$

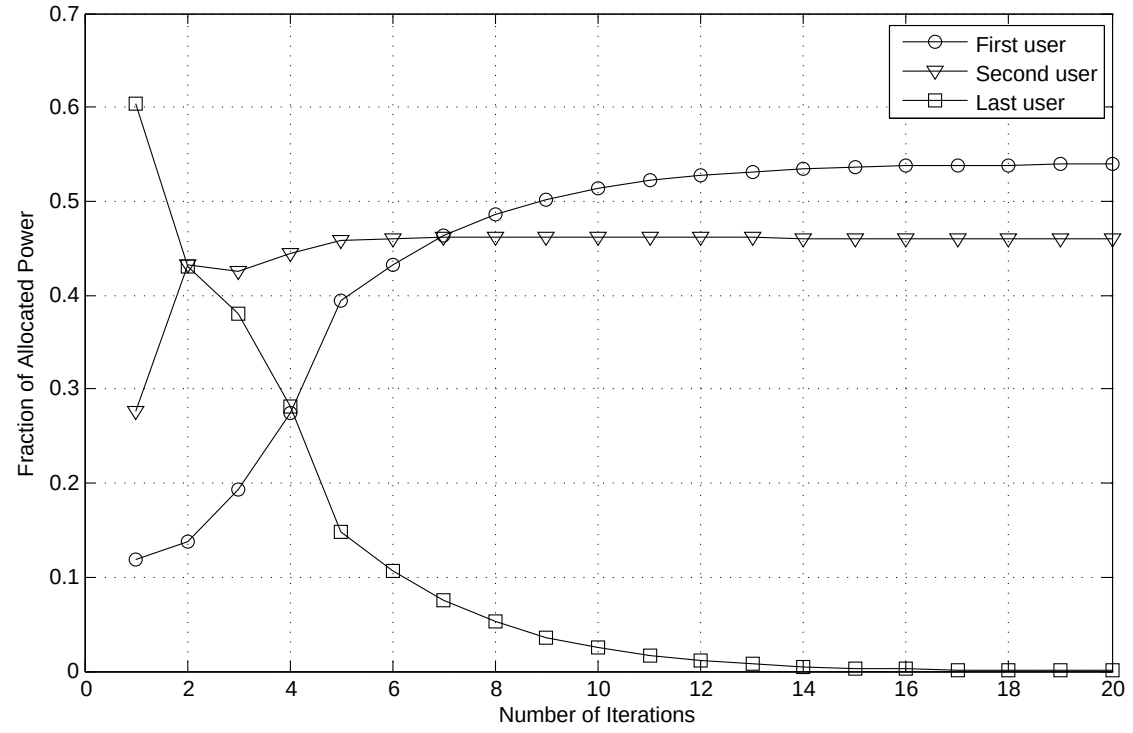
while in the second carrier user 2 is the one with better channel conditions —it is worthy to remark that the parameter xi depends on the channel frequency response, which depends at its turn on the multiple paths seen but the considered user, and the noise at the receiving antenna—, with

$$\xi_{2,2}>\xi_{2,1}>\xi_{2,3}.$$

Firstly, the water-filling through the frequency axis depicted in equation (15.5) allocates more



(a)



(b)

Figure 17.2: Fraction of allocated power for first, second and last users, with (a) $\xi_{n,\pi_n(1)} = \xi_{n,\pi_n(2)}/3 = \xi_{n,\pi_n(3)}/9$ and (b) $\xi_{n,\pi_n(1)} = \xi_{n,\pi_n(2)}/2 = \xi_{n,\pi_n(3)}/3$ at a given carrier, versus the number of iterations in the water-filling algorithm, at an average SNR = 0 dB, in a 2×1 MISO BC.

power to the first carrier rather than the second one. Specifically, $\mathcal{E}_1 = 1 - \mathcal{E}_2 = 0.67\mathcal{E}$. This has sense from both the information theory and MMSE point of views as seen in the MIMO channel in part III, as more power is allocated to the best resources.

Secondly, table 17.1 summarizes the result of the power allocation algorithm with 50 iterations —enough to ensure convergence, e.g. [VJG03]— for the three users along the two carriers, as well as the user rate achieved. As expected, the user ordering in the second carrier will be $\mathcal{O}_2 = \{3, 1, 2\}$, while in the first carrier $\mathcal{O}_1 = \{3, 2, 1\}$ under the best-last paradigm. We shall notice, again, that the power allocation algorithm is fair among the users. In this sense, user 3 and user 2 at the first and second carriers respectively, are the users that receive less energy at these bands. As always, the achieved rate for each user has contribution of the allocated power to that user and the position it has at each carrier as it will see the not yet precoded users as interferences. The third column depicts the overall performance for each user, in which it is seen that user 3 achieves more capacity while user 1, as it receives more interference at both carrier, achieves less capacity.

17.2 Capacity Region of the BC

With the aim of exemplifying the concept of capacity region, a simulation has been done for the $K = 2$ users 2×1 MISO BC with an average SNR of 0 dB, whose result is depicted in figure 17.3. The figure plots the set of possible simultaneously achievable rates for user 1 (first precoded) and 2 (last precoded) under the best-last paradigm, for which $\xi_{n,2} > \xi_{n,1}$ in the selected carrier.

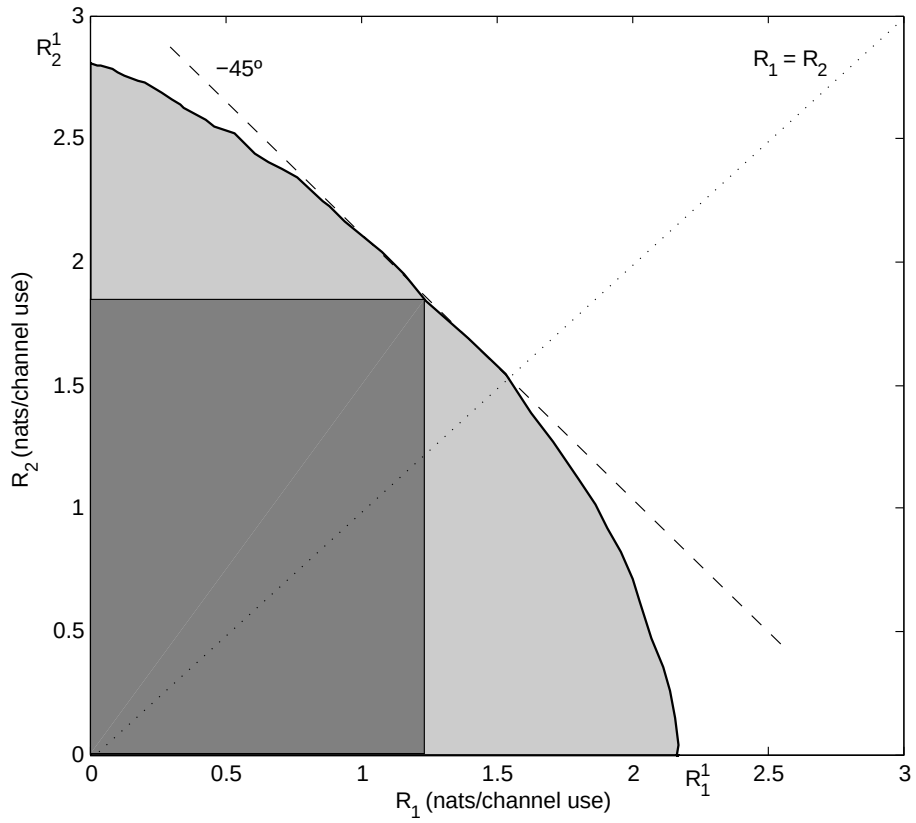


Figure 17.3: Capacity region for a $K = 2$ users BC channel with $\xi_{n,2} = 2\xi_{n,1}$ at the selected carrier with average SNR = 0 dB, in nats per channel use.

The black bold line defines the boundary of the capacity region of the BC, a convex hull in the K-space. This boundary is found by maximizing the sum of weighted achievable rates, i.e., determining the power allocation policy —in this case the transmitting variances $\mathbf{\Sigma}_{n,1}$ and $\mathbf{\Sigma}_{n,2}$, or, equivalently, $\boldsymbol{F}_n$— that maximize the scalar $\sum_k \alpha_k R_k$, where R_k is defined as (15.2) with the proper user order, in this case $\mathcal{O}_n = \{1, 2\}$. The result is light gray shadowed in the figure. As expected, it is a convex area in the 2-D plane and, from it, one may find several interesting points.

The first pair of points easily derived from the capacity region are the single-user rates, detailed in subsection 13.3.2, which correspond to the case where there is only one user transmitting. When one user is not transmitting, the other is able to transmit at a rate point located on its axis —for the second user, for instance, all the points located on the Y axis—, so the single-user rate is the maximum achievable rate, namely R_1^1 and R_2^2 in the figure and clearly with $R_2^1 > R_1^1$.

Secondly, with simultaneously transmission of both users, the rate of one user must decrease to offer capacity to the other user, fact that assures the convexity of the region. The capacity boundary depicts the achievable pair of rates for varying the pair of alphas, and follows the equation $\alpha_1 R_1 + \alpha_2 R_2 = \kappa$, where κ is a constant. As a result, the gradient of the capacity boundary becomes $\frac{-\alpha_2}{\alpha_1}$ and, in the case of sum-rate capacity —remember that this is the case studied in this work, for which $\alpha_1 = \alpha_2$—, the slope reads -1, as appreciated in figure 17.3. This means that the sum-rate capacity can always be found by finding the point in the capacity boundary that is tangent to the -45º slope line —dash line in the figure. If the users' channel conditions are very diverse, the sum-rate capacity will tens to favor the user with better channel conditions, as the -45º tangent will be located near to that user's single-user rate.

Thirdly, anther interesting point to discuss is the one with $R_1 = R_2$. This point can be found by tracing the $R_1 = R_2$ line —dotted line in the figure— and find the intersection with the capacity boundary. This strategy ensures that both users achieve the same performance but, if a user has very bad channel conditions —this is translated to a tight convex area—, the other user will have to transmit a lower rate to meet the requirement, since the intersection, in this case, will approximate to the user with worst channel conditions.

Finally, for a given power distribution, the capacity region becomes a square for the BC. This is represented by a dark gray square in figure 17.3 for the maximum sum-rate strategy. The water-filling algorithm gives the optimal power allocation to maximize the sum-rate capacity, that is, achieving the vertex of the dark gray square, which corresponds to $\alpha_1 = \alpha_2$. Figure 17.4 shows the evolution of the capacity region when increasing the transmitted power, whose evolution is predictable: the more transmitter power is available —or higher SNR—, the wider is the capacity region as the achievable rates increase with the SNR.

It is worthy to point out that, when selecting another user ethics strategy, the key resides in determining the set of priorities, $\{\alpha_k\}_k$ that reflects the strategy. For the case of maximum balanced rates, first the set of alphas that accomplish the requirement (13.3) is calculated and, after that, the power is distributed among users and carriers to maximize the corresponding rate (cf. [SVL05] for this methodology).

17.3 Single-Carrier Performance

This section analyzes the THP for the MISO BC, particularizing with $N = 1$. This is the widely studied scheme in the literature, as shows how the THP purely copes with the interference caused by users sharing frequency and time. The use of more carriers can be viewed from a design point of

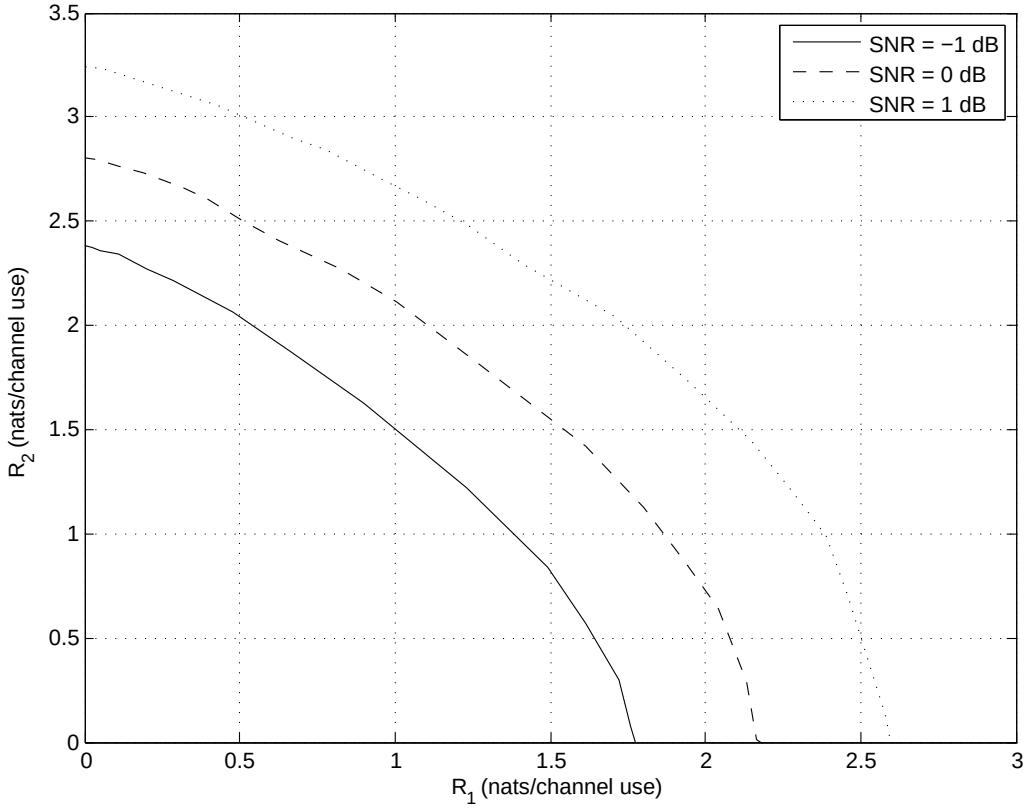


Figure 17.4: Capacity region for a $K = 2$ users BC channel with $\xi_{n,2} = 2\xi_{n,1}$ at the selected carrier, evolving with the average SNR, in nats per channel use.

view, as the system can place other users to that carriers or the same user to enhance the system's diversity. The three performance indicators —SER, NMSE and rate— have been obtained in a system with $T = K = 2$ users and $T = K = 3$ users; each one with users showing similar channel conditions, i.e. $\xi_{n,k} \approx \xi_n$; and, in second place, having different channel conditions. The first case reflects the fact that, though the first user is the one that receives more interference, it is the protected user by the system because, seeing that it does not interfere, more power can be allocated to it.

