Pre-Filtering in MC-CDMA Downlink Transmissions

Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica e Telecomunicazioni

Università di Pisa

A thesis submitted for the degree of

Dottore di Ricerca in Ingegneria dell'Informazione

Gennaio 2003 - Dicembre 2005

Author:

Ing. Luca Sanguinetti

Advisors:

Prof. Umberto Mengali

Prof. Michele Morelli

Acknowledgements

Foremost, I would like to thank my advisors Professor Umberto Mengali and Professor Michele Morelli for their guidance and support over the past three years. Professor Umberto Mengali is the one who convinced me to begin the intriguing and appealing journey of my Ph.D. He has always shown confidence in me and has given me a continuous and valuable support. Professor Michele Morelli has been my mentor and guide with whom I will be indebted for ever. His office door has been always open and I could talk to him no matter how busy he was. He encouraged me to look at problems from different view points and to leave conventional ways of thinking. Fortunately, I have beneficiated from his extraordinary motivation, great intuitions and technical insight. I just hope that my thinking and working attitudes have been shaped according to such outstanding qualities.

Many thanks to my friends and colleagues of the Department of Information Engineering at the University of Pisa for their countless discussions and for a pleasant working atmosphere. I am especially grateful to all members of the Transmission Systems Research Group. Special thanks are due to Professor Aldo D'Andrea who has always provided me with all means to carry out my research activities. I will always be indebted to my colleagues and friends of the lab for their helpful technical insights and valuable hints over all these three years. Among them, Simone has always been close to me not only

as a brother but also as a source of inspiring discussions and Antonio D'Amico and Marco Moretti have been always ready to listen to me and to have some fun together.

I would like to express my sincere thanks to all the members of the Mobile Radio Transmission Group from German Aerospace Center (DLR) in Oberpfaffenhofen. All of them welcomed me into the research group and treated me like a member of the group with no distinction, making my stay at DLR a pleasant and rewarding experience. Special thanks are due to my friend and colleague Ivan Cosovic with whom I had the fortune to work and collaborate. I believe that he has given me more than what he has received and I will always be grateful to him for the nice times we spent together.

Last but not least, I would like to thank my parents for the sacrifices they have made and for the inspiration and support they have provided throughout my life. Had they not instilled in me the value of higher education, I would not be writing this thesis today. Finally, I would like to thank Giulia, without whose patience, understanding, support, and countless sacrifices, I would not have been able to venture this far.

Pisa, March 2006 Luca Sanguinetti

Abstract

Future wireless communication systems are expected to support high-speed and high-quality multimedia services. In these applications the received signal is typically affected by frequency-selective fading, which must be properly counteracted to avoid a severe degradation of the system performance.

MC-CDMA is a multiplexing technique that combines OFDM with direct sequence CDMA. It is robust to frequency-selective fading thanks to the underlying OFDM modulation and exploits frequency diversity by spreading the data of different users in the frequency domain. For these reasons it is considered as a promising candidate for the physical layer of future high-speed wireless communications.

Recent publications show that MC-CDMA is particularly suitable for downlink transmissions, i.e., from the base station to the mobile terminals. In these applications orthogonal spreading codes are usually employed to provide protection against co-channel interference. In the presence of multipath propagation, however, signals undergo frequency-selective fading and the code orthogonality is lost. This gives rise to multiple-access interference, which strongly limits the system performance. In the past few years several advanced multiuser detection techniques have been proposed and discussed for interference mitigation. However, in spite of their effectiveness, all these

methods are quite unattractive for downlink applications since they would entail high complexity and excessive power consumption at the remote units.

As an alternative to multi-user detection, pre-filtering techniques can be employed in downlink transmissions to mitigate multiple-access interference and channel distortions. The idea behind pre-filtering is to vary the gain assigned to each subcarrier so that interference is reduced and the signal at the receiver appears undistorted. In this way, simple and low complex single-user detectors can be employed at the remote units, thereby moving most of the computational burden to the base station, where power consumption and computational resources are not critical issues.

In general terms, the main contribution of this dissertation is threefold. First, we propose and discuss several *linear* and *non-linear* pre-filtering schemes for the downlink of MC-CDMA systems equipped with multiple transmit antennas and operating in a time-division-duplex mode. The resulting schemes are derived according to different optimization criteria and aim at combating the detrimental effects of MAI while maintaining the complexity of the remote units as low as possible. A second contribution comes from providing a unified framework for investigating pre-filtering in the downlink of both MC-CDMA and OFDMA systems. The use of a unified framework comprising both MC-CDMA and OFDMA allows a fair comparison between these multiple-access technologies under the same operating conditions. It turns out that OFDMA outperforms MC-CDMA when the system resources are optimally assigned to the active users according to the actual channel realization. As we shall see, in order

to work properly, all the proposed schemes require explicit knowledge of the channel responses of the active users. In time-division-duplex systems this information can be achieved by exploiting the channel reciprocity between alternative uplink and downlink transmissions. If channel variations are sufficiently slow, the channel estimates of the active users can be derived at the base station during an uplink time-slot and reused for pre-filtering in the subsequent downlink time-slot. Thus, a third contribution comes from addressing the problem of channel acquisition in the uplink of an MC-CDMA system equipped with multiple receive antennas.

Contents

Notation

$(\cdot)^T$	denotes transpose operation
$(\cdot)^H$	denotes Hermitian transposition
$[\cdot]_{k,\ell}$	denotes the (k,ℓ)th entry of the enclosed matrix
int$\{x\}$	denotes the maximum integer not exceeding x
$\|\cdot\|$	denotes the Euclidean norm
tr$\{\cdot\}$	denotes the trace of a matrix
$\mathbf{I_N}$	denotes the identity matrix of order N
rank$\{\cdot\}$	denotes the rank of the enclosed matrix
$E\{\cdot\}$	denotes the expectation operator
$\mathbf{M}^{-1}$	denotes the inverse of a square matrix $\mathbf{M}$
Re$\{\cdot\}$	denotes the real part of a complex-valued quantity
Im$\{\cdot\}$	denotes the imaginary parts of a complex-valued quantity
$\|\cdot\|$	denotes the magnitude a complex-valued quantity

Acronyms

A/D	Analog-to-Digital
AGC	Automatic Gain Control
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BLAST	Bell-labs LAyered Space-Time
BPSK	Binary Phase Shift Keying
BS	Base Station
CDMA	Code-Division Multiple-Access
CIR	Channel Impulse Response
CP	Cyclic Prefix
CSI	Channel State Information
D/A	Digital-to-Analog
DD	Decision-Direct
DFT	Discrete Fourier Transform
DOA	Direction-Of-Arrival
DSP	Digital Signal Processing
EGC	Equal Gain Combining
ETSI	European Telecommunication Standardization Institute
EVD	Eigenvalue Decomposition
FDMA	Frequency-Division Multiple-Access

FFT	Fast Fourier Transform
HiperLAN	High Performance Local Area Network
IC	Interference Cancellation
ICI	Ideal Channel Information
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
ITU	International Telecommunication Union
LHS	Left-Hand-Side
LMS	Least-Mean-Square
LMS-SCE	Least-Mean-Square Structured Channel Estimator
LMS-UCE	Least-Mean-Square Unstructured Channel Estimator
L-TWF	Linear Transmit Wiener Filter
MAI	Multiple Access Interference
MC-CDMA	Multi-Carrier Code-Division Multiple-Access
MIL	Matrix Inversion Lemma
ML	Maximum-Likelihood
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
MSE	Mean Square Error
MSEE	Mean Square Estimation Error
MT	Mobile Terminal
m-TIR	Modified Total Interference Removal
MUD	Multi-user Detection
NL-TWF	Non-Linear Transmit Wiener Filter
NL-TZF	Non-Linear Transmit Zero-Forcing Filter
NSA	Non-Adaptive Subcarrier Assignment

OCA	Optimal Code Assignment
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple-Access
OOS	Optimal Ordering Strategy
OSA	Optimal Subcarrier Assignment
PD	Pure Despreading
PAPR	Peak-to-Average Power-Ratio
PCK	Perfect Channel Knowledge
PIC	Parallel Interference Cancellation
PPIC	Partial Parallel Interference Cancellation
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
QoS	Quality of Service
RHS	Right-Hand-Side
RLS	Recursive-Least-Square
RF	Radio Frequency
SCE	Structured Channel Estimator
SIC	Serial Interference Cancellation
SINR	Signal over Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
STBC	Space-Time Block Code
SUB	Single-User Bound
SUD	Single User Detection
UCE	Unstructured Channel Estimator
UMTS	Universal Mobile Telecommunication System
ULA	Uniform Linear Array
TDD	Time-Division-Duplex

TDMA	Time-Division Multiple-Access
TIR	Total Interference Removal
TH	Tomlinson-Harashima
THP	Tomlinson-Harashima Pre-coding
TWF	Transmit Wiener Filter
TZF	Transmit Zero-Forcing Filter
V-BLAST	Vertical Bell-labs LAyered Space-Time
WH	Walsh-Hadamard
WLAN	Wireless Local Area Network
ZF	Zero Forcing
ZF-L	ZF Linear
ZF-THP	Zero Forcing Tomlinson-Harashima Pre-coding

List of Figures

LIST OF FIGURES

List of Tables

Chapter 1

Introduction

The rapid growth of wireless communications and its pervasive use in all walks of the life are changing the way we communicate to each other. Cellular phones have become a incredible business tool and part of everyday life in most developed countries, and are rapidly supplanting antiquated wireline systems. In addition, wireless local area networks are currently supplementing or replacing wired networks in many homes, offices, and campuses. Many new applications, including wireless sensor networks, automated highways and factories, smart phones and appliances, and remote telemedicine, are emerging from research ideas to concrete systems. On the other hand, the design of an efficient, reliable, and robust wireless communication network for each of these emerging applications poses many technical challenges that must be addressed.

There are several fundamental aspects of wireless communications that make the design challenging and interesting. Firstly, the *radio spectrum*, which is a scarce and very expensive resource that must be allocated to many different applications and systems. In the U.S., companies spent over nine billion dollars for second generation cellular licenses, and the auctions in Europe for third generation cellular spectrum garnered around 100 billion dollars. The spectrum ob-

tained through these auctions must be used efficiently to get a reasonable return on its investment, and it must also be reused over and over in the same geographical area, thus requiring cellular system designs with high capacity and good performance. Secondly, the phenomena of *fading*. In wireless communications the transmitted signals arrive at the receive side through multipath propagations arising from scattering, reflection, and diffraction of the radiated energy by objects in the environment or refraction in the medium. Fluctuations of the signal power results from the constructive or destructive combination of these multipath components. In addition, the transmitted signal may experience random fluctuations in time if the transmitter, receiver, or surrounding objects are moving, due to changing reflections and attenuations. This means that the characteristics of the channel may also appear to change randomly with time making the design of reliable networks difficult. Thirdly, the *co-channel interference* arising from the frequency reuse. This may occur between transmitters communicating with a common receiver (e.g., the uplink of cellular system), between signals from a single transmitter to multiple receivers (e.g., downlink of a cellular system) or between different transmitter-receiver pairs (e.g., interference between users in different cells).

In the sequel, we give a brief history of mobile radio communications to illustrate how the success of wireless communications is related to innovations in transmission technologies. Then, the challenges faced in this work and the proposed innovations are discussed and eventually the structure of the thesis is presented.

1.1 Wireless communications

The pioneering era

Many scientists have been working on radio waves before Marconi established the first successful and practical radio system. Starting in 1894 with his first electrical experiments, in 1901 his radio telegraph system sent the first radio signals across the Atlantic ocean Brittain (1992). Since then the idea of mobile radio was investigated for practical implementations and until the 1920s mobile radio communications mainly made use of Morse Code. The first mobile radio telephones were employed in 1921 by the Detroit Police Department's radio bureau, that began experimentation with a band near 2 MHz for vehicular mobile services Bowers (1978).

These first pioneering experiments were followed by many others throughout the years, but although the overall performance was improved and new features were added, the costs of the calls remained prohibitive for commercial applications. The main limitation to further developments was the small system capacity. Systems were designed such that all mobile terminals accessed one radio access point that served the whole system. Such a configuration limited the maximum number of users that could be accepted in the system to the number of allocated frequency channels. Moreover the range of the system coincided with the area covered by the transmitting antenna at the single access point.

The cellular concept

The advent of the cellular concept (1947, AT&T Bell Laboratories Ring (1947)) was a crucial contribution in the development of mobile communications. In this idea, the served area is split in smaller regions, called cells. Due to its reduced dimensions, each cell requires a much lower power to be covered and, most

importantly, allows the reuse of the same frequency bandwidth in other (most often non-adjacent) cells. By taking advantage of the ground wave attenuation, proportional to a power (> 2) of the distance between cells, such a system is able to support a much larger number of users. As cell size decreases, system capacity increases because transmit power can be disproportionably reduced. Another major step toward the success of cellular communications was taken when in 1973 Dr. Martin Cooper for Motorola filed the first patent Cooper (Granted on September, 1975) describing a portable mobile telephone. Previously mobile communications were mainly designed for terminals mounted in cars or trains.

The first generation of commercial systems

However, mobile cellular systems were not introduced for commercial use until the early 1980s when the so-called first generation mobile systems were deployed all over the world. Semiconductor technology and microprocessors made smaller, light-weight, and more sophisticated mobile systems a practical reality for many more users. In the beginning, each country developed its own system. Analog cellular telephone systems experienced a very rapid growth in Europe, particularly in Scandinavia (Nordic Mobile Telephony (NMT), 1981), and the United Kingdom (Total Access Communication System (TACS), 1985). Advanced Mobile Phone Service (AMPS) was commercially started in 1979 in Japan and in 1983 in the United States.