17.3.1 Two Users Case

Similar Channel Conditions

Figures 17.5 and 17.6 show the user rates, SER and MSE for a 2 user system equipped with 2 transmitting antennas, versus the average SNR from -2 to 8 dB, when the channel conditions of both users —channel gain/attenuation and noise— are similar; in other words, $\xi_{n,1}$ and $\xi_{n,2}$ are statistically equal in average. Despite the user ordering it is assumed to be $\mathcal{O}_n = \{1, 2\}$, in the sequel we will always refer to the first and last users; or first, second and last for the three users case.

The first coherent result is the rate obtained for both users. In this particular case where both users have similar channel conditions, the capacity regions is expected to be symmetric to the line $R_1 = R_2$ —please refer to figure 17.3, having $R_1^1 = R_2^1$— and, as a result, the maximum sum-rate capacity, where the capacity boundary is tangent to the -45º slope is the same as the intersection of the boundary and the line $R_1 = R_2$. This means that, though the first user sees interference, the power allocation algorithm places more power to this user to cope with the interference it receives

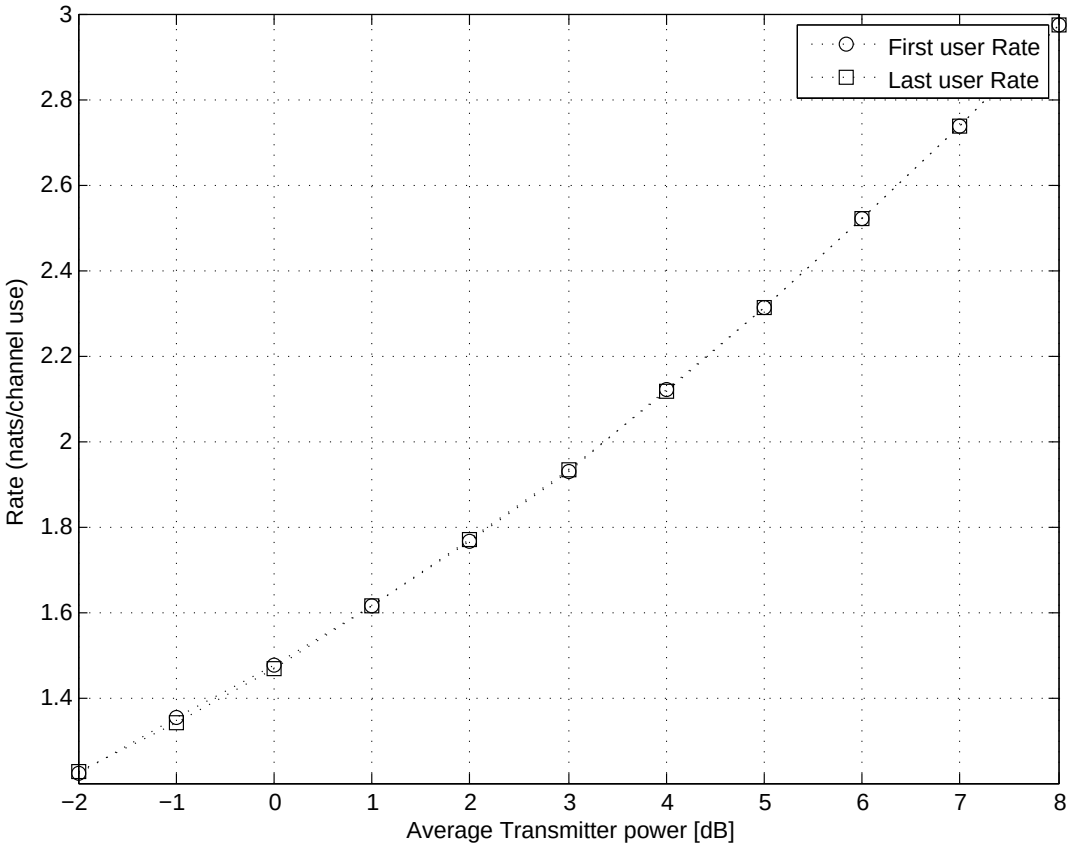


Figure 17.5: User rates for the $K = 2$ users $N = 1$ carrier THP 2×1 MISO BC versus average SNR, for each user, with similar channel conditions $\xi_{n,1} \approx \xi_{n,2}$, in nats per channel use.

—the second user—, while the second user is placed less energy since it is interference free. The balance is that, in this case, both achieve the same rate at high SNR.

Regarding the achieved SER and MSE by the first and the second users in figure 17.6, two considerations can be obtained.

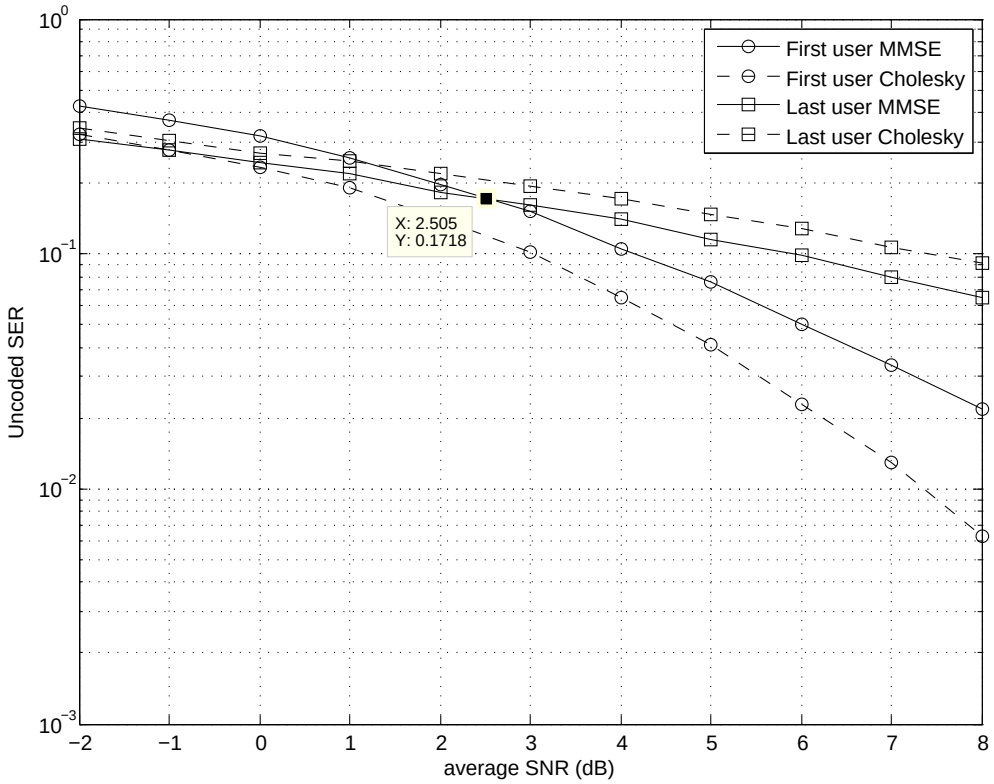
First of all, both MMSE and Cholesky —remember that both differ only in the expression of the feedback filter, i.e. equations (15.15) and (16.3)— have similar behavior versus the average SNR for the two users but, surprisingly, the MMSE solution is more conservative rather than the Cholesky solution, as the protected user —that is, the first user— performs better with Cholesky but the second user achieves worst SER and MSE with Cholesky. In the two user case, this is due to the value interference pre-subtraction of the first user to the second user, i.e. the element $(2,1)$ of matrix $\boldsymbol{B}_n$, $b_{n,2,1}$. On the one hand, the MMSE solution performs a complete subtraction as the value of the elements of the feedback matrix are exactly selected to perform complete interference pre-subtraction, as seen in equation (15.16). This means that the system spends more energy to assist the second user, leading to a fairer solution. On the other hand, the Cholesky factorization obtains the value of $b_{n,2,1}$ by the factorization of the design matrix which gives a lower triangular matrix. Since the design matrix is Hermitian, the information contained —which is nothing more than the inverse of the channel conditions for each user— can be represented with a triangular matrix, $\boldsymbol{L}_n$. However, the inverse of this matrix does not perform complete interference pre-subtraction, it minimizes the MSE as a whole. As a result, the system spends more energy, not to better precode the second user, but to precode it so that the impact that the second user will have to the first user is also minimized. In other words, $b_{n,2,1}$ is selected to reduce, not cancel, the interference that v_1 causes to s_2, but also to keep with bounds how the second user will interference the first one, as the transmitter knows the channel response from v_1 and v_2 to the users.

In second place, both SER and MSE plots reflect that the system has two behaviors: low and high SNR. At low SNR, the additive noise is more significant than the MUI and, as a consequence, the power allocation algorithm places energy to users proportionally to $\xi_{n,k}$. In this case, since both users have similar channel conditions, their performance is also similar at low SNR. In the next case with two users and diverse channel conditions, this fact is more representative. At high SNR, the additive noise can be neglected and the MUI plays the most important part. As expected, the system places energy proportionally to $A_{n,k}$. That is, more energy for the first user and less energy to the last user. Yet stated above, the user rates are equaled because the last user is interference free and the first user must cope with the interference caused by the last user. However, in terms of SER and MSE, the first user has better performance than the second user at a given SNR point. The SNR point that the SER of the first user crosses the SER of the last user is plotted, which would be a very desirable place to work and, if required, enhanced with the use of more carriers or a code.

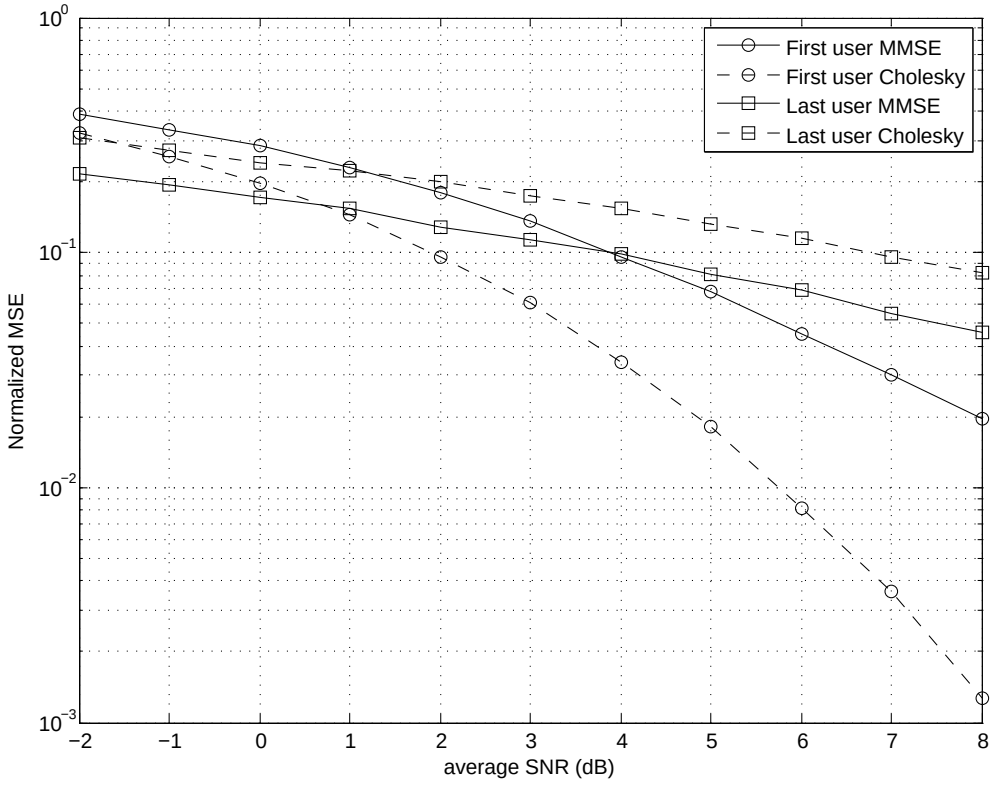
Diverse Channel Conditions

The same parameters are now obtained for the $T = K = 2$ user case with diverse channel conditions. Specifically, the second user —the last precoded— has a channel gain 3 dB larger than its partner. The consequences of this change can be appreciated on figures 17.7 for the rates and 17.8 for the SER and MSE.

As expected, the last user achieves more rate capacity than the first user. The explanation is supported by the new capacity region, which is depicted in figure 17.3 with $R_1^1 < R_2^1$, whose boundary point with tangent -45º is closer to the last precoded user. This is more significant at the low SNR regime for the same reason as before: the MUI can be neglected in front of the additive noise, which is



(a)



(b)

Figure 17.6: SER (a) and NMSE (b) for the $K = 2$ users $N = 1$ carrier THP 2×1 MISO BC versus average SNR, for each user, with similar channel conditions $\xi_{n,1} \approx \xi_{n,2}$.

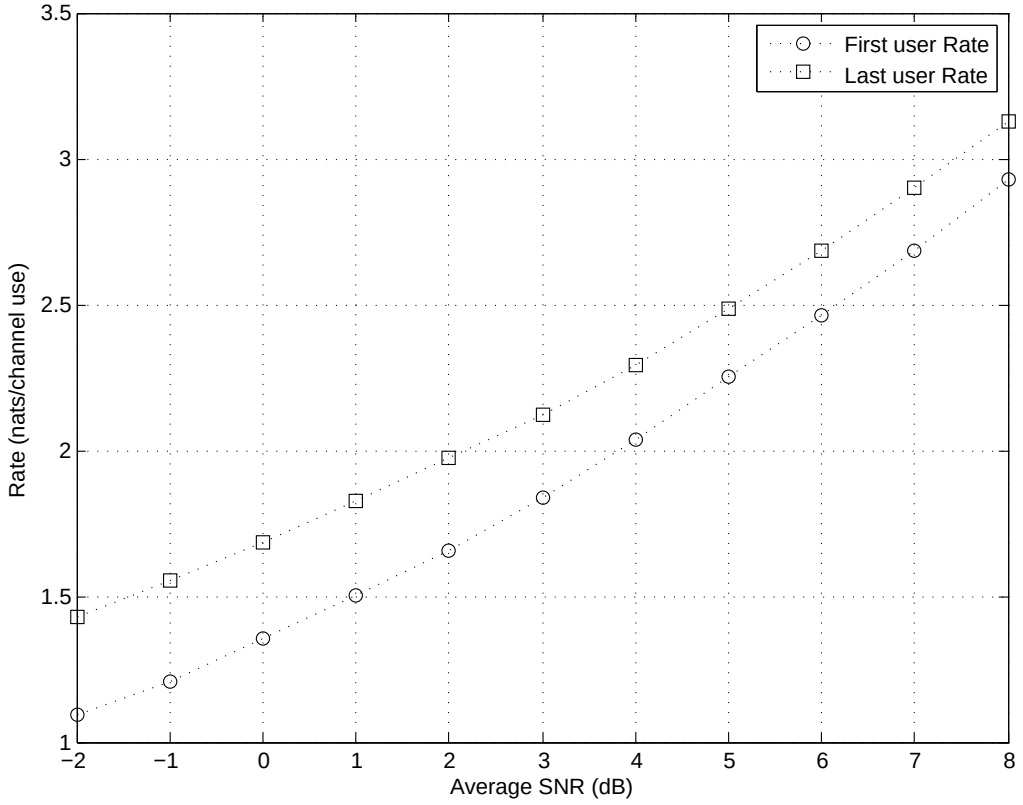
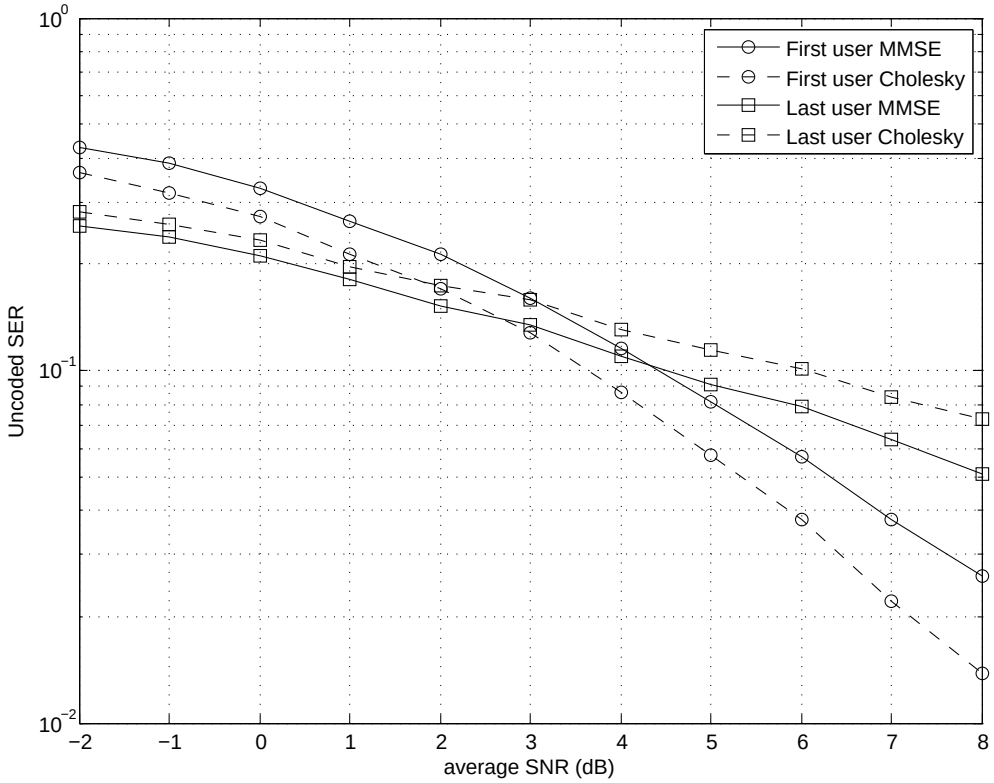


Figure 17.7: User rates for the $K = 2$ users $N = 1$ carrier THP 2×1 MISO BC versus average SNR, for each user, with diverse channel conditions $2\xi_{n,1} = \xi_{n,2}$, in nats per channel use.

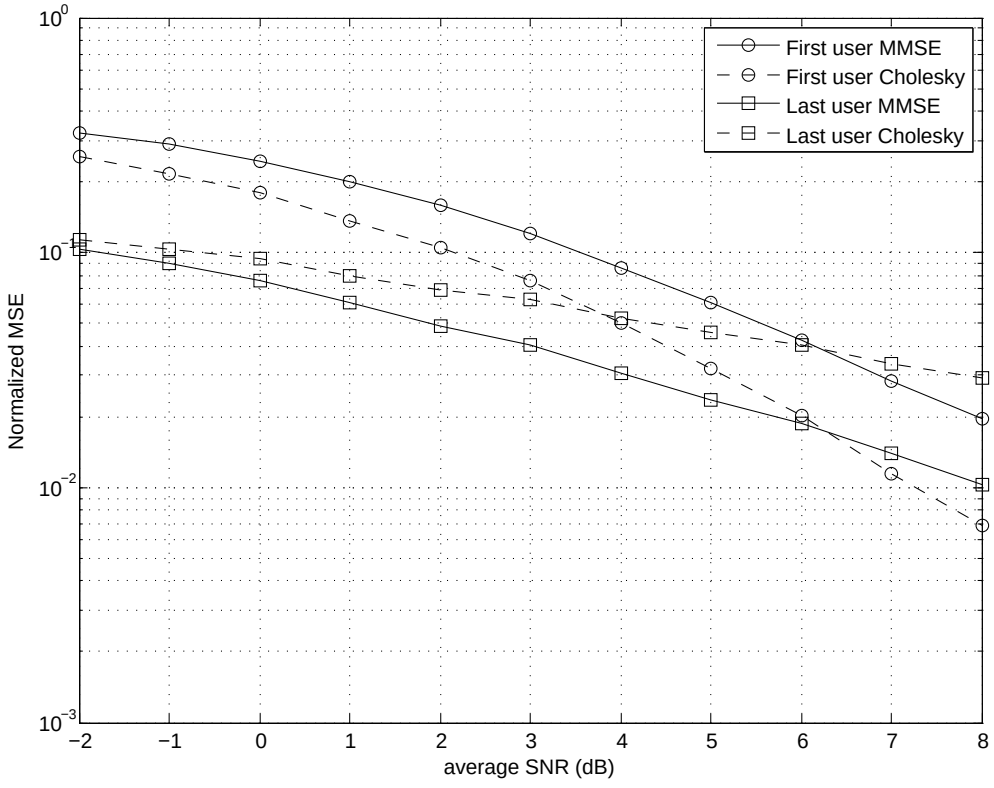
predominant. However, when increasing the SNR, the interference amongst the users increases with respect to the additive noise and the system tries to converge the rates. This means that the capacity region changes its boundary aspect with the SNR, so that the rate of the first user increases more rapidly compared to the last user. It can be slightly appreciated on figure 17.4 where, for the same SNR increment, the first precoded user's area timidly grows more than the last user's area.

Regarding the MSE and the SER in figure 17.8, both show the same conduct as studied in the former scheme: the MMSE solution is more conservative rather than the Cholesky factorization, and at high SNR the first user performs better compared to the second user. However, the more the channel conditions differ, the more both solutions approximate —see figure 17.9 for the SER in a system with $\xi_{n,2} = 10\xi_{n,1}$ to quantify this appreciation. This has sense since the value of $b_{n,2,1}$ in the Cholesky feedback filter approximates to the value of $b_{n,2,1}$ of the MMSE feedback filter when the channel conditions —noise and channel gain/attenuation— of the first user are significantly worst than the second user, because the best the last user channel conditions compared to the first one's, the less energy will be placed to the last user so the interference it has to cancel —that is, the first user precoded signal— is more relevant so it tends to cancel it all out.

Finally, it is worthy to highlight the fact that the point where the SER curves cross has increased in terms of SNR. The main reason is that now, at low SNR, the SER curves differ more than before and, therefore, need more SNR to cross. An immediate consequence is that the user order has a transcendent role and at the same high SNR as before, users' performance are more similar. This can be seen in both plots: on the one hand, the SER of the first user at an average SNR = 8 dB has increased, while the SER of the second user is quite better —compared to similar channel conditions—;



(a)



(b)

Figure 17.8: SER (a) and NMSE (b) for the $K = 2$ users $N = 1$ carrier THP 2×1 MISO BC versus average SNR, for each user, with diverse channel conditions $2\xi_{n,1} = \xi_{n,2}$.