2G: GSM

Driven by the success of NMT, which represented the first multinational cellular system, European countries decided to create a pan-European mobile service with advanced features and easy roaming. In 1982 the Conference of European Posts and Telecommunications (CEPT) formed the Group Spécial Mobile(GSM)

in order to develop a second-generation (2G) mobile cellular radio system. All first generation systems were analog and based on frequency division multiple access (FDMA). GSM was developed as a digital system, and the multiple access scheme was based on a hybrid between TDMA and FDMA. The new system could provide a much larger capacity and many more standardised services than the previous ones; this is one of the key reasons for the success of the GSM standard. It soon became a world standard, and by June 2001 there were more than 500 millions GSM subscribers worldwide. Eventually, also the United States followed the choice of adopting digital schemes for second generation mobile services in 1990 with IS-54 and in 1994 with IS-95. IS-95 is the first commercial mobile communication system based on the CDMA technology.

3G: the advent of CDMA

At the end of the 1990s, it became clear that the allocated frequencies for GSM were not sufficient to indefinitely support the growing number of users and the higher data rates required for multimedia services. Third generation (3G) systems were studied and developed all around the world. Under the umbrella of global institutions like the ETSI and the ITU, two main proposals emerged, both based on CDMA: the US CDMA-2000 and the Japanese-European UMTS. In a CDMA scheme all users share the same bandwidths and timeslots, and they are distinguished only by the use of orthogonal codes: each user's message is spread by a specific code sequence, quasi orthogonal to all other user sequences in the system.

CDMA has many advantages with respect to more conventional access schemes like FDMA and TDMA: higher spectral efficiency as well as higher flexibility and robustness against channel fading. On the other hand, it has the following main drawback. In practical applications the spreading sequences loose their orthog-

onality due to distortions introduced by the radio channel and, thus, multiple access interference occurs, thereby limiting the system capacity. The development of advanced multi-user receivers has been one of the solutions proposed to limit the destructive impact of the co-channel interference on system performance. The term MUD Verdú (1998) indicates those techniques that exploit the knowledge of channel parameters, spreading sequences and transmitted data of all users present in a cell to improve the receiver performance. The optimum multi-user detector obtains jointly optimum decisions for all users by selecting the most likely sequence of transmitted symbols Verdú (1998). Since the optimum multi-user detector is in practice too complex to be implemented, several sub-optimum schemes have been studied. Among them are: linear receivers like the decorrelator or zero-forcing receiver Lupas & Verdú (1990), the linear minimum mean square error receiver Madhow & Honig (1990) and interference cancellation schemes Patel & Holtzman (1990), Divsalar *et al.* (1998).

4G: coming soon

While 3G is just transforming itself into a reality from an engineers dream, novel communication and information services are being introduced almost daily and the demands for higher data rates and communication capacity is continuously growing. The demand for these services is increasing at an extremely rapid pace and these trends are likely to continue for several years. The radio spectrum available for wireless services is extremely scarce. As a consequence, a prime issue in current wireless systems is the conflict between the increasing demand for wireless services and the scarce electromagnetic spectrum. Spectral efficiency is therefore of primary concern in the design of future wireless data communication systems with the omnipresent bandwidth constraint. For these reasons, wireless designers are already working hard and constantly to develop innovative trans-

mission techniques which are able to face with these large number of challenges. These do not include only the limited availability of radio frequency spectrum and the impairments introduced by the time-varying wireless environment but, also, the need to guarantee better quality of services, fewer dropped calls, higher network capacity and user coverage. Research in multicarrier modulations as a means to meet such achievements, while data back to nineties, has experienced a boom of activity in recent years. Among all the existing multicarrier modulation techniques, MC-CDMA and OFDMA Fazel & Kaiser (2003) represent two promising candidates for 4G communications because of their attractive features such as high spectral efficiency and flexibility for integrated multimedia applications. These schemes can support multiple users by exploiting either code- or frequency-division multiplexing. In particular, in MC-CDMA the data of different users are spread in the frequency domain on an given group of subcarriers using pre-assigned signature sequences while in OFDMA an exclusive set of orthogonal subcarriers is allocated to each user.

Like all CDMA-based transmission techniques, the main impairment of an MC-CDMA system is the MAI which occurs in the presence of multipath propagation. Several advanced MUD techniques are already available for interference mitigation Moshavi (1996). Unfortunately, in spite of their effectiveness, these methods are not suited for downlink applications due to their relatively high computational complexity. As an alternative to MUD, interference mitigation can be accomplished at the transmit end using pre-filtering schemes. The idea behind pre-filtering is to adapt the transmitted signals to the channel conditions so that interference is reduced and the signal at the receiver appears undistorted. In this way, low complex detection schemes can be employed at the remote units while most of the computational burden is moved to the BS, where power consumption and complexity are not issues of concern. The aim of this dissertation is to

provide guidelines for the design of *linear* and *non-linear* pre-filtering schemes aimed at enhancing the performance of conventional MC-CDMA and/or OFDMA downlink transmissions.

1.2 Outline of Dissertation

The outline of the dissertation is as follows.

The present Chapter introduces the motivations and summarizes the structure of this dissertation.

Chapter 2 provides the basic concepts of OFDMA and MC-CDMA systems.

In Chapter 3, in order to make comparisons and underline the main contributions of this dissertation, we briefly revise some well-known data detection techniques employed in conventional MC-CDMA transmissions.

Chapter 4 discusses several linear pre-filtering techniques for the downlink of TDD MC-CDMA systems equipped with multiple transmit antennas. After discussing existing solutions, we derive three novel pre-coding schemes based on different approaches. The first one aims at minimizing the sum of the MSEs at all MTs and turns out to be the more efficient than the others although it requires knowledge of the noise power at the BS. The second one is based on the ZF criterion, thereby dispensing knowledge of the noise power at the BS. The last one looks for the minimum of an alternative cost function that depends on the SINRs of all active users.

In Chapter 5 we propose a unified framework comprising both MC-CDMA and OFDMA, and discuss non-linear pre-filtering based on THP. In designing the pre-coding matrices we aim at minimizing the sum of MSEs at all MTs under a constraint on the overall transmit power. The use of a unified framework comprising both MC-CDMA and OFDMA allows a comparison between these

multiple-access technologies under the same operating conditions.

An implicit assumption in any pre-filtering scheme is that CSI is available at the transmit side. In TDD systems this information can be achieved by exploiting the channel reciprocity between alternative uplink and downlink transmissions. If channel variations are sufficiently slow, the channel estimate derived at the BS during an uplink time-slot can be reused for pre-filtering in the subsequent downlink time-slot. Clearly, this solution is not suited for FDD networks, where channel estimates obtained at each MT must be fed back to the BS using an uplink control channel. For this reason, Chapter 6 deals with channel acquisition in the uplink of an MC-CDMA system. Channel acquisition is performed jointly with noise power estimation following two different approaches. The first assumes independently faded subcarriers while the second exploits the fading correlation across the signal bandwidth to improve the system performance. Both schemes are based on the ML criterion and exploit some training blocks carrying known symbols. We also address the problem of channel tracking using the LMS technique and data decisions provided by a PPIC receiver.

Chapter 2

Multi-carrier modulation techniques

As it is well-known, in frequency-selective fading channels the received signal is typically affected by intersymbol interference. The classical approach adopted in single carrier systems is time-domain equalization. However, the number of operations per signaling interval grows linearly with the number of interfering symbols or, equivalently, with the data rate. As a result, conventional time-domain equalizers are not suitable for high-speed transmissions with channel delay spreads extending over tens of symbol intervals. This has motivated the adoption of multicarrier modulations as a computationally efficient alternative for facing with the severe impairments of multipath propagation. The idea behind multicarrier modulations is to split a high-rate data stream into a number of low-rate streams which are transmitted in parallel on adjacent subchannels. As is shown in Figure 2.1, reducing the data rate or, equivalently, increasing the symbol duration, makes the frequency selective fading channel appear flat on each subcarrier, thereby limiting the intersymbol interference.

Thanks to their high spectral efficiency and flexibility in supporting integrated

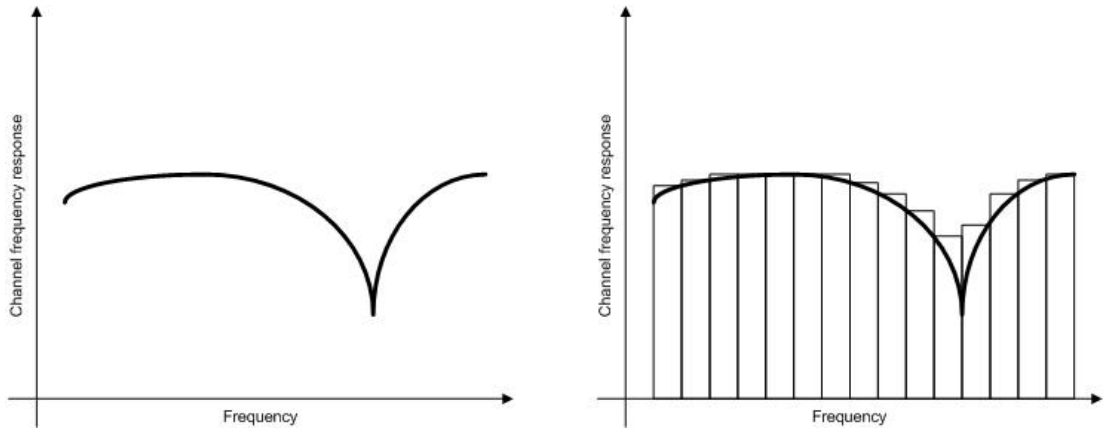

Figure 2.1: Channel frequency response. In multicarrier system each information-bearing symbol undergoes frequency flat fading channel.

multimedia services among the large variety of multicarrier modulations MC-CDMA and OFDMA are considered, as the most promising multicarrier multiple-access technologies for next generation high-speed wireless networks. In the next, we briefly revise the system structure of these two multiple access technologies.

2.1 OFDMA

OFDMA is a multiplexing technique in which several users simultaneously transmit their own data by modulating an exclusive set of orthogonal subcarriers. Its main advantage is that separating different users through FDMA techniques at the subcarrier level can mitigate MAI within a cell. Also, as mentioned earlier, compared with single-carrier multiple-access systems, OFDMA offers increased robustness to narrowband interference, allows straightforward dynamic channel assignment, and does not need adaptive time-domain equalizers, since channel equalization is performed in the frequency domain through one-tap multipliers.

This technique was originally implemented using a bank of Nyquist filters

which provide a set of continuous-time orthogonal basis functions. Using very fast and cost effective digital signal processors, OFDMA modulation is now implemented using DFT techniques. This has motivated the adoption of OFDMA as a standard for the new WLAN IEEE 802.11a and the ETSI HiperLAN/2. OFDMA has also been proposed for digital cable television systems.

Transmitter structure

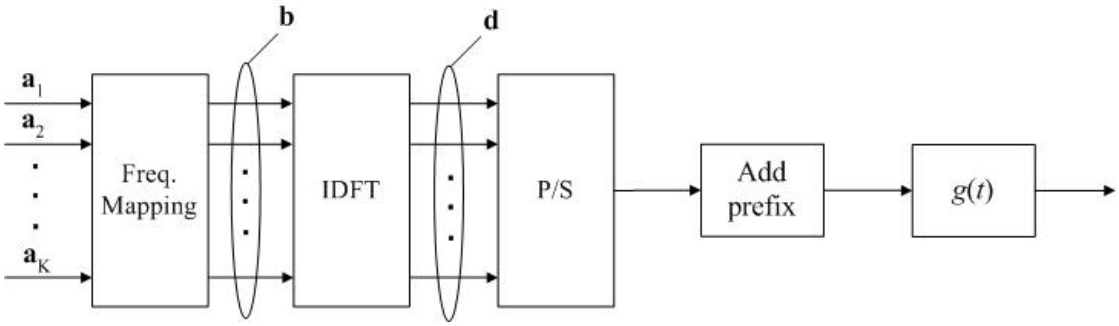

Figure 2.2: Block diagram of an OFDMA transmitter.

We consider the downlink of an OFDMA system employing N subcarriers and accommodating a maximum of K simultaneously active users. The base station communicates with each user on a set of $Q = N/K$ assigned subcarriers. The block diagram of the transmitter is shown in Figure 2.2. The nth symbol transmitted to the kth user is denoted as $a_k(n)$ and belongs to an M-QAM or M-PSK constellation with average energy σ_a^2. For convenience, we collect the users' data into Q-dimensional vectors $\mathbf{a}_k = [a_k(1), a_k(2), \ldots, a_k(Q)]^T$ for $k = 1, 2, \ldots, K$. The entries of each stream of data $\mathbf{a}_k$ are properly mapped on specified positions of an N-dimensional vector $\mathbf{b} = [b(1), b(2), \ldots, b(N)]^T$ with $b(n)$ defined as

$$b(n) = a_k(\ell) \quad \text{if } n = p_k(\ell) \tag{2.1}$$

where $p_k(\ell)$ denotes the index of the subcarrier modulated by the data symbol $a_k(\ell)$. Note that the indexes $p_k(\ell)$ for $\ell = 1, 2, \ldots, Q$ and $k = 1, 2, \ldots, K$ can be any integer in the interval $[0, N-1]$ but must satisfy the following relation

$$\sum_{k=1}^{K} \sum_{\ell=1}^{Q} p_k(\ell) = 1 \tag{2.2}$$

ensuring that no more than one user transmits on a given subcarrier.