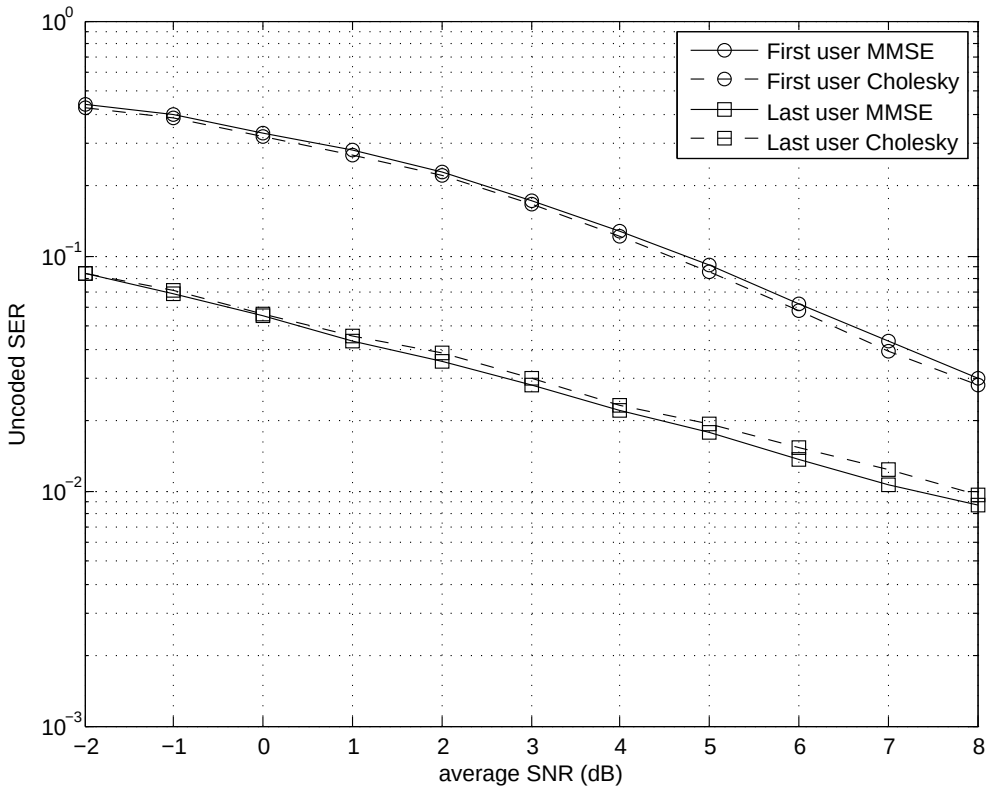


Figure 17.9: SER for the $K = 2$ users $N = 1$ carrier THP 2×1 MISO BC versus average SNR, for each user, with very diverse channel conditions $10\xi_{n,1} = \xi_{n,2}$.

and, on the other hand, the MSE plot is more compact.

17.3.2 Three Users Case

Since the following lines on three users are redundant to the two users case, this subsection provides the same study as the former subsection for the case of $K = 3$ simultaneously transmitting users in a system equipped with $T = 3$ transmitting antennas, at a selected carrier, varying the average SNR from -6 to 12 dB, with less detail.

With the simultaneity of three users, it is expected to require more energy to achieve similar performance, that is why the SNR range has been increased. The channel capacity region for three users is well-known to be a convex hull in the 3-dimensional space and it can be obtained analogous to the 2-user case.

Similar Channel Conditions

Figures 17.10 and 17.11 show the user rates, SER and NMSE for the three users in the case with similar channel conditions, namely $\xi_{n,1} \approx \xi_{n,2} \approx \xi_{n,3}$. The resulting plots are coherent when generalizing from the two users system previously analyzed.

On the one hand, the rate achieved by the three users is again the same because the capacity region is delimited by a surface whose intersection with the user axis is given by the values of the single-user rates, R_k^1, which are the same for the three users. The user rate achieved is obtained by the intersection of the convex surface with the line $R_1 = R_2 = R_3$, having the intersection point the plane $R_1 + R_2 + R_3 = \kappa$ as a tangent plane, with κ an irrelevant constant. Notice that the user rate

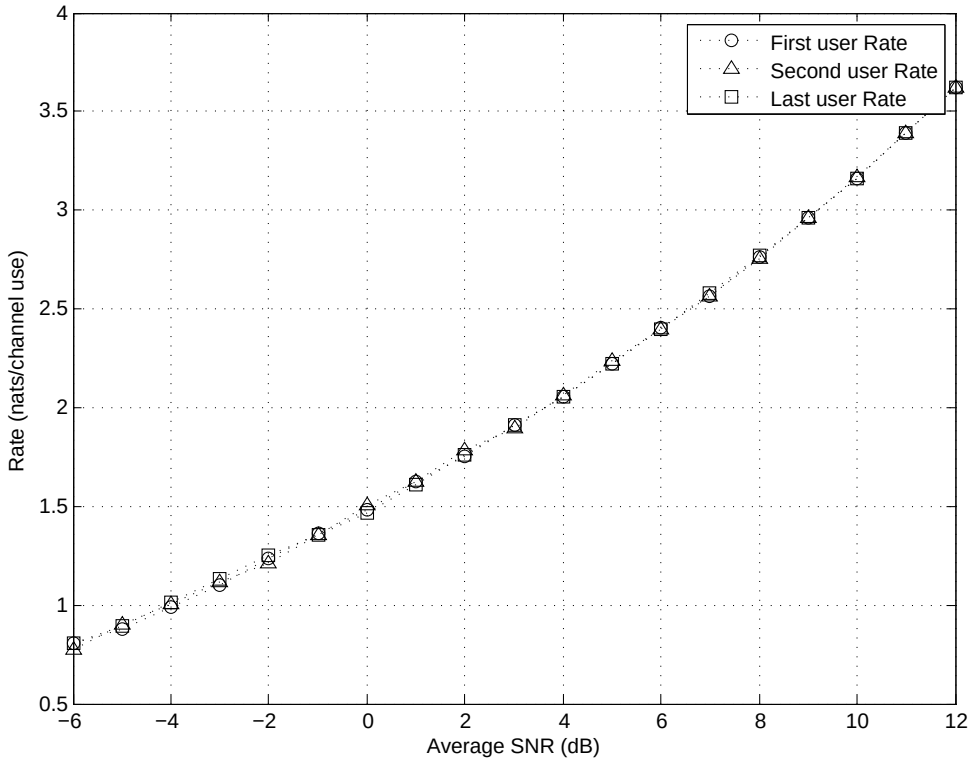


Figure 17.10: User rates for the $K = 3$ users $N = 1$ carrier THP 3×1 MISO BC versus average SNR, for each user, with similar channel conditions $\xi_{n,1} \approx \xi_{n,2} \approx \xi_{n,3}$, in nats per channel use.

at the same SNR point has been slightly reduced compared to the two users case. This is due to the same available energy has to be shared by three users, and not two.

On the other hand, the pair SER/MSE depict the same results as in the two users case: while at low SNR the performance is very similar for all the users, at high SNR the first precoded user out stands versus the others because more power is allocated to it. The Cholesky factorization treats better the first user, leading to a less fair solution, again.

Diverse Channel Conditions

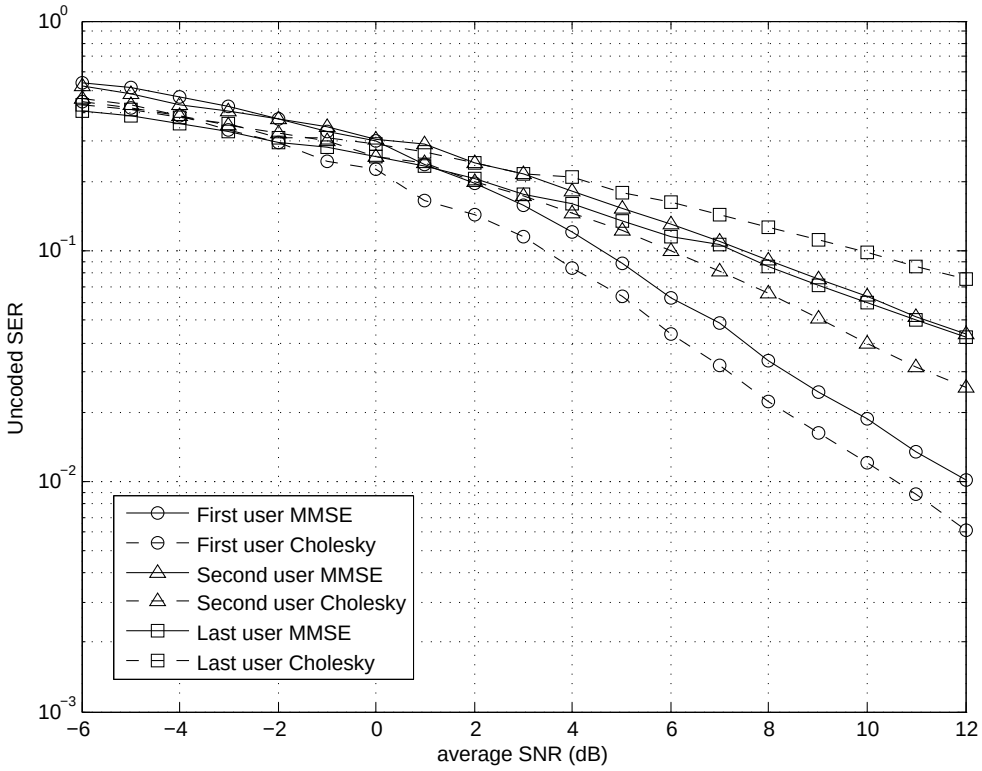
Figures 17.12 and 17.13 show the user rates, SER and NMSE for the three users in the case with different channel conditions, namely each user has 3 dB of better performance with respect to the previous one.

The rate distribution is consistent with the channel conditions: $R_1 < R_2 < R_3$. The difference is again more notable at low SNR regime, where the additive noise predominates in front of the MUI. For higher SNR, the parameter $A_{n,k}$,

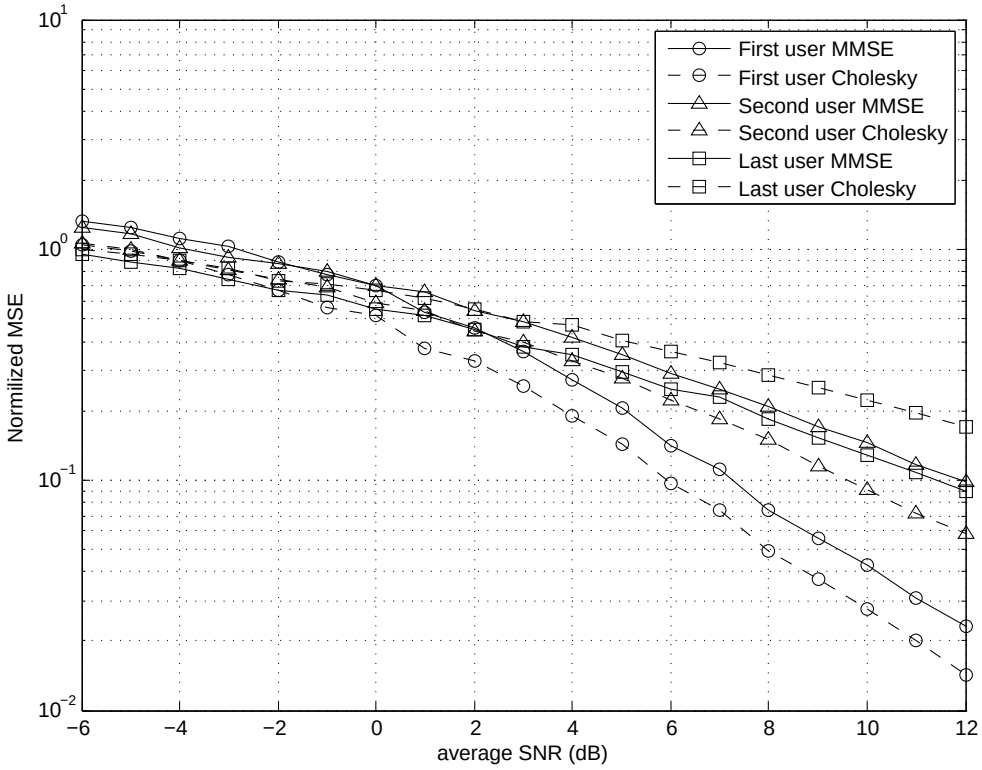
$$A_{n,k} \triangleq \left(\boldsymbol{h}_{n,k} \left(\sum_{j>k} \boldsymbol{\Sigma}_j \right) \boldsymbol{h}_{n,k} + \sigma^2_{n_{n,k}} \right),$$

is more affected by the interference term rather than the additive noise, for which the rate of the first user, though it has the worst channel conditions, is enhanced as it shows the higher $A_{n,k}$ of all three.

The same reasoning holds for the SER and MSE graphics: at low SNR the user that achieves best MSE and SER is the last precoded user since it has better channel conditions, but with higher SNR



(a)



(b)

Figure 17.11: SER (a) and NMSE (b) for the $K = 3$ users $N = 1$ carrier THP 3×1 MISO BC versus average SNR, for each user, with similar channel conditions $\xi_{n,1} \approx \xi_{n,2} \approx \xi_{n,3}$.

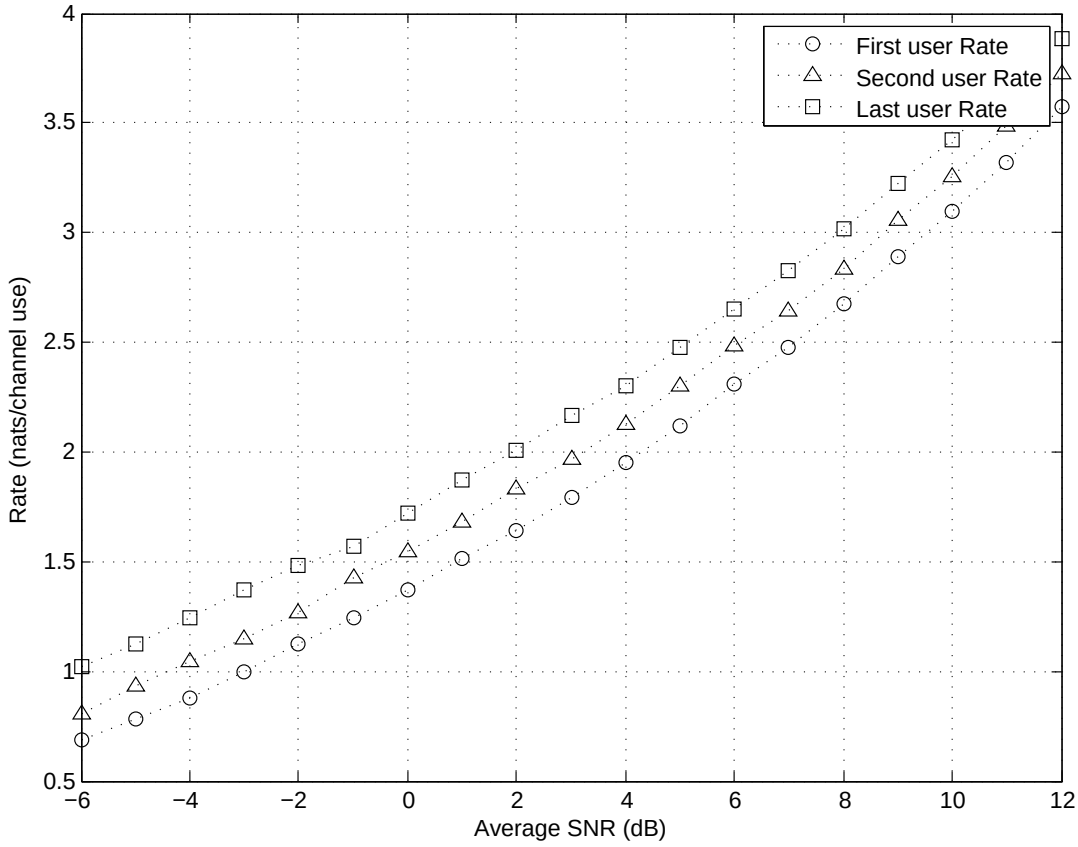
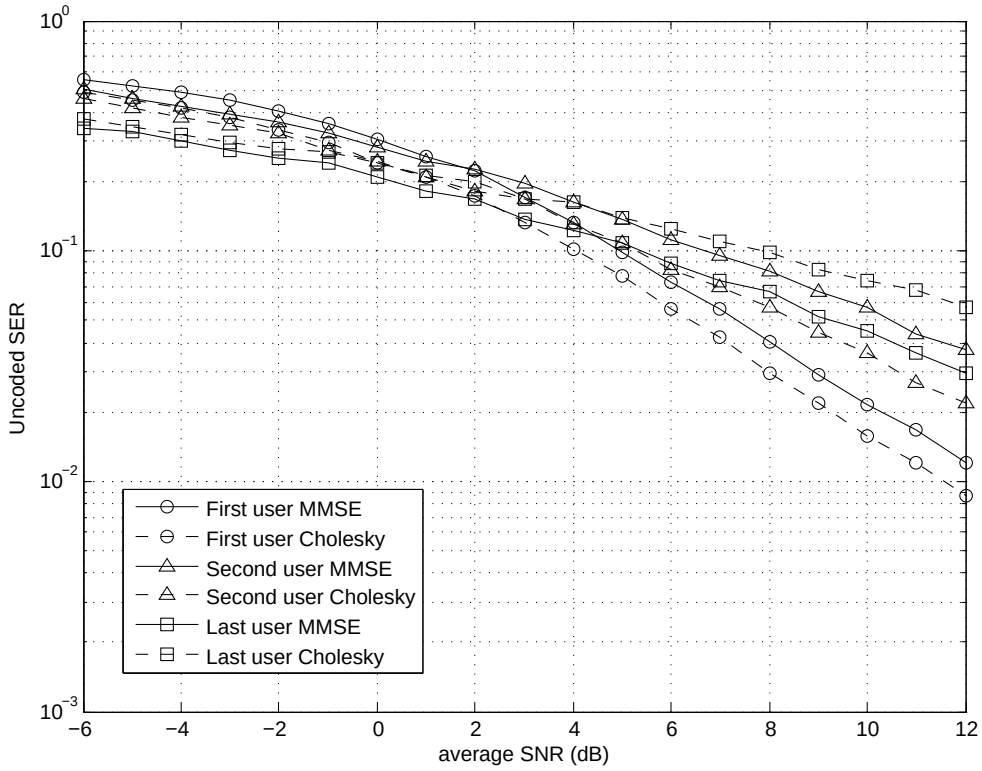
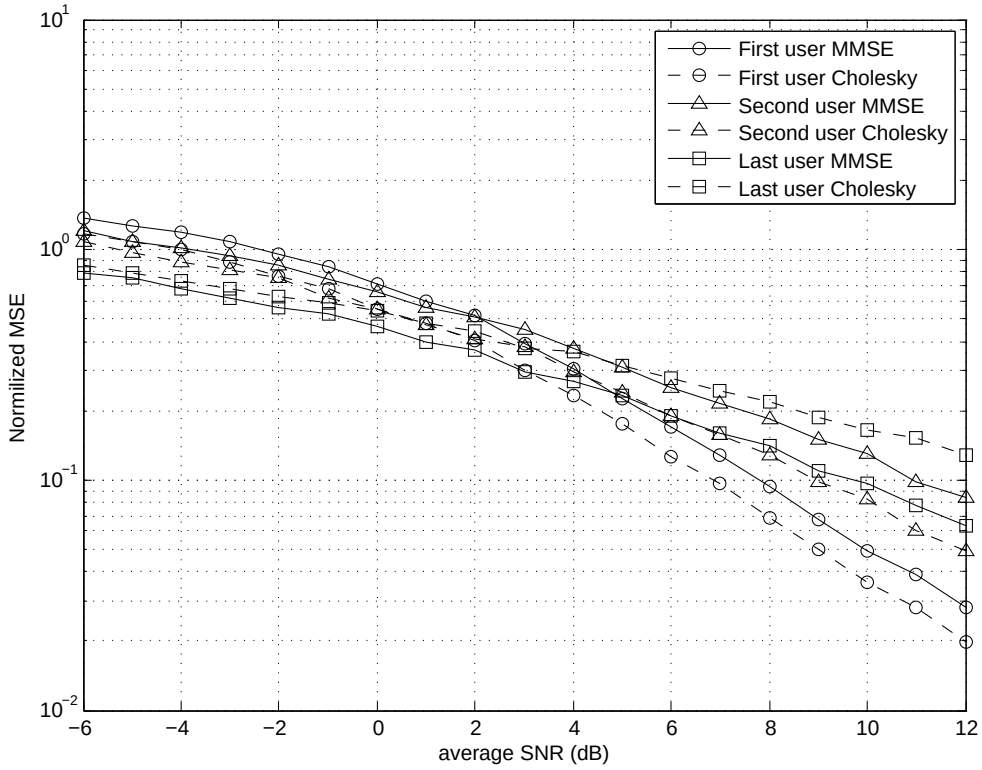


Figure 17.12: User rates for the $K = 3$ users $N = 1$ carrier THP 3×1 MISO BC versus average SNR, for each user, with diverse channel conditions $4\xi_{n,1} = 2\xi_{n,2} = \xi_{n,3}$, in nats per channel use.



(a)



(b)

Figure 17.13: SER (a) and NMSE (b) for the $K = 3$ users $N = 1$ carrier THP 3×1 MISO BC versus average SNR, for each user, with diverse channel conditions $4\xi_{n,1} = 2\xi_{n,2} = \xi_{n,3}$.

the papers are interchanged. Again, the point where the SER crosses depends on how different are the channel conditions among the users, as it defines how separated are the SER and the MSE at low SNR. Finally, the Cholesky factorization curves approach the MMSE with more proximity for diverse channel conditions than similar conditions.

17.4 Frequency Diversity

The use of OFDM in the THP system has two objectives: firstly, remove the temporal ISI by transmitting N symbols of larger duration, and secondly, providing a low cost solution —it is extremely cheap to implement the fast Fourier transform (FFT)— to achieve diversity of order N. How the system uses this diversity is a design aspect. As it happened in the point-to-point single-user approach, the user could use the N available carriers to transmit different information —maximizing the OFDM throughput— or use them to transmit less information in a more robust way —i.e. using a code or transmitting the same symbol. In a multi-user system something similar happens, as the system can use this N orthogonal bins to place more information, more users or none of these.

It has been insisted throughout the part that the Tomlinson-Harashima system is in charge of diminishing the MUI present when K active users —we understand by active that are constantly transmitting— share time and frequency, so they might be separated in space —namely spatially multiplexing. In this sense, THP has nothing to do among users that operate at different carrier, as long as they are considered to be orthogonal. For this reason and with the aim of illustration, as it has been done in the single-user MIMO channel, a system with $T = K$ users that share N frequencies use the frequency bin as frequency diversity to transmit the same QPSK symbol to enhance the SER. Notice that the user rate is given in nats per channel use, which means that each time the user accesses the channel is capable to place that amount of information. It is well-known that with the use of OFDM the duration of the symbol is substantially increased, in addition to the cyclic prefix; this means that the rate in nats per second is intrinsically linked to the rate in which the user access the channel. This justifies again the trade-off between throughput and diversity.

A Monte-Carlo simulation has been done in a THP system with $K = 2$ active users that share 4 carriers, in a 2×1 MISO BC channel. The channel conditions for the users are $\xi_{n,2} = 2\xi_{n,1}$ at each carrier, with $\xi_{1,k} > \ldots > \xi_{N,k}$ for each user. The achieved rates for both users are depicted in figure 17.14, while the SER and NMSE are shown in figure 17.15. The fist apparent change when using frequency diversity with $N > 1$ is that, with independence of frequency channels —this is reflected in full rank channel matrix—, the user rate is, approximately, multiplied by N. This means that the user is able to receive/place N OFDM symbols per channel use. Again, it is important to highlight that OFDM penalizes the duration of the symbol, which increases with a factor of N plus the cycle prefix. In a rich environment with a lot of multi-paths, the delay spread of the channel is high, i.e. $D_s = 5T$, being T the duration of the symbol. The OFDM modulation achieves orthogonality of channels if and only if the cycle prefix is larger than the channel memory: the delay spread. This means that, whatever is the number of points in OFDM, there is a loss in throughput. As stated above, it is up to the system to determine the trade-off between the throughput and the diversity. However, whatever is the strategy followed by the system, the rate capacity is a measure of the capacity of the channel, regardless the use we make of it. As a result, comparing figure 17.14 to figure 17.7, both figures showing the user rate for a $T = K = 2$ user system, it is easy to appreciate that, at any SNR point, the rate achieved using N carriers is the sum of the rate achieved by the N orthogonal channels. The behavior of the rates versus average SNR is the same: at high SNR the rates tend to approximate.