The resulting vector $\mathbf{b}$ is then fed to an OFDM modulator, which comprises an N-point IDFT unit and the insertion of an N_G-point CP larger than the channel impulse response. The cyclic prefix serves to eliminate inter-block interference and makes the linear convolution of the symbols with the channel look like a circular convolution, which is essential for demodulation based on DFT. This produces an $(N + N_G)$-dimensional vector $\mathbf{d} = [d(-N_G), d(-N_G + 1), \ldots, d(N-1)]^T$ of time domain samples where $d(n) = d(n + N)$ for $-N_G \leq n \leq -1$ and

$$d(n) = \sum_{\ell=1}^{N} b(\ell) e^{-j2\pi \frac{\ell n}{N}}, \quad 0 \leq n \leq N-1. \tag{2.3}$$

The resulting vector $\mathbf{d}$ is finally passed to a digital-to-analog (D/A) converter with impulse response $g(t)$ and signalling interval T_s. The complex envelope of the signal transmitted takes the form

$$s(t) = \sum_{n=-N_G}^{N-1} d(n) g(t - nT_s) \tag{2.4}$$

where $g(t)$ has a root-raised cosine Fourier transform with some roll-off α.

Receiver structure

Figure 2.3 illustrates the block diagram of an OFDMA receiver. The signals transmitted by the active users propagate through different channels and undergo

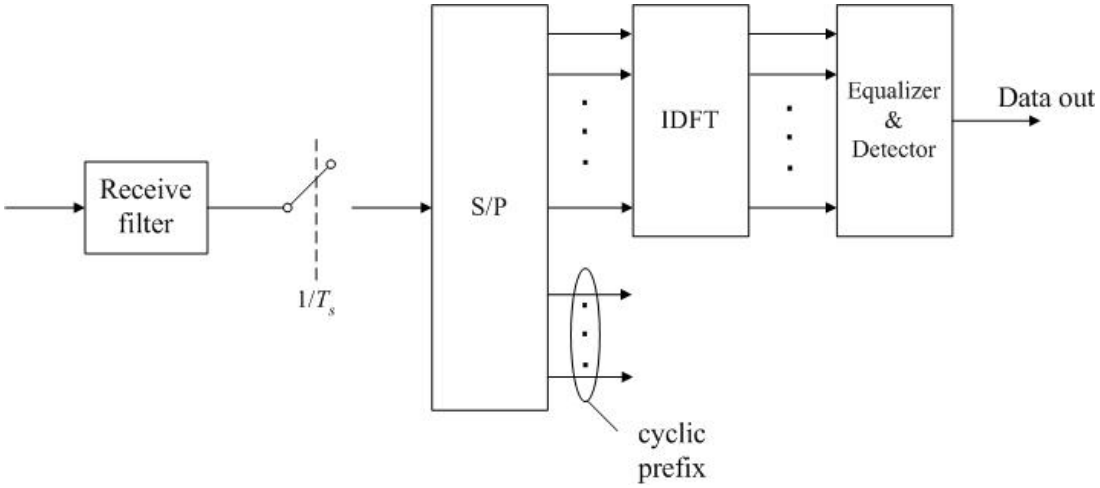

Figure 2.3: Block diagram of an OFDMA receiver.

frequency selective fading. At the receiver, the incoming waveforms are first filtered and then sampled at rate $1/T_s$. This produces

$$x(n) = \sum_{\ell=-N_G}^{N-1} d(\ell)h(n-\ell) + w(n) \tag{2.5}$$

where $w(n)$ is thermal noise and $h(n)$ is the sample of the overall channel impulse response $h(t)$ (including the physical channel, the shaping pulse at the transmitter and the whitening matched filter at the receiver) at time $t = nT_s$. Statistical models for the channel impulse response of a fading multipath channel have been described in details in literature over the past years. Since its out of the scope of this work to provide a description of such models, we refer to some excellent references on this field such as Proakis (1995)-Bello (1963).

After discarding the cyclic prefix and using a matrix notation, the samples at

the output of the matched filter can be expressed as follows

$$\begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix} = \begin{bmatrix} h(0) & 0 & \cdots & \cdots & 0 & h(L) & \cdots & h(1) \\ \vdots & \ddots & \ddots & & & 0 & \ddots & \vdots \\ \vdots & & h(0) & \ddots & & & \ddots & h(L) \\ h(L) & & \vdots & \ddots & 0 & & & 0 \\ 0 & \ddots & \vdots & & h(0) & \ddots & & \vdots \\ \vdots & \ddots & h(L) & & \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & h(L) & \cdots & \cdots & h(0) \end{bmatrix} \begin{bmatrix} d(0) \\ d(1) \\ \vdots \\ d(N-1) \end{bmatrix} \tag{2.6}$$

where L is the length in sampling periods of the channel impulse response. Inspections of the above equation reveals that thanks to the introduction of the CP at the transmitter and its removal at the receiver, the resulting time domain samples are related to the input data symbols through the channel matrix in (2.6). The latter is a circulant matrix, i.e., its rows are composed of cyclically shifted versions of a given sequence Horn & Johnson (1985). In other words, the effect of the cyclic prefix is to make the channel look like circular convolution instead of linear convolution, thereby completely removing the intersymbol interference.

From matrix theory it turns out that these kind of matrices have a very interesting and useful property Lancaster (1969). The eigenvectors are independent of the specific channel coefficients and are always given by complex exponentials. To be more precise, the EVD of the circulant channel matrix in (2.6) is

$$\begin{bmatrix} h(0) & 0 & \cdots & \cdots & 0 & h(L) & \cdots & h(1) \\ \vdots & \ddots & \ddots & & & 0 & \ddots & \vdots \\ \vdots & & h(0) & \ddots & & & \ddots & h(L) \\ h(L) & & \vdots & \ddots & 0 & & & 0 \\ 0 & \ddots & \vdots & & h(0) & \ddots & & \vdots \\ \vdots & \ddots & h(L) & & \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & h(L) & \cdots & \cdots & h(0) \end{bmatrix} = \mathbf{F}^H \mathbf{H} \mathbf{F} \tag{2.7}$$

where $\mathbf{F}$ is the $N \times N$ unitary DFT matrix whose entries are given by

$$[\mathbf{F}]_{n,\ell} = e^{-j2\pi\, n\ell/N} \qquad 0 \le n \le N-1\,, \quad 0 \le \ell \le N-1. \tag{2.8}$$

while $\mathbf{H} = \mathrm{diag}\,\{\mathrm{H}(0), \mathrm{H}(1), \ldots, \mathrm{H(n)}\}$ is $N \times N$ diagonal matrix with

$$H(n) = \sum_{\ell=0}^{L-1} h(\ell) e^{-j2\pi \frac{\ell n}{N}}. \tag{2.9}$$

Collecting the above fact together, we see that at the receiver the DFT outputs $y(n)$ for $n = 0, 1, \ldots, N-1$ can be written as

$$y(n) = H(n)b(n) + w(n) \tag{2.10}$$

or in matrix notation

$$\mathbf{y} = \mathbf{H}\mathbf{b} + \mathbf{w}. \tag{2.11}$$

From the above equation it follows that thanks to the multicarrier approach, the original frequency-selective channel with inter-symbol and inter-block interference is transformed into a set of parallel flat subchannels, which can be straightforwardly equalized at receiver side using a simple bank of one-tap multipliers.

2.2 MC-CDMA

MC-CDMA is a multiplexing technique that was introduced in September 1993 independently by Fazel (1993) and Yee *et al.* (1993) and combines OFDM with direct-sequence CDMA. Thanks to its high spectral efficiency and flexibility in supporting integrated multimedia services, it is considered as a promising multiple-access technology for next generation high-speed wireless networks Fazel & Kaiser (2003).

In the next, we briefly revise the transmitter and receiver structure of an MC-CDMA network.

Transmitter structure

We consider the downlink of an MC-CDMA network in which the total number of subcarriers, N, is divided into smaller groups of Q elements Fazel & Kaiser (2003). By exploiting the subcarriers of a given group, the BS simultaneously communicate with K active users ($K \leq Q$), which are separated by their specific spreading codes (typically chosen from an orthogonal set). Without loss of generality, we concentrate on a single group.

Figure 2.4 illustrates the transmit side of an MC-CDMA system. The symbol a_k of the kth user is spread over Q chips using a *unit-energy* spreading sequence $\mathbf{c}_k = [c_k(1), c_k(2), \ldots, c_k(Q)]^T$. Depending on the specific application, a large amount of different spreading codes can be used Popovic (1999). For example, in order to reduce the PAPR of the transmitted signal Goaly or Zadoff-Chu codes can be employed. The latter are very attractive especially for uplink transmissions where high PAPR gives rise to the need for highly linear power amplifiers at the transmitter (i.e., the mobile terminal), which directly translates into expensive devices, thereby limiting widespread industrial applications. However, a detailed

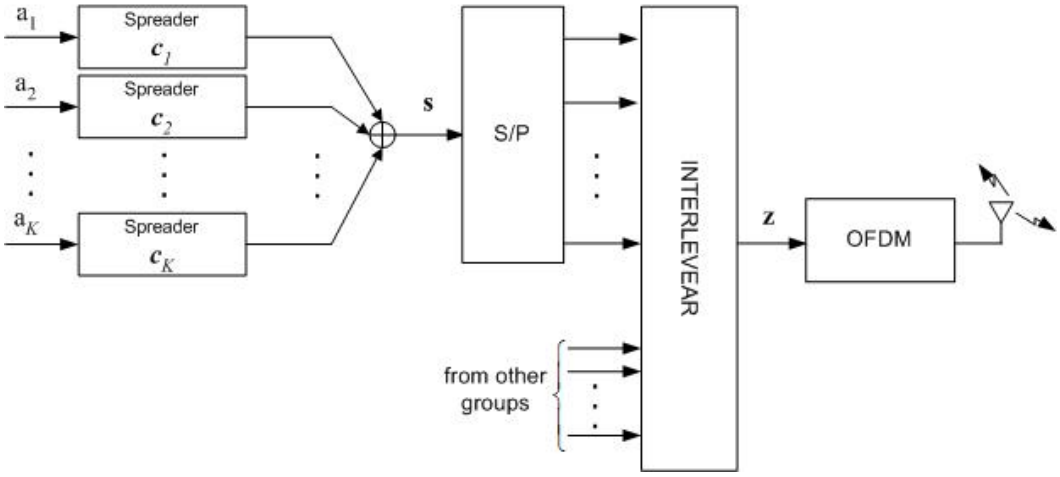

Figure 2.4: Block diagram of an MC-CDMA transmitter.

analysis of the intrinsic properties of different spreading codes is out of the scope of this dissertation. In the next, we resort to orthogonal spreading sequences taken from a set of WH codes. This is tantamount to setting $c_k(n) \in \{\pm 1/\sqrt{Q}\}$.

After spreading operation, the contributions of all users are summed chip-by-chip to form the following transmit signal vector

$$\mathbf{s} = \sum_{k=1}^{K} a_k \mathbf{c}_k. \tag{2.12}$$

which is then frequency interleaved with the contributions of the other groups. The resulting N-dimensional vector $\mathbf{z}$ is finally mapped on N subcarriers using an OFDM modulator, comprising an IDFT unit and the insertion of a CP.

Receiver structure

Figure 2.5 illustrates the receive side of an MC-CDMA system. Without loss of generality, we concentrate on the kth mobile terminal. The incoming waveform is first filtered and then passed to an OFDM demodulator. The latter discards the CP and performs an N-point DFT operation before passing the resulting

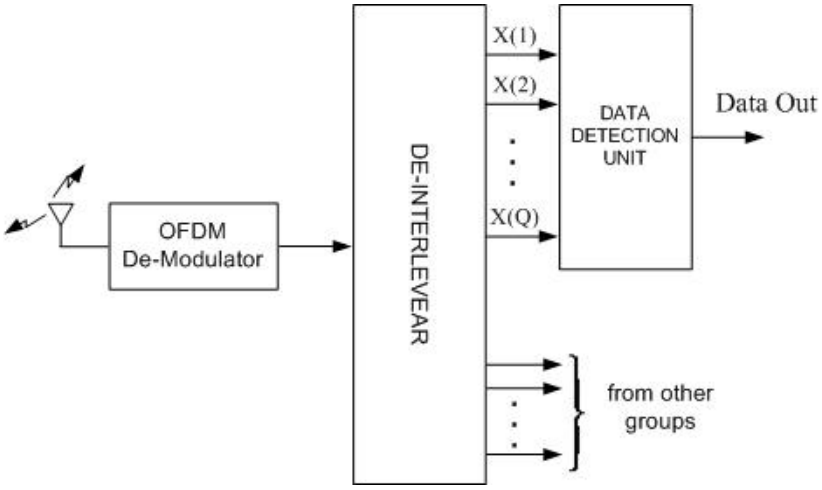

Figure 2.5: Block diagram of an MC-CDMA receiver.

frequency-domain samples to the de-interleaver unit. The output is then fed to the data detection unit which delivers an estimate of the transmitted data symbols.

Denoting $\mathbf{X} = [X(1),\, X(2), \ldots, X(Q)]^T$ the demodulator outputs corresponding to the Q subcarriers of the considered group and assuming ideal frequency and timing synchronization, we have

$$X(n) = H(n) \sum_{k=1}^{K} a_k c_k(n) + w(n) \quad 1 \leq n \leq Q \tag{2.13}$$

where $H(n)$ is the channel frequency response over the nth subcarrier and $w(n)$ is thermal noise, which is modelled as a Gaussian random variable with zero mean and variance σ_w^2. From the above equation, we see that, thanks to the underlying OFDM, the MC-CDMA received signal is not affected by inter-symbol interference. As in (2.10), the frequency selective channel introduces only a phase and amplitude distortion on each subcarrier. Differently from OFDMA systems where the active users are separated in the frequency domain, in MC-CDMA the channel distortions must be properly compensated to avoid that the code

orthogonality is lost and multiple-access interference occurs. The problem of data detection in the downlink of MC-CDMA systems will be discussed in the next Chapter where several solutions are presented.

Chapter 3

Data Detection in MC-CDMA

Data detection in the downlink of MC-CDMA transmissions may be accomplished using either SUDs or IC-based receivers. The former detect the user of interest without exploiting any information about the interfering users. This allows the remote units to maintain low computational complexity but makes the mobile terminals more exposed to the interference, which strongly degrades the system performance. This limits the use of SUDs especially in heavy-loaded systems and calls for more sophisticated detection schemes based on IC. In these schemes interfering signals are detected and subtracted from the received waveform before detection of the desired user's data. In spite of their effectiveness, however, IC-based schemes are quite unattractive for downlink applications due to their relatively large complexity. Moreover, in addition to the channel response, IC-based schemes require knowledge of the number of active users. More precisely, they need knowledge of which codes are effectively employed and how they are assigned to the active users in the current transmission. In practical applications only the BS has this information meaning that it should be sent to the mobile terminals using a feedback channel or estimated in some manner at the receivers.

In the sequel, we discuss several well-known SUDs and IC-based receivers and

provide some simulation results to highlight their effectiveness.

3.1 Single-user detection

As mentioned earlier, SUDs are usually employed in data detection of downlink transmissions to keep a low computational burden at the remote units. The block diagram of a SUD is sketched in Figure 3.1. Without loss of generality, we concentrate on the mth mobile terminal and assume that each mobile terminal is equipped with a single receive antenna.

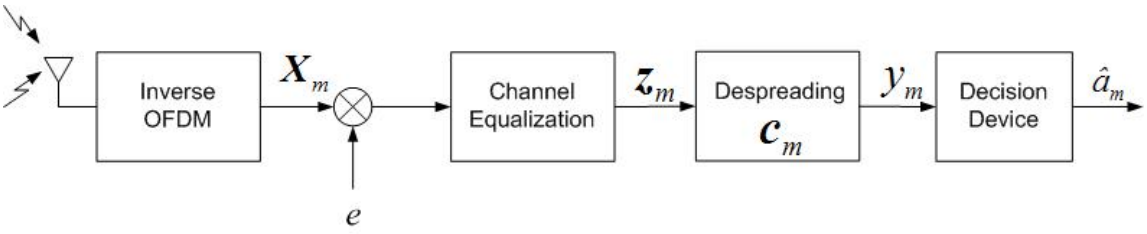

Figure 3.1: Block diagram of an MC-CDMA receiver employing SUD.

The incoming waveforms are implicitly combined by the receive antenna and passed to an OFDM demodulator, which eliminates the CP and performs a DFT operation. The DFT output $\mathbf{X}_m$ is first ideally scaled by a real parameter e that can be thought of as being part of an automatic gain control (which does not impair the signal-to-noise ratio at the receiver) and is then passed to an equalization unit which performs channel equalization. The latter operates in the frequency domain through one-tap multipliers and aims at mitigating on each subcarrier the phase and amplitude distortions due to the multipath channel. The output of the equalization unit takes the form

$$\mathbf{z}_m = \mathbf{A}_m \mathbf{X}_m \tag{3.1}$$

where $\mathbf{A}_m$ is a $Q \times Q$ diagonal matrix given by

$$\mathbf{A}_m = \begin{bmatrix} A_m(1) & 0 & \cdots & 0 \\ 0 & A_m(2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m(Q) \end{bmatrix} \tag{3.2}$$

whose entries represent the Q complex-valued coefficients employed to face with the channel distortions introduced on those subcarriers where a_m was transmitted on. As shown in Figure 3.1, the resulting vector $\mathbf{z}_m$ is then fed to a despreading unit which correlates its input with the spreading code $\mathbf{c}_m$ assigned to the mth user. This produces

$$y_m = \mathbf{c}_m^H \mathbf{z}_m \tag{3.3}$$

which is finally passed to a threshold device to obtain an estimate of a_m.

In the sequel we briefly revise some well-know channel equalization strategies.

3.1.1 PD

The PD channel equalization unit does not perform any operation. This is tantamount to setting

$$A_m(n) = 1 \quad n = 1, 2, \ldots, Q. \tag{3.4}$$

Note that this approach dispenses channel knowledge at the receive side.

3.1.2 MRC

The MRC equalization unit multiplies each subcarrier with a complex-valued coefficient proportional to the complex conjugate of the corresponding channel gain, i.e.,

$$A_m(n) = H_m^*(n) \quad n = 1, 2, \ldots, Q. \tag{3.5}$$

It can be shown that this approach is the best against the thermal noise since it leads to the maximization of the signal-to-noise ratio at the output of the SUD Proakis (1995). Hence, in a single user environment (i.e., without interference), where the main impairment is represented by the thermal noise, MRC is the optimal equalization strategy. In a multi-user system, where a large number of users are simultaneously active, MRC becomes the worst since it strongly enhances the multiple access interference Tse & Viswanath (2004).

3.1.3 EGC

In this case the channel equalization coefficients have unit amplitude and only compensate for the channel phase, i.e.,

$$A_m(n) = \frac{H_m^*(n)}{|H_m(n)|} \quad n = 1, 2, \ldots, Q. \tag{3.6}$$

From the above equation it follows that the EGC technique requires only information about the phase of the corresponding channel gain. This makes the EGC the simplest channel equalization strategy.

3.1.4 ZF

The ZF approach employs the following equalization coefficients

$$A_m(n) = \frac{1}{H_m(n)} \quad n = 1, 2, \ldots, Q \tag{3.7}$$

which aim at removing the channel distortions from each subcarrier. This leads to the complete elimination of the multiple access interference since the cancellation

of the channel gain restores the orthogonality among the spreading codes. However, this comes at price of an increase of the thermal noise (i.e., noise enhancement) when $H_m(n)$ is much less than unity. This phenomena highly degrades the system performance Fazel & Kaiser (2003).

3.1.5 MMSE

The MMSE technique leads to the minimization of the following cost function

$$J_m(n) = E\left\{|\varepsilon_m(n)|^2\right\} \quad n = 1, 2, \ldots, Q \tag{3.8}$$

where

$$\varepsilon_m(n) = A_m(n)X_m(n) - d_m \quad n = 1, 2, \ldots, Q \tag{3.9}$$

is the error between the transmitted symbol and the output of the equalizer. The minimum of (3.8) can be found using the principle of orthogonality. The latter states that the mean square error $J_m(n)$ is minimized when the equalizer coefficient $A_m(n)$ is chosen such that $\varepsilon_m(n)$ is orthogonal to $X_m^*(n)$, i.e.,

$$E\left\{\varepsilon_m(n)X_m^*(n)\right\} = 0. \tag{3.10}$$

This produces

$$A_m(n) = \frac{H_m^*(n)}{|H_m(n)|^2 + \sigma^2}. \tag{3.11}$$

Inspection of (3.8) reveals that, in addition to the channel response, the MMSE detector requires knowledge of the noise power, which must estimated in some manner. An alternative solution to overcome the additional complexity required for the estimation of σ^2 is to employ a *suboptimum* scheme in which the noise

power can be replaced with a threshold parameter λ that must be properly set by the system designers Fazel & Kaiser (2003).

3.2 Interference Cancellation Receivers

As already mentioned, an effective and alternative solution to single-user detectors is represented by IC-based receivers. The latter are *multistage* non-linear detection schemes that can be largely divided into two main categories. The first comprises those receivers in which the multiple-access interference is estimated using tentative data decisions and subtracted out in *parallel* from the received signal before detecting the user of interest. For this reason they are refereed to as PIC receivers. The second category comprises those schemes based on *successive* interference cancellation (SIC) in which the interfering signal are *successively* decoded and subtracted from the received samples to reduce the amount of multiple-access interference.

As shown in Fazel & Kaiser (2003), PIC receivers are more attractive than SIC counterpart for those applications such as downlink transmissions in which the interfering signals have the same power. In the past few years, several advanced PIC receivers have been proposed and discussed. A complete description of each one is out of the scope of this thesis. Hence, in the next we briefly revise the PPIC detector proposed in Divsalar *et al.* (1998), which achieves a reasonable trade-off between performance and complexity.

3.2.1 PPIC

The PPIC receiver is a multistage receiver in which MAI is estimated using tentative data decisions and *partially* subtracted from the DFT output. Assuming that K users are simultaneously active, at the ℓth stage the PPIC detector computes

the vectors Divsalar *et al.* (1998)-Peijun & Rappaport (1998)

$$\mathbf{Z}_k^{(\ell)}(m) = \mathbf{X}(m) - \sum_{\substack{j=1 \\ j\neq k}}^{K} \gamma_j \hat{a}_j^{(\ell-1)}(m)\mathbf{u}_j(m) \qquad k = 1, 2, \ldots, K \tag{3.12}$$

where $\{\hat{a}_j^{(\ell-1)}(m)\}$ are data decisions from the previous stage. An estimate of $a_k(m)$ is then obtained passing the following statistic to a threshold device

$$y_k^{(\ell)}(m) = p^{(\ell)}\mathbf{u}_k^H(m)\mathbf{Z}_k^{(\ell)}(m) - (1 - p^{(\ell)})y_k^{(\ell-1)}(m) \tag{3.13}$$

where $(\cdot)^H$ denotes Hermitian transposition and $p(\ell)$ is a design parameter that determines the fraction of the MAI that is canceled out Divsalar *et al.* (1998)-Peijun & Rappaport (1998). It is less than unity in the early stages (when data decisions are not reliable) and approaches one at the final stage. The conventional *"brute force"* PIC detector which tries to fully cancel the MAI at each stage can be viewed as a special case of the PPIC and it is obtained setting $p(\ell) = 1$ for all ℓ.

As it is known, the performance of the PPIC detector depends heavily on the quality of the initial estimates $\{\hat{a}_k^{(0)}(m)\ ;\ k = 1, 2, \ldots, K\}$. The latter can be obtained using a SUD as initial stage.

3.3 Performance Analysis

Computer simulations have been run to assess the performance of the aforementioned SUDs and PPIC receiver. The system parameters are as follows.

The total number of subcarriers is $N = 64$ and WH codes of length $Q = 8$ are used for spreading purposes. We assume an uncoded transmission in which the information bits are mapped onto QPSK symbols using a Gray map. The number of active users is $K = 8$ (i.e., fully-loaded system) and the wireless channel is

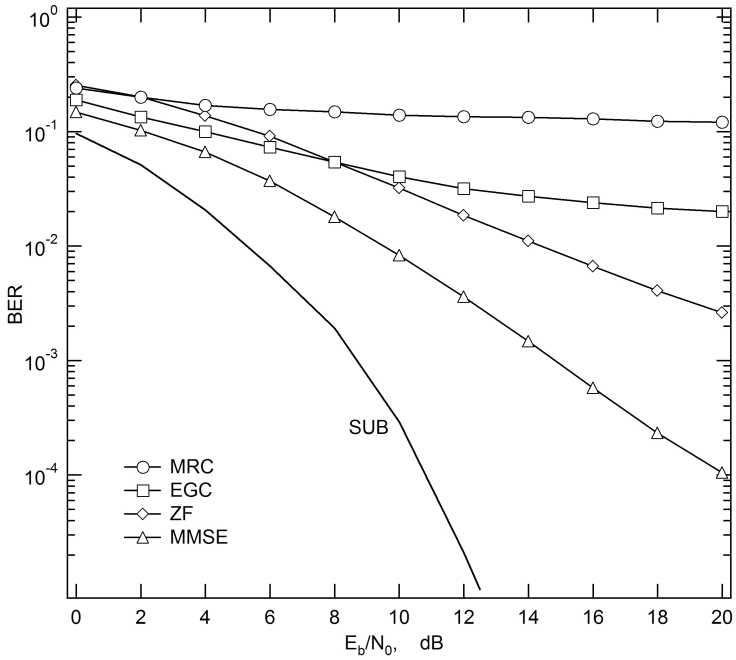

Figure 3.2: BER performance vs. E_b/N_0 for an MC-CDMA system employing different SUDs.

modelled as uncorrelated Rayleigh fading. Perfect channel knowledge is assumed at the receiver end. The performance of the discusses schemes is evaluated in terms of average BER computed over all active users versus E_b/N_0, where E_b is the energy per bit and $N_0/2$ is the two-sided noise spectral density at the receiver side.

Figure 3.2 compares the BER of the aforementioned SUDs. The SUB is also shown for comparison. The latter corresponds to the performance of the MRC detector when there is only one user active in the network (recall that in this case the main impairment at the receive side is the thermal noise and MRC becomes the best detection strategy). The results of Figure 3.2 show that the MMSE detector outperforms the other schemes. At $BER = 10^{-2}$, it achieves a gain of approximately 6 dB with respect to the ZF while the loss compared to

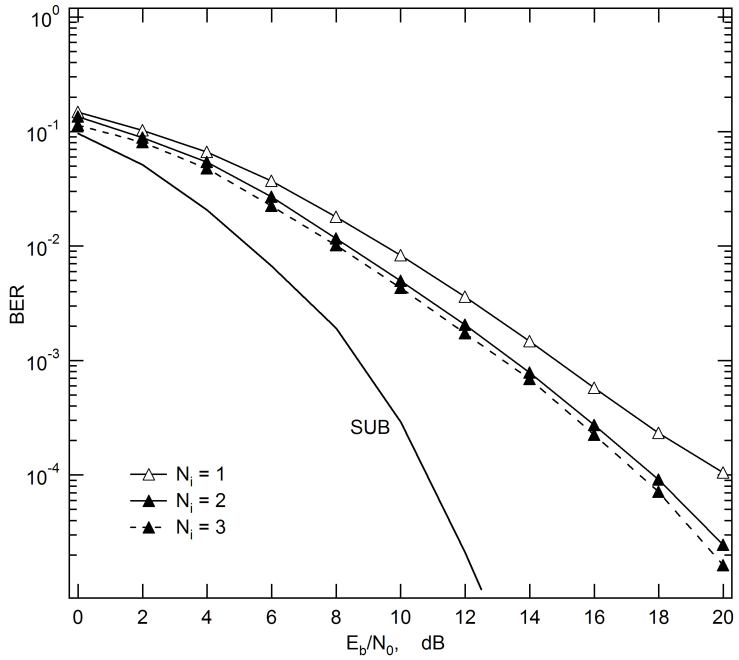

Figure 3.3: BER performance vs. E_b/N_0 for an MC-CDMA system employing a PPIC detector

the SUD is about 3 dB. As already mentioned before, this comes at price of an increased complexity with respect to the other schemes due to the fact that the MMSE detector requires knowledge of the noise power. As expected, the MRC strategy performs poorly since in a fully-loaded system it strongly enhances the MAI. Moreover, we see that, although the ZF approach leads to the complete elimination of MAI, the noise enhancement, which occurs in the presence of deeply faded subcarriers, has a significant impact on the system performance. Finally, we observe that employing the EGC strategy to keep the complexity of the remote unit as low as possible results into a high-error floor.