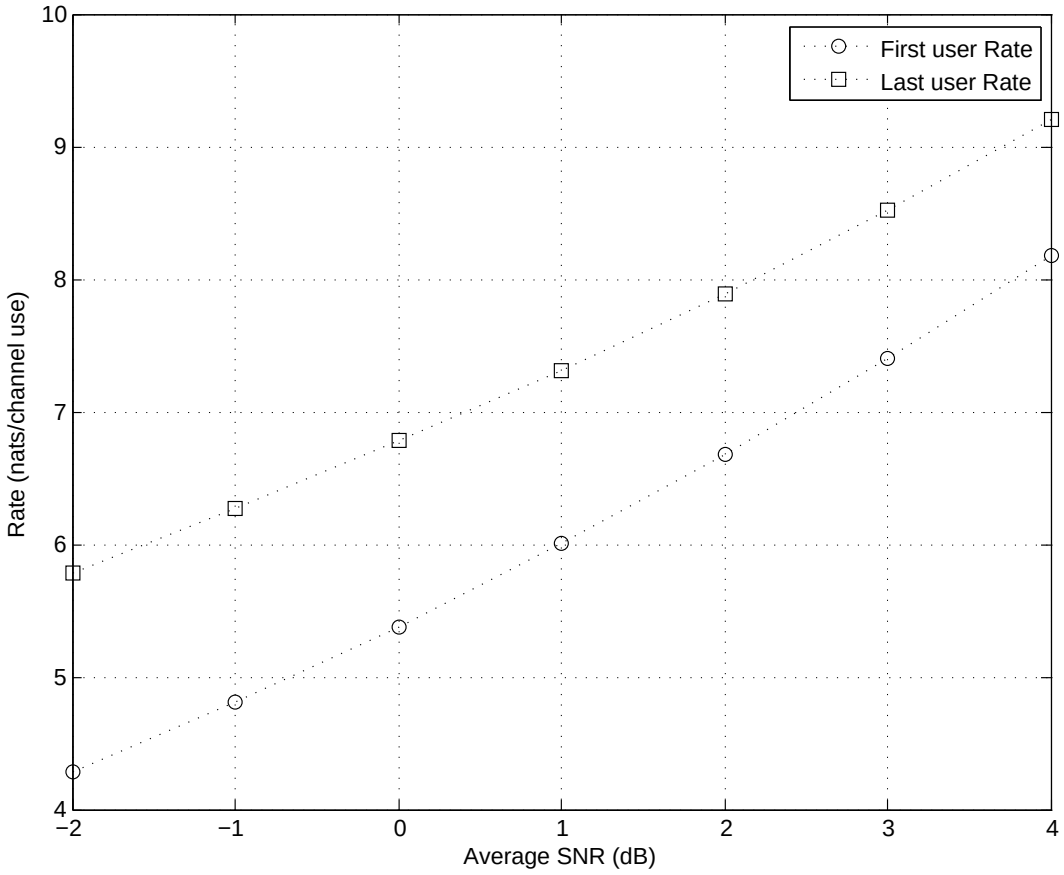
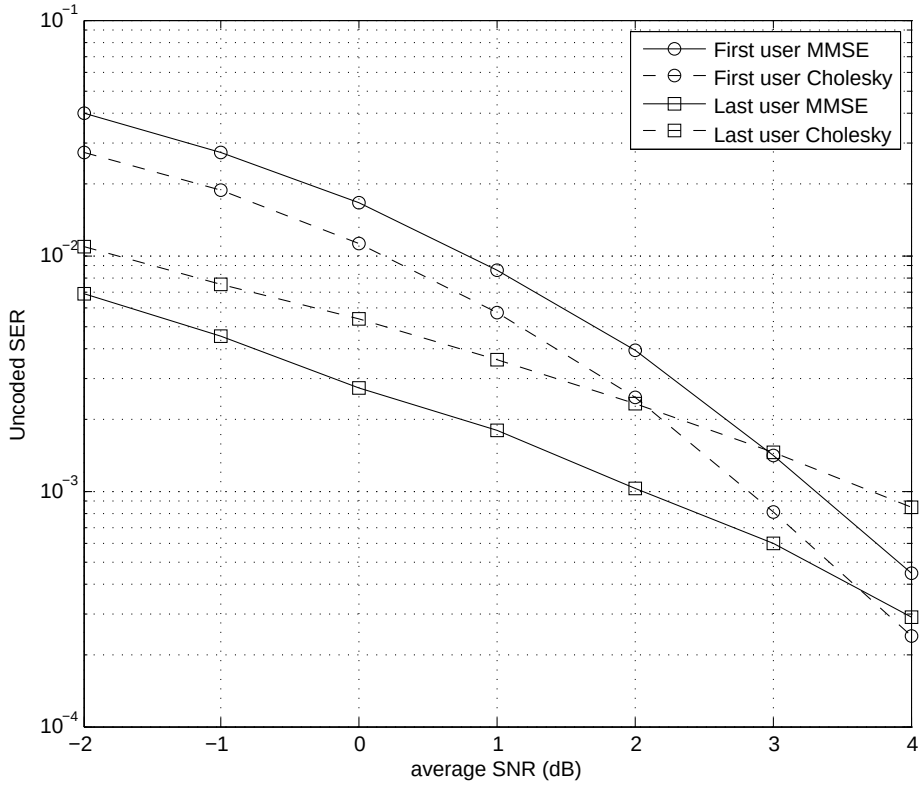
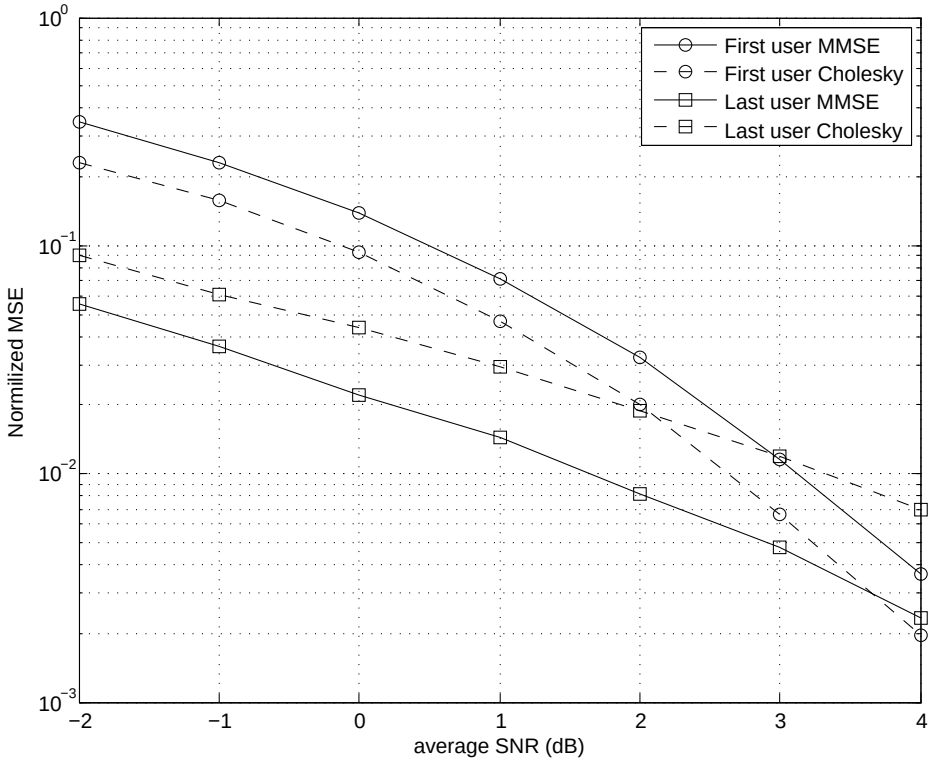


Figure 17.14: User rates for the $K = 2$ users $N = 4$ carriers THP 2×1 MISO BC versus average SNR, for each user, with diverse channel conditions $2\xi_{n,1} = \xi_{n,2}$, in nats per channel use.



(a)



(b)

Figure 17.15: SER (a) and NMSE (b) for the $K = 2$ users $N = 4$ carriers THP 2×1 MISO BC versus average SNR, for each user, with diverse channel conditions $2\xi_{n,1} = \xi_{n,2}$.

Finally, seeing the achieved SER and NMSE in figure 17.15, we appreciate that the use of frequency diversity enhances exponentially the performance. Since it is a scenario with diverse channel conditions, the first user achieves better performance at high SNR. Nonetheless, the performance is notably improved equally for both users, since the probability of detecting an error is reduced keeping in mind the fact that the probability of having bad transmission at more than one carrier at the same time is very low.

17.5 Number of Active Users in the System

This last section analyzes the effect of sharing the resources with other users in a system equipped with $T = 3$ transmitting antennas at one single-carrier. Since the number of active users in this system is limited by $K \leq T$, in this case the system can cope with 1, 2 or 3 users. Figure 17.16 shows the SER and the achieved rate for a selected user in three scenarios.

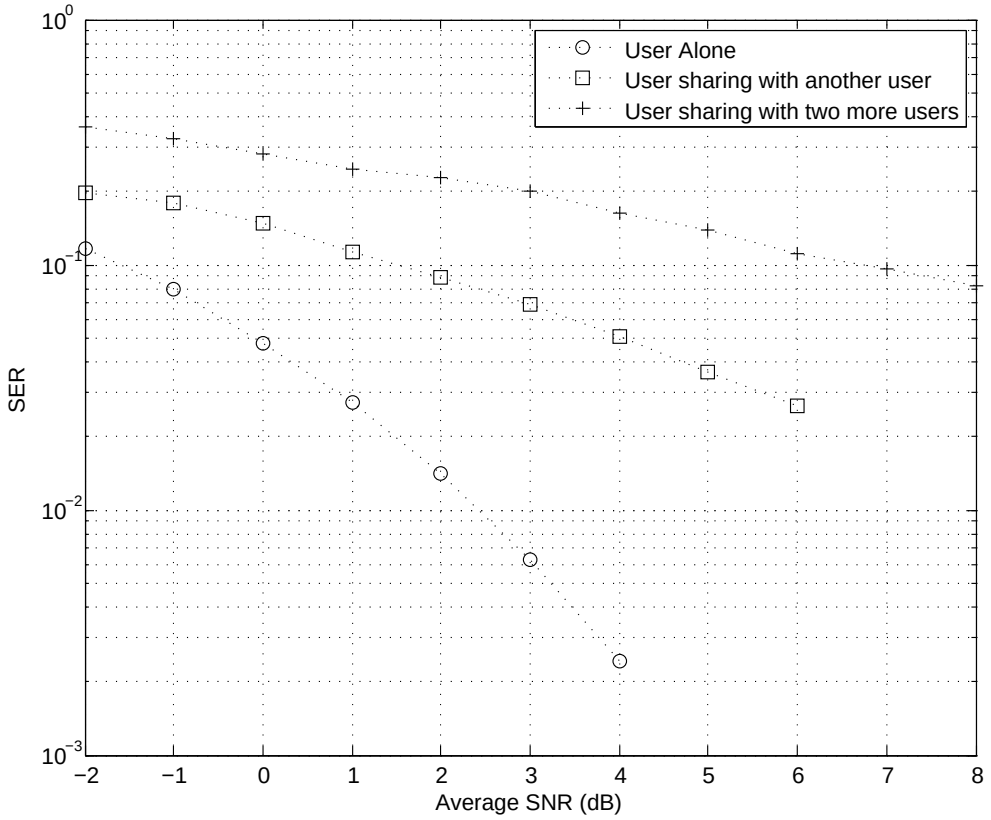
The first one (circles) depict the performance of a user who is alone at the 3-transmitting antennas. As shown in the point-to-point single-user MIMO channel, the user of more transmitting antennas for the same user results in an increasing of performance as spatial diversity is obtained. In other words, the THP does nothing as there is no interference to remove, so the power allocation filter copies the same information to the 3 transmitting antennas and these copies are optimally received by the same user which, since filter $\boldsymbol{F}$ is designed to optimally face the channel, receives maximum combination of them.

The second scenario (squares) simulates the entry of a second user, with worst channel conditions that the considered user. This means that the user whose performance is depicted in the figure becomes the last precoded. As we know, hence, it will achieve better SER and rate compared to the new user at high SNR but, since now both user share the same resources, there exists a penalty in performance. Namely, the SER increases and the rate decreases with respect to the first scenario because the available transmit energy must be shared for both users.

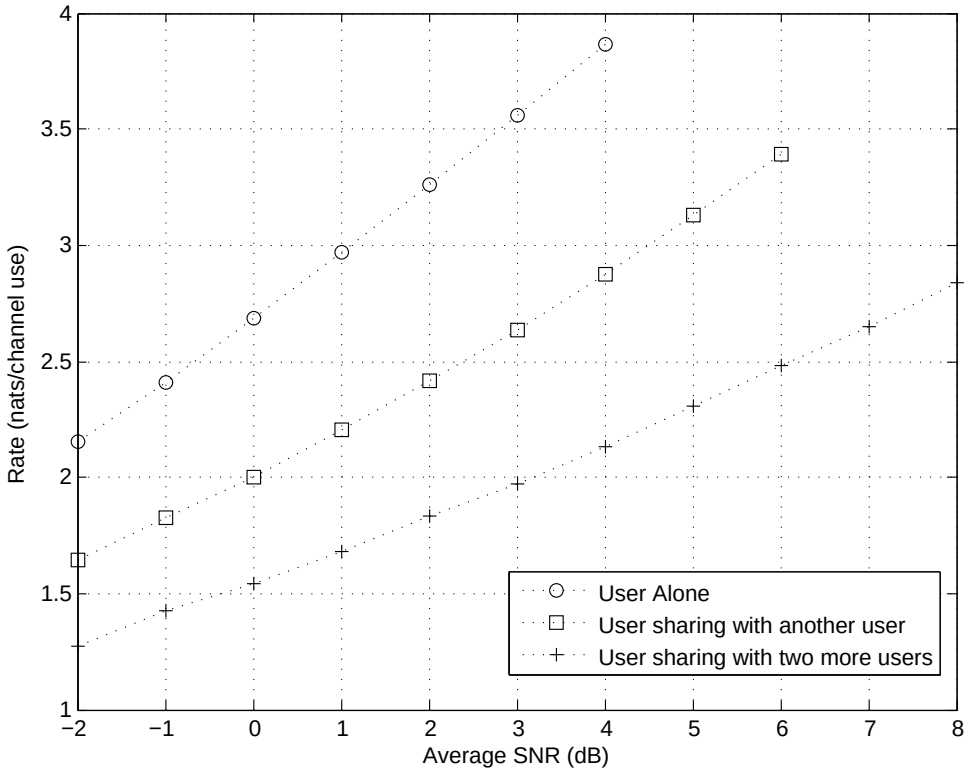
Finally, the third scenario (crosses) corresponds to the case in which $K = 3$ users co-exist in the system. Now, this third user has better channel conditions and, as a result, though it is not appreciated in the figures, it is placed in the last position in the precoding order. As expected, the performance of the user of interest down performs at the same level of average SNR.