Figure 3.3 investigates the performance of the PPIC detector for a various number of iterations N_i. The SUB is again shown for comparison. As it occurs with all IC-based schemes, the PPIC needs an initial estimate of the transmitted

data. The latter is taken as the output of a MMSE single-user detector. We see that increasing N_i from 1 to 2 improves the system performance of approximately 2 dB while marginal improvements are achieved passing from 2 to 3.

Figure 3.4 compares the BER of MMSE-SUD, PIC and PPIC. The number of iterations is fixed to 3. Again a MMSE single-user detector is employed as initial stage for both IC-based schemes. The results show that the best performance are achieved with the PPIC receiver. In particular, at a $BER = 10^{-3}$ it achieves a gain of approximately 2 dB compared with the MMSE-SUD while it reduces to 0.5 dB with respect to the PIC receiver. However, it is worth noting that this gain comes for free since PPIC and PIC have comparable complexity.

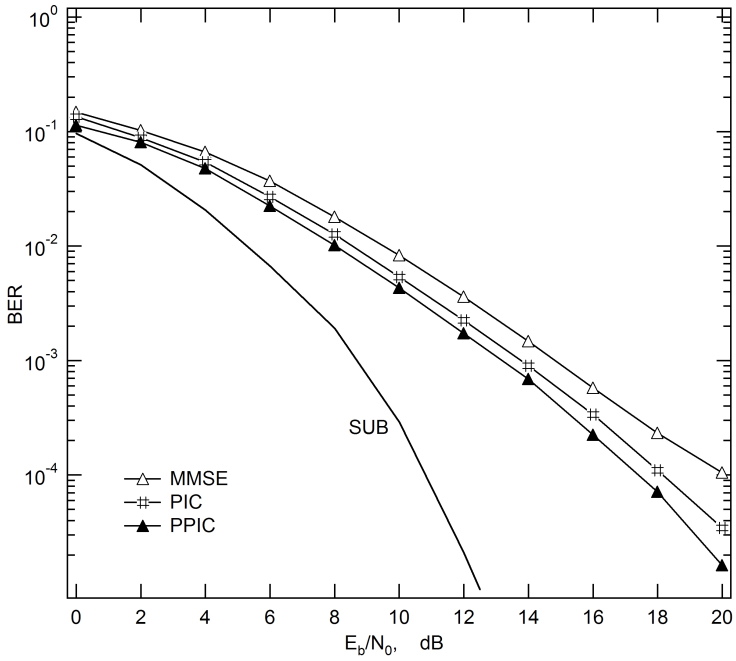

Figure 3.4: Comparison between IC-based schemes and MMSE single user detector.

Chapter 4

Linear Pre-Filtering

As an alternative to SUD strategies or IC-based techniques, pre-filtering can be adopted to cope with multipath propagation and MAI. The idea behind pre-filtering is to vary the gain assigned to each subcarrier so that interference is reduced and the signal at the receiver appears undistorted. In this way, channel estimation is not necessary at the MTs and low-complexity detection schemes may be employed. In order to work properly, however, pre-filtering techniques require CSI at the transmitter side. This assumption is reasonable in TDD systems (such as IEEE802.11a and HiperLAN/II) due to the channel reciprocity between alternative uplink and downlink transmissions. If the analog transmit/receive front ends are correctly calibrated Keusgen *et al.* (2001) and channel variations are sufficiently slow (as occurs in indoor applications), the channel estimate obtained at the BS during the uplink transmission can be reused for transmit filtering in the subsequent downlink time-slot.

So far the topic of signal pre-filtering in MC-CDMA downlink transmissions has received little attention Silva & Gameiro (2003)-Sälzer & Mottier (2003). In particular, the method discussed in Silva & Gameiro (2003) is based on the ZF criterion and selects the pre-filtering coefficients so as to completely eliminate

the MAI at all MTs. The resulting scheme dispenses with channel knowledge at the receiver and has good performance when multiple transmit antennas are employed. The method discussed in Sälzer & Mottier (2003) aims at maximizing the SINR at the MTs. Unfortunately, this leads to a complicated joint optimization problem for the transmit filters of all active users. To overcome this drawback, the authors propose a suboptimum approach based on a modified SINR. In Irmer *et al.* (2003) the pre-filtering coefficients are designed so as to minimize the average BER at all receivers in a CDMA system. This approach is expected to outperform the previous schemes but leads to a complicated non-linear optimization problem that cannot be solved in closed form.

In the next we discuss pre-filtering techniques for the downlink of TDD MC-CDMA systems equipped with multiple transmit antennas. In designing the pre-filtering coefficients we follow three different approaches. The first aims at minimizing the sum of the MSEs at all MTs and turns out to be the more efficient although it requires knowledge of the noise power at the BS. To overcome this problem, we propose a second pre-filtering technique which is based on the ZF criterion. The latter dispenses knowledge of the noise power at the BS but performs remarkably worse with respect to the first in case of a single transmit antenna since it suffers of the problem of noise enhancement. Finally, the third pre-filtering scheme looks for the minimum of an alternative cost function that depends on the SINRs of all active users. As we will see, this problem turns out to be too complex and has no closed-form solutions. To solve this problem, we follow a suboptimal procedure in which a ZF approach is employed first to completely eliminate the MAI and the result is then exploited to minimize the cost function in closed form.

In doing so we impose a constraint on the *overall* transmit power allocated to all active users. Note that this is in contrast to the methods in Silva & Gameiro

(2003)-Sälzer & Mottier (2003), where *separate* power constraints are posed for each active user. Simulations indicate that in this way significant performance improvements can be achieved with respect to existing pre-filtering schemes.

All the above schemes mitigate MAI and channel distortions and allows the MTs to employ simple and power efficient SUD techniques. In this way, most of the computational burden is moved to the BS, where power consumption and complexity are not issues of concern. It is worth noting that, in contrast to W.B. Jang & Pickholtz (1998), we do not attempt to perform joint transmit-receive optimization since, although powerful, this approach would lead to a complicated iterative procedure that cannot be implemented in practical systems. Indeed, the main goal is to derive the pre-filtering coefficients assuming that a given SUD scheme is employed at the MT, so as to improve the reliability of conventional (i.e., without signal pre-filtering) MC-CDMA downlink transmissions without increasing the complexity at the remote device.

4.1 System model

4.1.1 Transmitter structure

We consider the downlink of an MC-CDMA network operating in a TDD mode. The BS is equipped with N_T antennas and the total number of subcarriers, N, is divided into smaller groups of Q elements Fazel & Kaiser (2003). By exploiting the subcarriers of a given group, the BS simultaneously communicate with K active users ($K \leq Q$), which are separated by their specific spreading codes (typically chosen from an orthogonal set). Without loss of generality, we concentrate on a single group and assume that the Q subcarriers are uniformly spread over the signal bandwidth so as to better exploit the channel frequency diversity. We denote $\{i_n\,;\ 1 \leq n \leq Q\}$ the subcarrier indexes in the considered group and a_k

the symbol transmitted to the kth user. For convenience, we collect the users' data into a K-dimensional vector $\mathbf{a} = [a_1, a_2, ..., a_K]^T$.

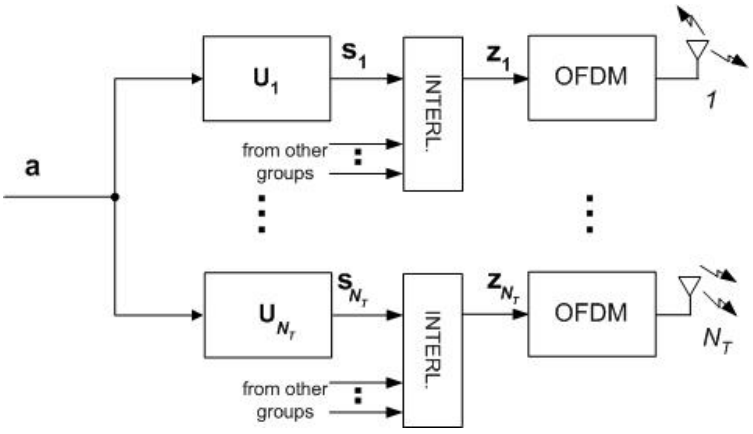

Figure 4.1: Block diagram of an MC-CDMA transmitter employing linear pre-filtering.

Figure 4.1 shows the transmit side of the system under investigation. Each a_k $(k = 1, 2, \ldots, K)$ is spread over Q chips using a unit-energy WH code $\mathbf{c}_k = [c_k(1),\, c_k(2), \ldots,\, c_k(Q)]^T$ and passed to N_T parallel pre-filtering units, one for each antenna branch. We denote $a_k\mathbf{u}_{k,i}$ the pre-filtered chips of the kth user over the ith branch, where $\mathbf{u}_{k,i} = [u_{k,i}(1),\, u_{k,i}(2), \ldots,\, u_{k,i}(Q)]^T$ is a vector with entries

$$u_{k,i}(n) = c_k(n)v_{k,i}(n) \tag{4.1}$$

and $v_{k,i}(n)$ is the pre-filtering coefficient over the nth subcarrier. The contributions of all users are then summed to form the following multi-user signal at the ith antenna

$$\mathbf{s}_i = \sum_{k=1}^{K} a_k\mathbf{u}_{k,i} \qquad i = 1, 2, \ldots, N_T \tag{4.2}$$

or, equivalently,

$$\mathbf{s}_i = \mathbf{U}_i \mathbf{a} \qquad i = 1, 2, \ldots, N_T \tag{4.3}$$

where $\mathbf{U}_i = [\mathbf{u}_{1,i} \;\; \mathbf{u}_{2,i} \cdots \mathbf{u}_{K,i}]$ is the ith forward matrix of Fig. 4.1. In the following we aim directly at the design of $\mathbf{u}_{k,i}$ rather than $\mathbf{v}_{k,i}$.

Finally, each $\mathbf{s}_i$ is frequency interleaved with the contributions from the other groups and mapped on the corresponding subcarriers using an OFDM modulator, which comprises an IDFT unit and the insertion of a CP to avoid interference between adjacent blocks.

4.1.2 Receiver structure

The block diagram of an MC-CDMA receiver is sketched in Figure 3.1. The N_T incoming waveforms are implicitly recombined by the single receive antenna and passed to an OFDM demodulator, which eliminates the CP and performs a DFT operation. Without loss of generality, we concentrate on the mth MT. Neglecting timing and frequency synchronization errors, the output $\mathbf{X}_m = [X_m(1),\, X_m(2), \ldots, X_m(Q)]^T$ from the DFT unit takes the form

$$\mathbf{X}_m = \sum_{i=1}^{N_T} \sum_{k=1}^{K} \mathbf{H}_{m,i} \mathbf{u}_{k,i} a_k + \mathbf{n}_m \tag{4.4}$$

where $\mathbf{n}_m = [n_m(1),\, n_m(2), \ldots, n_m(Q)]^T$ represents thermal noise and it is modelled as a Gaussian vector with zero-mean and covariance matrix $\sigma^2 \mathbf{I}_Q$ while $\mathbf{H}_{m,i}$ is a diagonal matrix

$$\mathbf{H}_{m,i} = \mathrm{diag}\{H_{m,i}(i_1),\, H_{m,i}(i_2), ..., H_{m,i}(i_Q)\} \tag{4.5}$$

collecting the channel frequency response between the ith transmit antenna and the mth MT over the Q subcarriers.

To proceed further, $\mathbf{X}_m$ may be rewritten as

$$\mathbf{X}_m = \mathbf{H}_m \mathbf{U}\mathbf{a} + \mathbf{n}_m \tag{4.6}$$

where $\mathbf{H}_m = [\mathbf{H}_{m,1}\ \mathbf{H}_{m,2} \cdots \mathbf{H}_{m,N_T}]$ is an $Q \times QN_T$ channel matrix. To keep the complexity of the MT at a reasonable level, the decision statistic for a_m is obtained by feeding $\mathbf{X}_m$ to a linear single user detector. This produces

$$y_m = e \cdot \mathbf{q}_m^{\mathrm{H}} \mathbf{X}_m \tag{4.7}$$

where $\mathbf{q}_m = [q_m(1), q_m(2), ..., q_m(Q)]^T$ is a *unit-norm* vector that performs both channel equalization and signal despreading. The scalar e in (4.7) can be thought of as being part of an automatic gain control which does not impair the signal-to-noise ratio at the receiver.