17.6 Other System Parameters

As a reminder, the single-user SISO and MIMO channels have provided with very understanding results that can easily be applied to the multi-user scenario. Examples of this are the number of available OFDM carriers, N, which produces a better diversity in the coding scheme; the correlation parameter at the transmitter, ρ_T, which defines the level of uncorrelation in the different inner paths at the same carrier; or the noise distribution that, since it only affects the determination of $\boldsymbol{F}$ it is analogous to consider different channel conditions in terms of channel gain or attenuation. All these system parameters are expected to evolve under the same nature in the multi-user channel.



(a)



(b)

Figure 17.16: SER (a) and rates (b) for the $K = \{1, 2, 3\}$ users $N = 1$ carrier THP 3×1 MISO BC versus average SNR, for a selected user, with diverse channel conditions.

Part V

Conclusions and Future Work

Chapter 18

Conclusions

In this thesis, Tomlinson-Harashima precoding (THP) [Tom71, HM72] has been considered under analysis and design for three communication systems with the scope of providing an ambitious technique that satisfies the needs of the current telecommunication market. Indeed, THP is a greatly versatile solution that makes an excellent use of the system resources to achieve very high capacity, certainly the essential point in covering the present demand of high quality of service (QoS) in wireless multi-user services.

Though Tomlinson-Harashima based precoders already exist in wire line communication —to be cited the American high bit rate digital subscriber line (HDSL) or the 10 Gbit 10GBASE-T Ethernet IEEE 802.3an standard—, its application to wireless communication is motivated by the fact that it is regarded as a low-complexity practical implementation of dirty paper coding (DPC) [Cos83], able to spatially multiplex several users. Thus, it puts on the table the space as a promising dimension which, combined to other techniques such as orthogonal frequency division multiplexing (OFDM) or code domain multiplexing (CDM), might increase the capacity of the system. For instance, THP could be used in cellular systems to place K users at the same cell, sharing the same frequency and the same scrambling code, so that the blocking probability or dropping probability by handover would perceive an extreme reduction. Likewise, the use of two transmitting antennas in a IEEE 802.11n wireless network would enhance the downlink and uplink data rates by performing the precoding at the access point, letting the mobile devices —e.g. a laptop— as simple as possible significantly improving the cost and the battery consumption.

Former examples are some of the possible scenarios where THP would walkover. The following paragraphs state the conclusions obtained from the work done in parts II, III and IV, where different channel conditions have been simulation regarding the model provided by the ETSI Hiperlan II standard [MS98]. The main scope has been to predict the behavior of THP in diverse situations which, in a full layered system, would abruptly improve the performance of the upper layers.

The results carried out in part II place the THP in a very advantageous position compared to other equalization techniques, albeit the precoder and the receiver introduce losses due to the modulo operation. First of all, since THP does not suffer from error propagation as decision feedback equalization (DFE), the use of a convolution or constellation-expanding codes —usually coupled with OFDM modulation— is very welcome. Secondly, the introduction of a dither signal added to the user data, together with the effect of the non-linear modulo operation, is a low-complexity solution to the problem of instantaneous power increase required when the transmitter must face a deep channel fade. Interestingly, the transmitter power is bounded by the modulo operation but also controlled to a fixed

value by the dithering, avoiding that the linear amplifier works out of the linear regime. In third place, it is noticed that the filter optimization makes a clever use of the temporal diversity provided by the frequency-selective channel: in a rich environment, the same transmitted symbol is received a number of times with different signature and, in view of it, the precoding technique rearranges the information to obtain a constructive addition of the copies after transmission. This can be done because despite being regarded as inter-symbol interference (ISI), the knowledge on the channel makes possible to treat it as temporal diversity. Lastly, probably the most important conclusion is that many ideas and consequences obtained in this part become a conceptual basis for any THP system —and thus the two other channels analyzed in this thesis— by reason of having employed vectorial and matrix notation. Among them, the optimization of the minimum mean-square error (MMSE) and zero-forcing (ZF) feedback and receive filters, the role and influences of the non-liner modulo operation, the proposal and use of a dither signal, the accurate mathematical development of the Cholesky filters, and the linearization of the model are of a big value.

The trio multiple-input multiple-output (MIMO), OFDM, and THP is absolutely a choosy combination because it provides an scalable, robust, high capacity and reliable wireless solution. As stated throughout the thesis, since the MIMO channel models many communication problems —besides multiple antennas —, the work done in part III is very general and comprehensively compatible. It is seen that the MMSE solution in MIMO-OFDM systems outperforms compared to the ZF counterpart due to a key point in MIMO systems: the power allocation. On the one hand, MMSE cleverly allocates the available power according to, approximately, the channel conditions in proportion. In other words, the MMSE solution may omit the noisiest channels to better exploit the better ones. On the other hand, ZF waists resources to allocate energy to the noisiest channels with the aim of fulfilling the ZF condition. The introduction of MIMO and OFDM offers an evident flexibility to the system in managing the resources, but with subtle consequences. First, OFDM achieves up to N times more resources, at the penalty of multiplying by roughly N the symbol period to assure robustness in front of temporal channel conditions. And second, though MIMO provides more capacity [GJJV03], it requires a more complex treatment on the transmitted signal, the placement of more antennas and, with a fixed power budget, the distribution of it along the inputs. Despite these issues, MIMO and OFDM are and will be worthy to be considered in many systems, as their resources may be adapted to achieve high channel diversity.

Finally, the broadcast channel (BC) part allows to deduce that an arbitrary number of users, K, can be spatially multiplexed in a system with $T \geq K$ transmitting antennas by making use of the resources provided by MIMO and OFDM, as well as THP implementing DPC. The main conclusion is that, contrary to single-user, the feedback filter and modulo operation are exclusively in charge of precoding the users. In other words, with uncooperative receivers, maximum combination at reception is not possible. As a result, both MMSE and ZF are the same, since the precoding filter concentrates only in canceling the multi-user interference (MUI). Another interesting point to highlight is that asymmetric signal processing is a powerful technique, as far as channel state information (CSI) is available, because it light weights the complexity, cost and power consumption of the receivers. The issue of user ordering has been considered of a big interest, since it has severe consequences on the fairness among users and on the power allocation algorithm. Thanks to information theory considerations [Cov72], it is seen that only the not yet precoded users interfere to the already precoded users. Therefore, more power is allocated to the first user —since it is the one that receives more interferences and the one that does not interfere—, while the last precoded user receives less power —as it interferes to the others despite being interference free; defining the best-last user ordering strategy as the fairest choice.

Chapter 19

Future Work

Lastly, this chapter proposes five lines of further investigation, some of them strictly derived from the present thesis, whereas others —despite being parallel to THP— promise boundless results when crossed with the Tomlinson-Harashima structure. Cited in the following list, all of them turn out to be of a big worthiness in communication systems.

1. Regarding the importance of CSI, THP should be considered together with a channel estimation subsystem capable of both predicting the value of the channel by transmitting known sequences, or pilots and optimally providing it to the THP filters. The estimation of the channel should take into consideration the characteristics of the channel, namely if it is a flat or frequency selective; as well as the degree of channel temporal variation. This investigation line would close the study of THP with complete CSI and, thanks to it, it would be possible to determine the robustness of the optimal filters in cases when the estimation is not accurate.

2. A second line of investigation, already started by Utschick in [DBU07] for the multi-user single-carrier flat MIMO channel, would be the design of THP —with the combination MIMO and OFDM— for the BC with partial or statistical CSI. Partial CSI deals with the knowledge —in other words, estimation— of the statistics of the channel conditioned to a certain received signal, whereas statistical CSI consists of knowing the full channel probability density function (PDF). It has been shown that these techniques outperform when the channel variations are high compared to the block transmission duration. This is the particular case in many wireless systems whose channel coherence time is small compared to the filtering order.

3. Along the thesis, it has been assumed that every user knows how and from where it will receive the information. Namely, in the single-user MIMO-OFDM channel, this deals with negotiating how many carriers and how many symbols per carrier will be transmitted. In the multi-user BC, this issue gains importance when coupled with orthogonal frequency division multiple access (OFDMA), since each user must known which set of carriers is the transmitter using to communicate. This line of research should suggest the optimal way in which this information is send between the transmitter and the receiver(s) in a THP system. Evidently, a possible solution is to dedicate some of the resources to the transmission of control information such as reserved carriers; or, alternatively, mix it with the user data using, e.g., a spread spectrum technique. In any case, the optimal solution requires the transmission of the minimum amount of information, in view that the control signals suppose a large payload in many systems.

4. For the multi-user part, checking the performance of Tomlinson-Harashima based precoding schemes with other power allocation policies could be researched. The water-filling algorithm presented in this thesis is optimal for the sum-rate capacity, which means that the sum of the rates of all the users is maximized. Other approaches, stated in subsection 13.3.2, should be in mind such as equal rates or balanced rates. Usually, these approaches consist in determining the priority coefficients $\{\alpha_k\}_k$ prior to the rate optimization process. For the balanced rates, Sartenaer [SVL05] provides the full water-filling algorithm for multi-carrier SISO channel. This could be investigated together with THP and also developed for multiple-input single-output (MISO) or MIMO BCs, taking into account that depending on the type of service offered by the system a different QoS must be attained.

5. Finally, robust codes such as convolution codes and the use of interleavers could be investigated in the single-user cases coupled with the Tomlinson-Harashima structure. The placement of the feedback loop at the receiver was motivated by the fact that decision error propagation is not possible. However, the modulo operation at the receiver may generate modulo error and, in convolution coding, this can be translated to inaccurate values of the log likelihood ratio (LLR) provided by the soft demapper. In this sense, besides investigating how codes improve THP, it should be interesting to check how THP affects the performance of the code, as well as possibly design codes with more sophisticated structures taking into consideration the precoder structure.

Bibliography

[ADC95] N. Al-Dahir and J. M. Cioffi, *MMSE decision-feedback equalizers: finite-length results*, IEEE Transactions on Information Theory **41** (1995), no. 4, 961–975.

[BF07] H. K. Bizaki and A. Falahati, *Power loading by minimizing the average symbol error rate on MIMO THP systems*, The 9th International Conference on Advanced Communication Technology, vol. 2, February 2007, pp. 1323–1326.

[BF08] ———, *Tomlinson-Harashima precoding with imperfect channel state information*, The Institution of Engineering and Technology **2** (2008), no. 1, 151–158.

[BS00] S. Barbarossa and A. Scaglione, *Signal processing advances in wireless and mobile communications*, Trends in Single and Multi-User Systems, vol. 2, ch. Time-Varying Fading Channels, Prentice-Hall, Inc., New Jersey, 2000.

[BV04] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[CC04] I. B. Collings and I. V. L. Clarkson, *A low-complexity lattice-based low-PAR transmission scheme for DSL channels*, IEEE Transactions on Communications **52** (2004), no. 5, 755–764.

[Cio02] J. M. Cioffi, *Digital communications: Signal processing*, EE379 lecture notes, Stanford University, 2002.

[Cos83] M. H. Costa, *Writing on dirty paper*, IEEE Transactions on Information Theory **29** (1983), 439–441.

[Cov72] T. M. Cover, *Broadcast channels*, IEEE Transactions on Information Theory **18** (1972), no. 1, 2–14.

[CS03] G. Caire and S. Shamai, *On the achievable throughput of a multi-antenna Gaussian broadcast channel*, IEEE Transactions on Information Theory **49** (2003), no. 7, 1691–1706.

[CTHC08] C-B. Chae, T. Tang, R. W. Heath, and S. Cho, *MIMO relaying with linear processing for multi-user transmission in fixed relay networks*, IEEE Transactions in Signal Processing **56** (2008), no. 2, 727–738.

[DBU07] F. A. Dietrich, P. Breun, and W. Utschick, *Robust Tomlinson-Harashima precoding for the wireless broadcast channel*, IEEE Transactions on Signal Processing **55** (2007), no. 2, 631–644.

[ESZ05] U. Erez, S. Shamai, and R. Zamir, *Capacity and lattice strategies for canceling known interference*, IEEE Transactions on Information Theory **51** (2005), no. 11, 3820–3833.

[FE91] G. D. Forney and M. V. Eyuboğlu, *Combined equalization and coding using precoding*, IEEE Communications Magazine **29** (1991), no. 12, 25–34.

[FG98] G. Foschini and M. Gans, *On limits of wireless communication in a fading environment when using multiple antennas*, Wireless Personal Communications **6** (1998), 311–335.

[FGVW99] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, *Simplified processing for high spectral efficiency wireless communication employing multi-element arrays*, IEEE Journal on Selected Areas in Communications **17** (1999), no. 11, 1841–1852.

[FK00] S. D. G. J. Foschini and J. M. Kahn, *Fading correlation and its effect on the capacity of multi element antenna systems*, IEEE Transactions on Communications **48** (2000), 502–513.

[Fle87] R. Fletcher, *Practical methods of optimization*, John Wiley and Sons, New York, 1987.

[FTK07a] Y. Fu, C. Tellambura, and W. A. Krzmien, *Precoding for multi-user orthogonal space-time block-coded OFDM downlink over spatially-correlated channels*, Wireless Communications and Networking Conference, March 2007, pp. 1308–1313.

[FTK07b] ———, *Transmitter precoding for ICI reduction in closed-loop MIMO OFDM systems*, IEEE Transactions on Vehicular Technology **56** (2007), no. 1, 115–125.

[FWLH02] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, *Space-time transmission using Tomlinson-Harashima precoding*, January 2002, To be presented at 4th International ITG Conference on Source and Channel Coding, Berlin.

[FYL07] C-H. F. Fung, W. Yu, and T. J. Lim, *Precoding for the multi-antenna downlink: multi-user SNR gap and optimal user ordering*, IEEE Transactions in Communications **55** (2007), no. 1, 188–197.

[GA06] Ana García-Armada, *SNR gap approximation for M-PSK-based bit loading*, IEEE Transactions on Wireless Communications **5** (2006), no. 1, 57–60.

[Gal68] R. G. Gallager, *Information theory and reliable communication*, John Wiley, New York, 1968.

[GC00] G. Ginis and J. M. Cioffi, *A multi-user precoding scheme achieving cross-talk cancellation with application to DSL systems*, Proc. of the Asilomar Conference on Signals, Systems and Computers, vol. 2, October 2000, pp. 1627–1631.

[GC02] ———, *Vectored transmission for digital subscriber line systems*, IEEE Journal on Selected Areas in Communications **20** (2002), no. 5, 1085–1104.

[GJJV03] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, *Capacity limits for MIMO channels*, IEEE Journal on Selected Areas in Communications **21** (2003), no. 5, 684–702.

[GS99] M. R. Gibbard and A. B. Sesay, *Asymmetric signal processing for indoor wireless LAN's*, IEEE Transactions on Vehicular Technology **48** (1999), no. 6, 2053–2064.

[GSV05] D. Guo, S. Shamai, and S. Verdu, *Mutual information and minimum mean-square error in gaussian channels*, IEEE Transactions on Information Theory **51** (April 2005), no. 4, 1261–1282.

[HJU05] R. Hunger, M. Joham, and W. Utschick, *Extension of linear and non-linear transmit filters for decentralized receivers*, Proc. European Wireless Conf., vol. 1, April 2005, pp. 40–46.

[HM72] H. Harashima and H. Miyakawa, *Matched-transmission technique for channels with intersymbol interference*, IEEE Transactions on Communications **20** (1972), 774–780.

[JBU04] M. Joham, J. Brehmer, and W. Utschick, *MMSE approaches to multi-user spatio-temporal Tomlinson-Harashima precoding*, Proc. of the ITG Conference on Source Channel Coding, January 2004, pp. 387–394.

[JRV+05] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, *Sum power iterative water-filling for multi-antenna Gaussian broadcast channels*, IEEE Transactions on Information Theory **51** (2005), no. 4, 1570–1580.

[JU05] M. Joham and W. Utschick, *Ordered spatial Tomlinson-Harashima precoding*, Smart Antennas—State-of-the-Art, EURASIP Book Series on Signal Processing and Communication, vol. 3, EURASIP, Hindawi Publishing Corporations, New York, 2005.

[JUN05] M. Joham, W. Utschick, and J. A. Nossek, *Linear transmit processing in MIMO communications systems*, IEEE Transactions on Signal Processing **53** (2005), no. 8, 2700–2712.

[KJUB05] K. Kusume, M. Joham, W. Utschick, and G. Bauch, *Efficient Tomlinson-Harashima precoding for spatial multiplexing on flat MIMO channel*, Proc. of the International Conference on Communications, vol. 3, May 2005, pp. 2021–2025.

[KJUB07] ———, *Cholesky factorization with symmetric permutation applied to detecting and precoding spatially multiplexed data streams*, IEEE Transactions on Signal Processing **55** (2007), no. 6, 3089–3103.

[KLV07] A. Kalakech, J. Louveaux, and L. Vandendorpe, *Applying the balanced capacity concept to DSL systems*, to appear at Proc. Int. Conf. Communications (ICC'07) (2007).

[KWZ02] S. Kung, Y. Wu, and X. Zhang, *Bezóut space-time precoders and equalizers for MIMO channels*, IEEE Transactions on Signal Processing **50** (2002), no. 10, 2499–2514.

[Lia05] A. P. Liavas, *Tomlinson-Harashima precoding with partial channel knowledge*, IEEE Transactions on Communications **53** (2005), no. 1, 5–9.

[LK05] J. Liu and W. A. Krzymien, *Improved Tomlinson-Harashima precoding for the downlink of multiple antenna multi-user MIMO systems*, IEEE Wireless Communications and Networking Conference **1** (13-17 March 2005), 466–472 Vol. 1.

[LLLS07] P.-H. Lin, S.-C. Lin, H.-T. Liu, and H.-J. Su, *Peak to average power ratio reduction for multi-carrier systems using dirty paper coding*, Journal of Communications **2** (2007), no. 3, 9–16.

[LTV06] A. Lozano, A. M. Tulino, and S. Verdú, *Optimum power allocation for parallel Gaussian channels with arbitrary signal constellations*, IEEE Transactions on Information Theory **52** (2006), no. 7, 3033–3051.

[Mil07] Elisabeth Million, *The Hadamard product*, University of Puget Sound lecture notes, April 2007.

[MN99] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*, revised ed., John Wiley and Sons, New York, 1999.

[MO79] A. W. Marshall and I. Olkin, *Inequalities: theory of majorization and its applications*, Academic, New York, 1979.

[MS76] J. E. Mazo and J. Salz, *On the transmitted power in generalized partial response*, IEEE Transactions on Communications **24** (1976), no. 3, 348–352.

[MS98] J. Mebdo and P. Schramm, *Channel models for HIPERLAN/2 in different indoors scenarios*, ETSI, March 1998.

[MS07] M. Morelli and L. Sanguinetti, *A unified framework for Tomlinson-Harashima precoding in MC-CDMA and OFDMA downlink transmissions*, IEEE Transactions on Communications **55** (2007), no. 10, 1963–1972.

[MW05] T. Michel and G. Wunder, *Optimal and low complex suboptimal transmission schemes for MIMO-OFDM broadcast channels*, IEEE International Conference on Communications **1** (2005), no. 16-20, 438–442.

[OS75] A. V. Oppenheim and R. W. Schafer, *Digital signal processing*, Prentice-Hall, Englewood Cliff, New Jersey, 1975.

[PCL03] D. Pérez Palomar, J. M. Cioffi, and M. A. Lagunas, *Joint tx-rx beamforming design for multi-carrier MIMO channels: a unified framework for convex optimization*, IEEE Transactions on Signal Processing **51** (2003), no. 9, 2381–2401.

[Pro00] J. G. Proakis, *Digital communications*, 4th ed., Science/Engineering/Math, McGraw-Hill, August 2000.

[RC98] G. G. Raleigh and J. M. Cioffi, *Spatio-temporal coding for wireless communication*, IEEE Transactions on Communications **46** (1998), 357–366.

[RG07] A. Artés Rodríguez and F. Pérez González, *Comunicaciones digitales*, ch. Detección en canales con interferencia intersimbólica, Pearson Education, 2007.

[Sch64] L. Schuchman, *Dither signals and their effect on quantization noise*, IEEE Transactions on Communications **12** (1964), no. 4, 162–165.

[SD08] M. B. Shenouda and T. N. Davidson, *A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding*, IEEE Journal on Selected Areas in Communications **26** (2008), no. 2, 401–411.

[Sha48] C. E. Shannon, *A mathematical theory of communications*, Bell Systems Lab **27** (1948), 379–423.

[SL96] S. Shamai and R. Laroia, *The inter-symbol interference channel: lower bounds on capacity and channel precoding loss*, IEEE Transactions on Information Theory **42** (1996), 1388–1404.

[SM00] L. L. Scharf and C. T. Mullis, *Canonical coordinates and the geometry of interference, rate and capacity*, IEEE Transactions on Signal Processing **48** (2000), no. 3, 824–831.

[SM07] L. Sanguinetti and M. Morelli, *Non-linear precoding for multiple-antenna multi-user downlink transmissions with different QoS requirements*, IEEE Transactions on Wireless Communications **6** (2007), no. 3, 852–856.

[SS77] A. Sripad and D. Snyder, *A necessary and sufficient condition for quantization errors to be uniform and white*, IEEE Transactions on Acoustics, Speech, and Signal Processing **25** (1977), no. 5, 442–448.

[SS04] O. Simeone and U. Spagnolini, *Linear and non-linear preequalization/equalization for MIMO systems with long term channel state information at the transmitter*, IEEE Transactions on Wireless Communications **3** (2004), no. 2, 373–378.

[SSB+02] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, *Optimal designs for space-time linear precoders and decoders*, IEEE Transactions on Signal Processing **50** (2002), no. 5, 1051–1064.

[Stü01] G. L. Stüber, *Principles of mobile communications*, 2nd ed., Kluwer Academic Publishers, Boston, 2001.

[SVL05] T. Sartenaer, L. Vandendorpe, and J. Louveaux, *Balanced capacity of wire line multi-user channels*, IEEE Transactions on Communications **53** (2005), no. 12, 2029–2042.

[TdPE02] S. Thoen, L. Van der Perre, and M. Engels, *Adaptive loading for OFDM/SDMA-based wireless networks*, IEEE Transactions on Communications **50** (2002), no. 11, 1789–1810.

[Tel99] E. Telatar, *Capacity of multi-antenna Gaussian channels*, Europ. Transactions on Telecommunications **10** (1999), no. 6, 585–569.

[Tom71] M. Tomlinson, *New automatic equalizer employing modulo arithmetic*, Electronic Letter **7** (1971), no. 5/6, 138–139.

[TUBN05] P. Tejera, W. Utschick, G. Bauch, and J. A. Nossek, *Sum-rate maximizing decomposition approaches for multi-user MIMO-OFDM*, IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, 2005, pp. 213–235.

[UL07] P. S. Udupa and J. S. Lehnert, *Optimizing zero-forcing precoders for MIMO broadcast systems*, IEEE Transactions on Communications **55** (2007), no. 8, 1516–1624.

[VJG03] S. Vishwanath, N. Jindal, and A. Goldsmith, *Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels*, IEEE Transactions of Information Theory **49** (2003), no. 10, 2658–2668.

[Wei94] L-F. Wei, *Generalized square and hexagonal constellations for inter-symbol interference channels with generalized Tomlinson-Harashima precoders*, IEEE Transactions on Communications **42** (1994), no. 9, 2713–2721.

[Wey49] H. Weyl, *Inequalities between the two kinds of eigenvalues of a liner transformation*, Proc. of the National Academy of Science **35** (1949), 408–411.

[WFGV98] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, *V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel*, Proc. of the International Symposium on Signals, Systems, and Electronics (1998), 295–300.

[WFH04] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, *Lattice-reduction-aided broadcast precoding*, IEEE Transactions on Communications **52** (2004), no. 12, 2057–2060.

[WFVH04] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, *Precoding in multi-antenna and multi-user communications*, IEEE Transactions on Wireless Communications **3** (2004), no. 4, 1305–1316.

[Wid61] B. Widrow, *Statistical analysis of amplitude-quantized sample-data systems*, Transactions on AIEE (Applications an Industry) **79** (1961), 555–568.

[WK04] Y. Wu and S. Kung, *Signal detection for MIMO-ISI channels: a unitary linear recovery approach*, IEEE Transactions on Signal Processing **52** (2004), no. 3, 703–720.

[YC04] W. Yu and J. M. Cioffi, *Sum capacity of Gaussian vector broadcast channels*, IEEE Transactions of Information Theory **50** (2004), no. 9, 1875–1892.

[YRBC04] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, *Iterative water-filling for Gaussian multiple-access channels*, IEEE Transactions of Information Theory **50** (2004), no. 1, 145–152.

[ZK05] X. Zhang and S. Kung, *Capacity bound analysis for FIR Bézout equalizers in ISI MIMO channels*, IEEE Transactions on Signal Processing **53** (2005), no. 6, 2193–2204.

[ZL07] Y. Zhu and K. B. Letaief, *Frequency domain equalization with Tomlinson-Harashima precoding for broadband MIMO systems*, IEEE Transactions on Wireless Communications **6** (2007), no. 12, 4420–4431.