Substituting (4.6) into (4.7) yields

$$y_m = e \cdot \mathbf{g}_m^H \mathbf{U}\mathbf{a} + e \cdot w_m \tag{4.8}$$

where $\mathbf{g}_m = \mathbf{H}_m^H \mathbf{q}_m$ is an QN_T-dimensional vector that depends on the channel coefficients and data detection strategy while $w_m = \mathbf{q}_m^H \mathbf{n}_m$ is a Gaussian random variable with zero-mean and variance σ^2. As mentioned earlier, $\mathbf{q}_m$ performs both channel equalization and despreading and it is designed according to conventional SUD techniques, such as pure despreading, maximum ration combining and equal gain combining. This is tantamount to setting

1) Pure despreading

$$q_m(n) = c_m(n) \quad n = 1, 2, ..., Q \tag{4.9}$$

2) Maximum Ratio Combining

$$q_m(n) = c_m(n) \cdot \frac{\sum_{i=1}^{N_T} H_{m,i}(i_n)}{\sqrt{\frac{1}{Q}\sum_{\ell=1}^{Q}\left|\sum_{i=1}^{N_T} H_{m,i}(i_\ell)\right|^2}} \quad n = 1, 2, ..., Q \tag{4.10}$$

3) Equal Gain Combining

$$q_m(n) = c_m(n) \cdot \frac{\sum_{i=1}^{N_T} H_{m,i}(i_n)}{\left|\sum_{i=1}^{N_T} H_{m,i}(i_n)\right|} \quad n = 1, 2, ..., Q. \tag{4.11}$$

As seen from the above equations and discussed in Chapter 3, only pure despreading strategy dispenses from channel knowledge at the receiver, thereby keeping the complexity of the MT at a very low level. Both other schemes require explicit knowledge of the channel parameters.

As seen from (4.4), in downlink transmissions all the signals arriving from the BS to a given MT propagate through the same channel. This facilitates the channel estimation task, which may be accomplished with the same methods employed for OFDM applications, i.e., using known symbols inserted in both the frequency and time dimensions Kaiser & Hoeher (1997) or employing a channel-sounding approach in which a train of pulses is periodically transmitted Cacopardi *et al.* (1997). In the sequel, perfect channel knowledge is assumed at the MTs.

Stacking the decision statistics of all users into a single vector $\mathbf{y} = [y_1, y_2, ..., y_K]^T$, we get

$$\mathbf{y} = e \cdot \mathbf{G}^H \mathbf{U}\mathbf{a} + e \cdot \mathbf{w} \tag{4.12}$$

where $\mathbf{w} = [w_1, w_2, ..., w_K]^T$ is a Gaussian vector with zero-mean and covariance matrix $\sigma^2 \mathbf{I}_Q$ whereas $\mathbf{Q}$ is the following matrix with dimensions $N_T Q \times K$

$$\mathbf{G} = \begin{bmatrix} \mathbf{H}_{1,1}^H \mathbf{q}_1 & \mathbf{H}_{2,1}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,1}^H \mathbf{q}_K \\ \mathbf{H}_{1,2}^H \mathbf{q}_1 & \mathbf{H}_{2,2}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,2}^H \mathbf{q}_K \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{1,N_T}^H \mathbf{q}_1 & \mathbf{H}_{2,N_T}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,N_T}^H \mathbf{q}_K \end{bmatrix}. \tag{4.13}$$

Inspections of (4.12) indicates that $\mathbf{y}$ depends on the pre-filtering sequences through $\mathbf{U}$. In the next we design $\mathbf{U}$ according to several criteria.

4.2 Linear pre-filtering techniques

4.2.1 MMSE design

The optimality criterion employed for the design of $\mathbf{U}$ is based on the minimization of the following mean square error

$$J = E\left\{ \ \|\mathbf{y} - \mathbf{a}\|^2 \right\} \tag{4.14}$$

where the statistical expectation must be computed with respect to data symbols and thermal noise. In doing so, we impose a constraint on the *overall* transmit power allocated to *all* active users. This means that $\mathbf{U}$ must fulfill the following equation

$$\mathrm{tr}\left\{\mathbf{U}^{\mathrm{H}}\mathbf{U}\right\} = \mathrm{K}. \tag{4.15}$$

Note that this is in contrast to the methods discussed in Silva & Gameiro (2003)-Sälzer & Mottier (2003) where *separate* constraints are posed for *all* active users, i.e., $\|\mathbf{u}_k\|^2 = 1$ for $k = 1, 2, \ldots, K$. As is intuitively clear, the constraint (4.15) ensures more degrees of freedom and allows the BS to give more power to the weakest users so as to jointly perform pre-filtering and power allocation.

4.2.1.1 Constrained MMSE optimization

As shown in Appendix A, in minimizing the cost function (4.14) under the power constraint (4.15) we can employ different approaches. For example, we can follow the line of reasoning employed in Choi & Murch (2004) or, alternatively, the more conventional technique presented in Kusume *et al.* (2005). Since in both cases we arrive at the same final result, in the sequel we adopt the approach of Choi & Murch (2004) which is mathematically simpler. We begin by substituting (4.12) into (4.14) and computing the statistical expectation over $\mathbf{a}$ and $\mathbf{w}$. Assuming that the transmitted symbols belong to a PSK constellation (i.e., $|a_k|^2 = 1$) and bearing in mind that $\mathbf{a}$ and $\mathbf{w}$ are statistically independent, it follows that

$$J = \mathrm{tr}\left\{\left(\mathbf{G}^{\mathrm{H}}\mathbf{V} - \mathbf{I}_{\mathrm{K}}\right)\left(\mathbf{G}^{\mathrm{H}}\mathbf{V} - \mathbf{I}_{\mathrm{K}}\right)^{\mathrm{H}}\right\} + \mathrm{e}^2\sigma^2\mathrm{K} \tag{4.16}$$

where we have defined $\mathbf{V} = e \cdot \mathbf{U}$. In these circumstances the power constraint (4.15) can also be rewritten as

$$\mathrm{tr}\left\{\mathbf{V}^{\mathrm{H}}\mathbf{V}\right\} = \mathrm{e}^2 \cdot \mathrm{K} \tag{4.17}$$

so that (4.16) reduces to

$$J = \mathrm{tr}\{(\,\mathbf{G}^{\mathrm{H}}\mathbf{V} - \mathbf{I}_{\mathrm{K}})(\,\mathbf{G}^{\mathrm{H}}\mathbf{V} - \mathbf{I}_{\mathrm{K}})^{\mathrm{H}} + \sigma^2\mathbf{V}^{\mathrm{H}}\mathbf{V}\}. \tag{4.18}$$

Our objective consists of determining the matrix $\mathbf{V}$ that minimizes the RHS of (4.18). The solution to this problem is found by setting to zero the derivative of J with respect to $\mathbf{V}$ and reads Lutkepohl (1996)

$$\mathbf{V} = \mathbf{G}(\mathbf{G}^H\mathbf{G} + \sigma^2\mathbf{I}_K)^{-1}. \tag{4.19}$$

Finally, bearing in mind that $\mathbf{U} = (1/e) \cdot \mathbf{V}$, we have

$$\mathbf{U} = \sqrt{\frac{K}{\mathrm{tr}\{\mathbf{V}^{\mathrm{H}}\mathbf{V}\}}} \times \mathbf{V}. \tag{4.20}$$

In the sequel, the matrix $\mathbf{U}$ is referred to as TWF. It is fair to say that TWF is reminiscent of the method proposed by Sälzer and Mottier (S&M) in Sälzer & Mottier (2003). The latter employs the following pre-filtering coefficients

$$\mathbf{u}_k = \gamma_k \cdot \mathbf{G}\left(\mathbf{G}^H\mathbf{G} + \sigma^2\mathbf{I}_K\right)^{-1}\mathbf{d}_k \tag{4.21}$$

where γ_k is chosen so as to meet the set of constraints $\|\mathbf{u}_k\|^2 = 1$ $(k = 1, 2, \ldots, K)$, while $\mathbf{d}_k$ denotes the K-dimensional vector with entries

$$d_k(n) = \begin{cases} 1 & \text{if} \quad n = k \\ 0 & \text{otherwise}\,. \end{cases} \tag{4.22}$$

4.2.1.2 Unconstrained MMSE optimization with power scaling

A possible drawback of TWF is that it requires knowledge of the noise power σ^2 which is not available at the BS. To overcome this problem, in the next we adopt a ZF approach similar to that employed in Brandt-Pearce & Dharap (2000). In particular, the pre-filtering matrix $\mathbf{U}$ is first derived so as to minimize the sum of the mean square errors at all MTs and then it is suitably scaled so as to meet the constraint on the overall transmit power. This means that we must look for the minimum of the following cost function

$$J = \mathrm{tr}\left\{\left(\mathbf{G}^{\mathrm{H}}\mathbf{V} - \mathbf{I}_{\mathrm{K}}\right)\left(\mathbf{G}^{\mathrm{H}}\mathbf{V} - \mathbf{I}_{\mathrm{K}}\right)^{\mathrm{H}}\right\} \tag{4.23}$$

which produces

$$\mathbf{V} = \mathbf{G}\left(\mathbf{G}^H\mathbf{G}\right)^{-1}. \tag{4.24}$$

Substituting (4.24) into (4.12) yields

$$\mathbf{y} = e \cdot \mathbf{a} + \mathbf{w} \tag{4.25}$$

from which it is seen that the MAI has been completely eliminated from the decision statistics. To maintain the same transmit power as in the case where no pre-filtering is performed, the transmitted signals must be scaled or, equivalently, matrix $\mathbf{U}$ in (4.24) must be replaced by

$$\mathbf{U} = \sqrt{\frac{K}{\mathrm{tr}\left\{(\mathbf{G}^{\mathrm{H}}\mathbf{G})^{-1}\right\}}} \cdot \mathbf{G}\left(\mathbf{G}^{H}\mathbf{G}\right)^{-1}. \tag{4.26}$$

The pre-filtering algorithm (4.26) is reminiscent of the TIR scheme discussed in Silva & Gameiro (2003), which reads

$$\mathbf{u}_k = \sqrt{\frac{1}{(\mathbf{G}^{H}\mathbf{G})^{-1}_{k,k}}} \cdot \mathbf{G}\left(\mathbf{G}^{H}\mathbf{G}\right)^{-1}\mathbf{d}_k \tag{4.27}$$

where $\mathbf{d}_k$ denotes the K-dimensional vector with entries

$$d_k(n) = \begin{cases} 1 & \text{if} \quad n = k \\ 0 & \text{otherwise}\,. \end{cases} \tag{4.28}$$

It should be noted that S&M using (4.21) reduces to TIR as σ^2 becomes vanishing small. In the sequel, the algorithm using the pre-filtering matrix $\mathbf{U}$ given in (4.26) is called TZF in the sequel.

Remarks

The following remarks are of interest:

1) Inspection of (4.20) and (4.26) reveals that the crux in the calculations of $\mathbf{U}$ is the inversion of $\mathbf{G}^{H}\mathbf{G}$. The latter has dimensions $K \times K$, independently of the spreading factor and the number of transmit antennas. Also, a single matrix

inversion is needed to compute all the pre-filtering coefficients. This makes the proposed algorithm suitable for practical implementations.

2) From (4.25) we see that the BER is the same for all the users and, for a QPSK constellation, reads

$$BER = Q\left(e\sqrt{\frac{1}{2\sigma^2}}\right) \tag{4.29}$$

where $Q(\alpha) = \left(1/\sqrt{2\pi}\right) \cdot \int_{\alpha}^{\infty} e^{-\left(\beta^2/2\right)} d\beta$. This makes the proposed scheme particularly suited for commercial applications, where fair treatment of the active users is typically recommended.

4.2.1.3 Numerical results

Computer simulations have been run to assess the performance of the proposed schemes. The investigated multi-carrier system is inspired by the HiperLAN/II standardization project Medbo (1998) and employs a total of $N = 64$ subcarriers. The latter are divided into 8 groups, each containing $Q = 8$ elements. To better exploit the channel frequency diversity, the subcarriers belonging to the considered group are uniformly distributed over the signal bandwidth with indexes $i_n = 8(n-1)$ for $n = 1, 2, \ldots, 8$. The signal bandwidth is $B = 20$ MHz, so that the useful part of each block has duration $T = N/B = 3.2$ s. The sampling period is $T_s = T/N = 5 \cdot 10^{-2}$ μs and a cyclic prefix of length $T_G = 0.8$ μs is employed to eliminate interblock interference. This corresponds to an extended block (including the cyclic prefix) of 4 μs. The channel frequency response between the ith transmit antenna and the mth MT is expressed by

$$H_{m,i}(i_n) = \sum_{\ell=0}^{L-1} h_{m,i}(\ell)\, e^{-j2\pi\, i_n \ell/N} \qquad n = 1, 2, \ldots, Q \tag{4.30}$$

where $\mathbf{h}_{m,i} = [h_{m,i}(0),\, h_{m,i}(1), \ldots, h_{m,i}(L-1)]^T$ represents the corresponding the discrete-time channel impulse response (CIR) of length $L = 8$. The entries of $\mathbf{h}_{m,i}$ are modeled as independent Gaussian random variables with zero-mean and power

$$\mathrm{E}\left\{|h_{m,i}(\ell)|^2\right\} = \lambda \cdot \exp(-\ell/4) \qquad \ell = 0, 1, \ldots, 7 \tag{4.31}$$

where λ is chosen such that the average energy of $\mathbf{h}_{m,i}$ is normalized to unity, i.e., $\mathrm{E}\left\{|\mathbf{h}_{m,i}|^2\right\} = 1$. The CIRs are kept fixed over the downlink time-slot (slow fading) but vary independently from slot to slot. The transmit antennas are adequately separated so as to make vectors $\mathbf{h}_{m,i}$ statistically independent for different values of i and m.

We assume an uncoded transmission in which the information bits are mapped onto QPSK symbols using a Gray map. The number of active users in the considered group of subcarriers is $K = 8$ (fully-loaded system) and perfect channel knowledge is assumed at both the transmit and receive ends.

The performance of the proposed pre-filtering schemes is evaluated in terms of average BER computed *over all active users* versus E_T/N_0, where E_T is the *transmitted* energy per bit and $N_0/2$ is the two-sided noise spectral density. For a fair comparison with a variable number of transmit antennas, E_T/N_0 has been scaled (in dB) by $10\log(N_T)$ so as to cancel out the corresponding array gain.

Figure 4.2 illustrates the BER of TWF and TZF in conjunction with a PD technique. The curves are labelled TWF+PD and TZF+PD, respectively. The number of transmit antennas is $N_T = 1$, 2 or 4 and the performance over the AWGN channel is also shown as a benchmark. Comparisons are made with the TIR and S&M schemes. As expected, the system performance improves with N_T. We see that TWF+PD outperforms the other schemes for $N_T = 1$. As the number of transmit antennas increases TWF+PD and TZF+PD have virtually

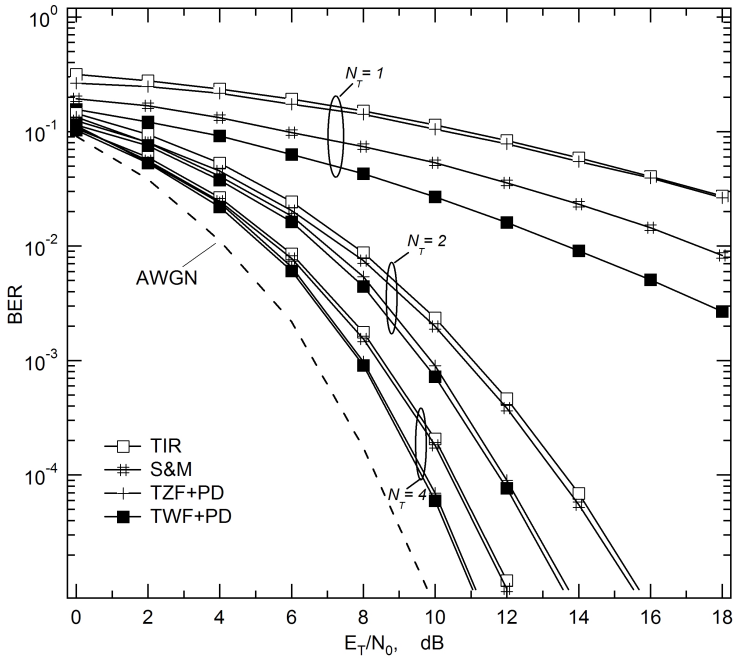

Figure 4.2: BER vs. E_T/N_0 with $K = 8$ for different linear pre-filtering techniques and $N_T = 1$, 2 or 4 antennas.

the same performance and outperform both TIR and S&M.

Figure 4.3 shows the BER of the previous schemes for $N_T = 1$ and 4 in the presence of power imbalance. In particular, we consider a system in which the averages energies of the channel responses are 0 dB for users #1,2,3 and 4, +3 dB for users #5, 6 and, finally, -3 dB for the last two users. Again, TWF+PD and TZF+PD perform remarkably better than the other schemes and achieves a gain of approximately 2 dB at an error rate of 10^{-3}. The reason is that both schemes re-distribute the transmission power among available users, thereby allowing the BS to jointly perform signal pre-filtering and power allocation. Vice versa, both TIR and S&M allocate the same power to all users and, thus, results a very poor performance for the weak users.

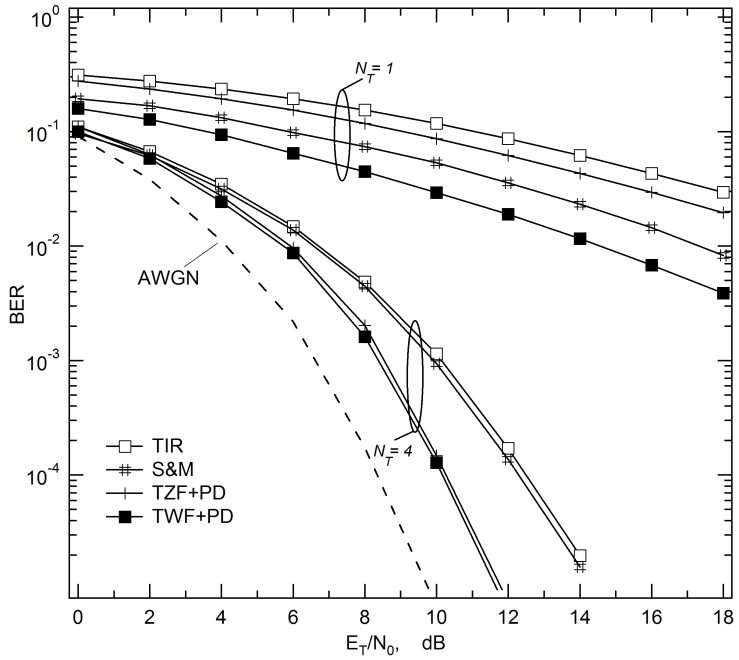

Figure 4.3: BER vs. E_T/N_0 with $K = 8$ for different linear pre-filtering techniques and $N_T = 1$ or 4 in the presence of power imbalance.

Figure 4.4 shows the performance of the proposed schemes in conjunction with different SUD schemes and a single transmit antenna. Since the impact of a given SUD scheme on the system performance is expected to be the same for both TZF and TWF, we concentrate our analysis on TWF. This allow us of reducing the number of curves for each picture, thereby improving the readability of each one. The BER of a conventional MMSE single-user receiver (i.e., without any pre-filtering) is also shown for comparison. We see that TWF+EGC gives the best results. In particular, the gain with respect to the conventional MMSE receiver is approximately 6 dB an error rate of 10^{-2}. As expected, TWF+PD performs poorly compared to the other schemes as it does not exploit CSI at the receiver side.

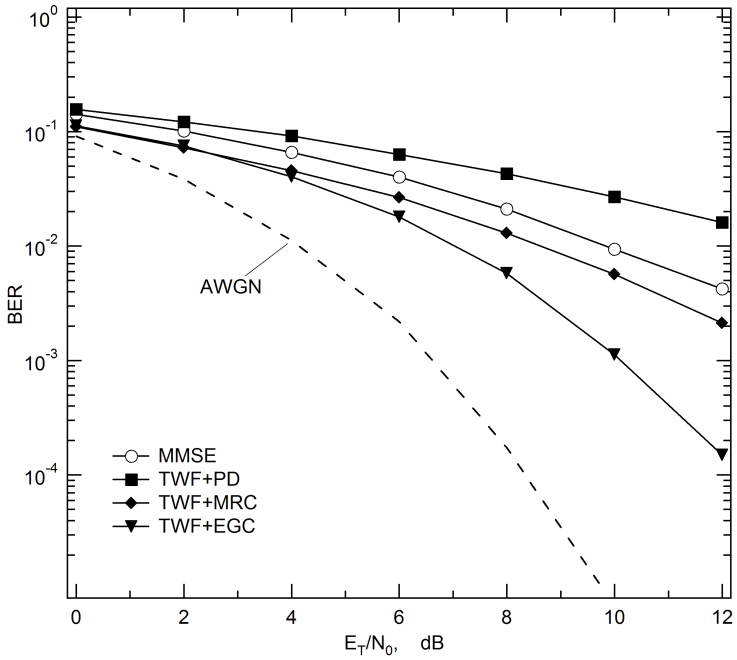

Figure 4.4: BER vs. E_T/N_0 for the TWF scheme with $K = 8$, $N_T = 1$ and different SUD schemes.

The results of Figure 4.5 have been obtained in the same operating conditions of Figure 4.4, expect that two transmit antennas are now employed. The curve labelled STBC+MMSE refers to the transmission technique proposed by W .Sun *et al.* in Sun *et al.* (2003), which employs the Alamouti space-time coding scheme at the transmitter and an MMSE-based detector at the receiver. Note that this technique does not require any CSI at the transmitter and, therefore, it cannot provide any array gain. For this reason, the corresponding curve has not been scaled by $10\log(N_T)$. In contrast to the results of Figure 4.6, we see that now TWF+MRC has the best performance and achieves a gain of approximately 1.5 dB with respect to TWF+PD at an error rate of 10^{-3}. Also, we see that TWF outperforms STBC+MMSE for E_T/N_0 >6 dB, irrespective to the detection strategy employed at the receiver.

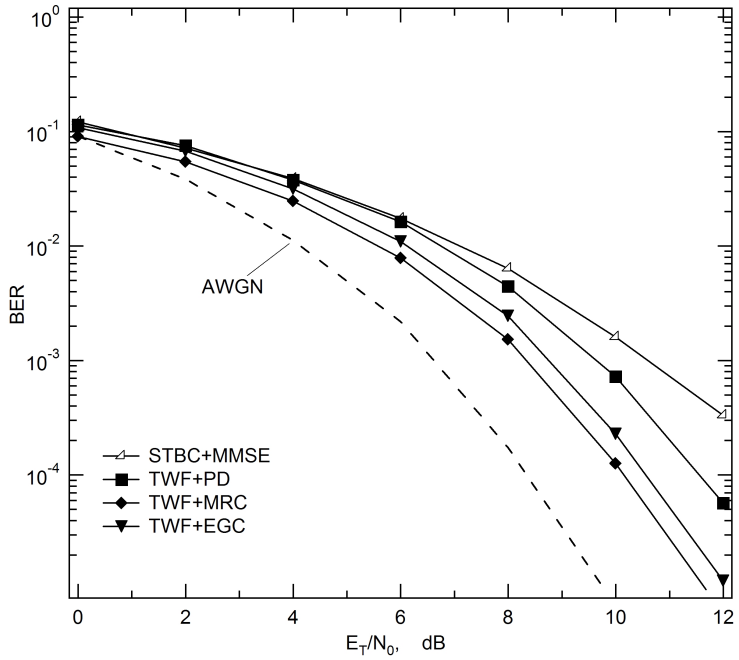

Figure 4.5: BER vs. E_T/N_0 with $K = 8$ for the TWF scheme using $N_T = 2$ and different SUD schemes.

The impact of channel variations on the system performance is addressed in Figure 4.6, where the BER is shown as a function of the mobile speed v for $E_T/N_0 = 10$ dB and $N_T = 2$. For all the considered schemes, the matrix $\mathbf{U}$ is computed at the beginning of each downlink slot and is kept fixed over the entire slot while the channel varies continuously due to Doppler effects. The slot has duration 1.0 ms and consists of 300 blocks. The receiver has ideal CSI at each time instant (i.e., we assume that the channel variations are tracked perfectly at the receiver). As expected, the BER degrades as v increases but the performance loss is negligible for mobile speeds up to 5 m/s. Hence, it can be concluded that the considered schemes are well suited for indoor applications.

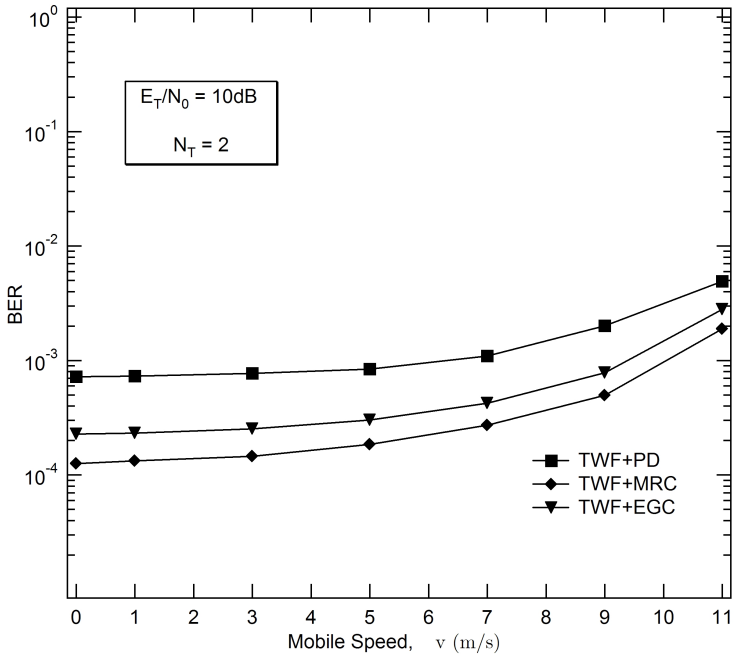

Figure 4.6: BER vs. the mobile speed v with $K = 8$ for the TWF scheme using $N_T = 2$ antennas and different SUD schemes.

4.2.2 Minimum SINR-based design

In the next, we propose an alternative pre-coding technique where the pre-filtering coefficients are designed so as to minimize a proper cost function that depends on the SINRs of all active users. Since the problem turns out to be too complex, we propose a suboptimal procedure in which the ZF approach is employed first to completely eliminate the MAI and the result is then exploited to minimize a proper cost function in closed form.

Letting $e = 1$ and substituting (4.4) into (4.7) yields

$$y_m = \sum_{i=1}^{N_T} \sum_{k=1}^{K} a_m \mathbf{g}_{m,p}^H \mathbf{u}_{k,i} + \mathbf{q}_m^H \mathbf{n}_m. \tag{4.32}$$

At this stage it is useful to separate the contribution of the mth user from that

of the others. This produces

$$y_m = \underbrace{a_m(\mathbf{g}_m^H \mathbf{u}_m)}_{\substack{\text{desired}\\ \text{signal}}} + \underbrace{\sum_{k=1,\, k\neq m}^{K} a_k(\mathbf{g}_m^H \mathbf{u}_k)}_{\text{MAI}} + \underbrace{\mathbf{q}_m^H \mathbf{n}_m}_{\substack{\text{thermal}\\ \text{noise}}}. \tag{4.33}$$

The above equation indicates that both the desired signal and the MAI depend on the pre-equalizer coefficients which must be designed so as to mitigate the MAI while enhancing the desired signal component. A common optimality criterion employed for the design of the pre-filtered spreading sequences $\mathbf{u}_m$ is based on the SINR at the MTs. Assuming statistically independent data symbols with zero-mean and unit variance (i.e., $|a_k|^2 = 1$) from (4.33) we see that the SINR at the mth MT is given by

$$SINR_m = \frac{\left|\mathbf{g}_m^H \mathbf{u}_m\right|^2}{\sigma^2 + \sum\limits_{k=1,\, k\neq m}^{K} \left|\mathbf{g}_m^H \mathbf{u}_k\right|^2} \tag{4.34}$$

where we have used the fact that $\mathbf{q}_m$ is an unit norm vector, i.e., $\mathbf{q}_m^H \mathbf{q}_m = 1$. Intuitively, we would like to maximize $SINR_m$ for $m = 1, 2, \ldots, K$. Therefore, a good optimality criterion for the design of $\{\mathbf{u}_k\,;\, k = 1, 2, \ldots, K\}$ seems to be the maximization of the following sum of SINRs

$$J_1 = \sum_{k=1}^{K} SINR_k. \tag{4.35}$$

As shown later, however, maximization of J_1 does not lead to any useful pre-filtering algorithm. As an alternative, we propose to select $\{\mathbf{u}_k\}$ so as to minimize the modified cost function

$$J_2 = \sum_{k=1}^{K} (SINR_k)^{-1} \tag{4.36}$$

under the power constraint (4.15). Collecting (4.34) and (4.36), we see that direct minimization of is analytically infeasible as it leads to a joint optimization problem for the pre-filtered sequences of all active users. As mentioned before, in order to overcome this problem, we follow a suboptimal procedure in which MAI is eliminated first and the result is then exploited to minimize J_2 in closed form.

MAI elimination

Inspection of (4.33) reveals that $a_k(\mathbf{g}_m^H \mathbf{u}_k)$ is the interference that the signal of a given user k produces at the mth MT. Therefore, the complete elimination of the MAI induced by the kth user at all other MTs leads to the following set of equations

$$\begin{cases} \mathbf{g}_k^H \mathbf{u}_m = 0 & 1 \leq k \leq K \quad \text{and} \quad k \neq m \\ \mathbf{g}_m^H \mathbf{u}_m = \lambda_m \end{cases} \tag{4.37}$$

where λ_m is a positive real parameter that must be chosen so as to meet the power constraint in (4.15). To proceed, we see that $\mathbf{u}_m$ can also be written as

$$\mathbf{u}_m = \lambda_m \mathbf{u'}_m \tag{4.38}$$

where $\mathbf{u'}_m$ satisfies the following system of K linear equations and $N_T Q$ unknowns

$$\mathbf{G}^H \mathbf{u'}_m = \mathbf{d}_m \tag{4.39}$$

where $\mathbf{d}_m$ is a K-dimensional vector with entries

$$d_m(n) = \begin{cases} 1 & \text{if} \quad n = m \\ 0 & \text{otherwise}\,. \end{cases} \tag{4.40}$$

Since $K \leq Q$, the above system may be underdetermined (more unknowns than equations). In such a case there exists an infinite number of sequences $\mathbf{u'}_m$ satisfying (6.25). The problem of finding the best $\mathbf{u'}_m$ is addressed later.

Design of the pre-filtering coefficients

From (4.38), we see that the power constraint (4.15) may be rewritten as

$$\sum_{k=1}^{K} \lambda_k^2 \left\| \mathbf{u'}_k \right\|^2 = K \tag{4.41}$$

Also, collecting (4.34), (4.36) and (4.37) produces

$$J_2 = \sigma^2 \sum_{k=1}^{K} \frac{1}{\lambda_k^2}. \tag{4.42}$$

We proceed by computing the coefficients $\{\lambda_k \,;\, k = 1, 2, \ldots, K\}$ that minimize J_2 under the constraint (4.41). The solution to this problem is found looking for the minimum of the following cost function with Lagrange multiplier μ

$$J = \sigma^2 \sum_{k=1}^{K} \frac{1}{\lambda_k^2} + \mu \left\{ K - \sum_{k=1}^{K} \lambda_k^2 \left\| \mathbf{u'}_k \right\|^2 \right\}. \tag{4.43}$$

Taking the derivative of J with respect to λ_m and setting it to zero produces

$$\lambda_m = \frac{\gamma}{\sqrt{\left\| \mathbf{u'}_m \right\|}} \tag{4.44}$$

where γ is found substituting (4.44) into (4.41) and reads

$$\gamma = \sqrt{\frac{K}{\sum_{k=1}^{K} \left\| \mathbf{u'}_k \right\|}}. \tag{4.45}$$

At this stage we are left with the problem of finding the best sequence $\mathbf{u'}_m$ satisfying the linear system (4.39), where the optimality criterion is still the minimization of J_2 . To this purpose, we substitute (4.44)-(4.45) into (4.42) and obtain

$$J_2 = \frac{\sigma^2}{K} \left[\sum_{k=1}^{K} \left\| \mathbf{u'}_k \right\| \right]^2. \tag{4.46}$$

Clearly, the sequence $\mathbf{u'}_m$ that minimizes J_2 is the minimum-norm solution of (4.39) and reads

$$\mathbf{u'}_m = \mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}\mathbf{d}_m \tag{4.47}$$

From the above expression it follows that $\|\mathbf{u'}_m\| = \sqrt{\rho_{m,m}}$, where we have denoted the -entry of $(\mathbf{G}^H\mathbf{G})^{-1}$. In summary, collecting (4.38) and (4.44)-(4.47), we see that the set of pre-filtered spreading sequences leading to the complete elimination of MAI and minimizing is given by

$$\mathbf{u}_m = \sqrt{\frac{K}{\sum_{k=1}^{K}\sqrt{\rho_{k,k}\cdot\rho_{m,m}}}} \times \mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}\mathbf{d}_m \qquad m = 1, 2, \ldots, K. \tag{4.48}$$

The above equation indicates that proposed pre-filtering algorithm is reminiscent of the TIR scheme discussed in Silva & Gameiro (2003) as given in (4.27). The only difference between (4.27) and (4.48) is in the multiplicative scalar parameter, which is designed in (4.27) so as to meet the set of power constraints $\|\mathbf{u}_k\| = 1$ for $k = 1, 2, \ldots, K$. For this reason, the proposed algorithm is called the modified TIR (m-TIR) in the sequel.

Equations (4.44) and (4.45) provide the coefficients λ_m that minimize J_2 under the power constraint (4.41). At this stage it is interesting to compute the coefficients λ_m that maximize the cost function in (4.35). After some manipulations, it is straightforward to show that

$$\lambda_m = \begin{cases} \sqrt{K/\left\|\mathbf{u'}_\ell\right\|^2} & \text{if } m = \ell \\ 0 & \text{otherwise} \end{cases} \tag{4.49}$$

where

$$\ell = \arg\max_{1\le k\le K}\{\|\mathbf{u'}_k\|\}\,. \tag{4.50}$$

Clearly, the above solution has no practical interest as it allocates the overall transmit power to a single user.

4.2.2.1 Numerical results

We consider an indoor scenario inspired by the TDD HiperLAN/2 standard. The transmitted symbols belong to a QPSK constellation and are obtained from the information bits through a Gray map. The total number of subcarriers is $N = 64$ and WH codes of length $Q = 8$ are used for spreading purposes. The signal bandwidth is $B = 20$ MHz, so that the useful part of each MC-CDMA block has length $T = N/B = 3.2$ μs. A cyclic prefix of $T_G = 0.8$ μs is adopted to eliminate inter-block interference. This corresponds to an extended block (including the cyclic prefix) of 4 μs. The downlink time-slot has duration 1.0 ms and consists of 250 blocks. All the CIRs $\mathbf{h}_{k,i}$ have the same length $L_{k,i} = 8$. The entries of $\mathbf{h}_{m,i}$ have power

$$\mathrm{E}\left\{|h_{m,i}(\ell)|^2\right\} = \exp(-\ell) \qquad \ell = 0, 1, \ldots, 7 \tag{4.51}$$

and are generated by filtering statistically independent white Gaussian processes in a third-order low-pass Butterworth filter. The 3-dB bandwidth of the filter is taken as a measure of the Doppler shift $f_D = f_0 v/c$, where $f_0 = 5$ GHz is the carrier frequency, v denotes the MT velocity and $c = 3 \times 10^8$ m/s is the speed of light.

The performance of the proposed pre-filtering algorithm is computed in terms of averaged bit-error-rate computed over all active users. Comparisons are made with the TIR scheme discussed in Silva & Gameiro (2003) and the S&M method proposed in Sälzer & Mottier (2003).

Figure 4.7 illustrates the BER of the considered pre-filtering techniques E_T/N_0 vs. σ^2, where E_T/N_0 is the transmitted energy per bit and is the two-sided noise

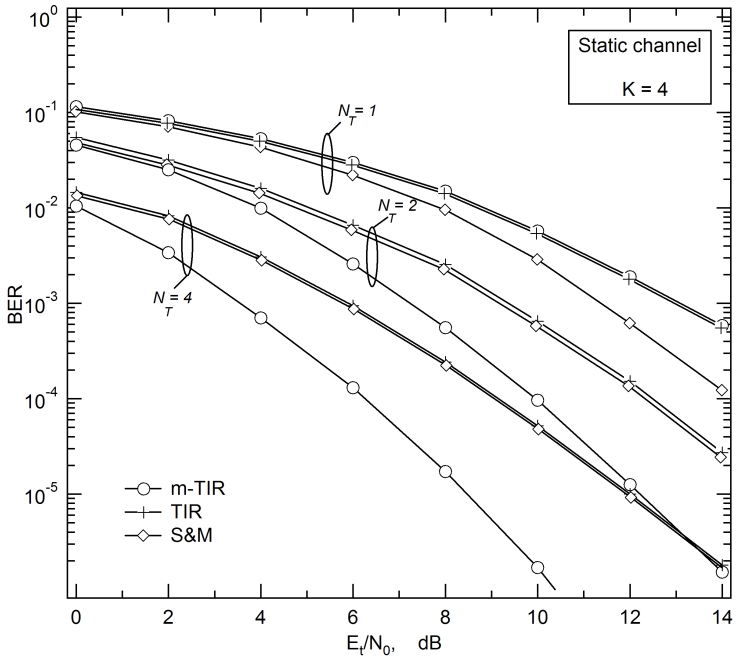

Figure 4.7: BER performance of the m-TIR scheme over a static channel with four active users.

spectral density at the receiver side. There are 4 active users (half-load) and the number of transmit antennas is $N_T = 1$, 2 or 4. The CIRs are kept fixed over the downlink time-slot (static channel) but vary from slot to slot. We see that the m-TIR gives the best results irrespective of the number of antennas. TIR and S&M have virtually the same performance and, for an error probability of 10^{-3}, they loose 2 dB with respect to m-TIR. Note that doubling the number of antennas entails a gain of approximately 4 dB for all schemes.

Figure 4.8 shows results obtained in the same operating conditions of Figure 4.7, except that the system is now full-loaded ($K = 8$). As we can see, the m-TIR outperforms the other schemes for $N_T = 2$ and 4 whereas S&M takes the lead when $N_T = 1$. The loss of the considered schemes with respect to the corresponding curves of Figure 1 is approximately 1.5 dB when $N_T = 2$ and reduces

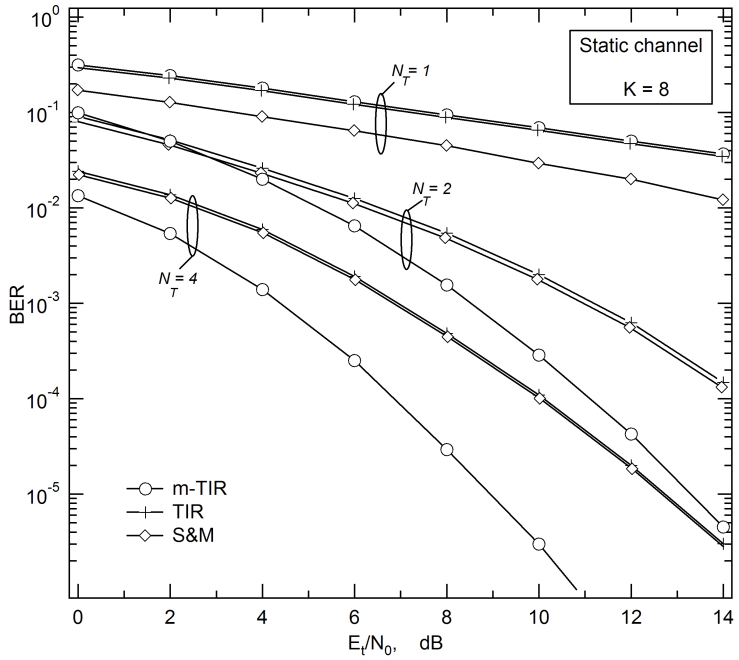

Figure 4.8: BER performance of modified TIR over a static channel with eight active users.

to 1 dB with $N_T = 4$. Much larger degradations occur in the case of a single transmit antenna.

The impact of channel variations on the system performance is addressed in Figure 4.9, where the BER is shown as a function of the mobile speed v for $E_T/N_0 = 10$ dB. The system is full-loaded and the number of antennas is either $N_T = 2$ or 4. For all the considered schemes, the sequences $\mathbf{u}_k$ $(k = 1, 2, \ldots, K)$ are computed at the beginning of each downlink slot and are kept fixed over the slot while the CIRs vary from block to block due to Doppler effects. As expected, the BER degrades as v increases but the performance loss is negligible for mobile speeds up to 3 m/s. This means that the considered schemes are well suited for indoor applications.

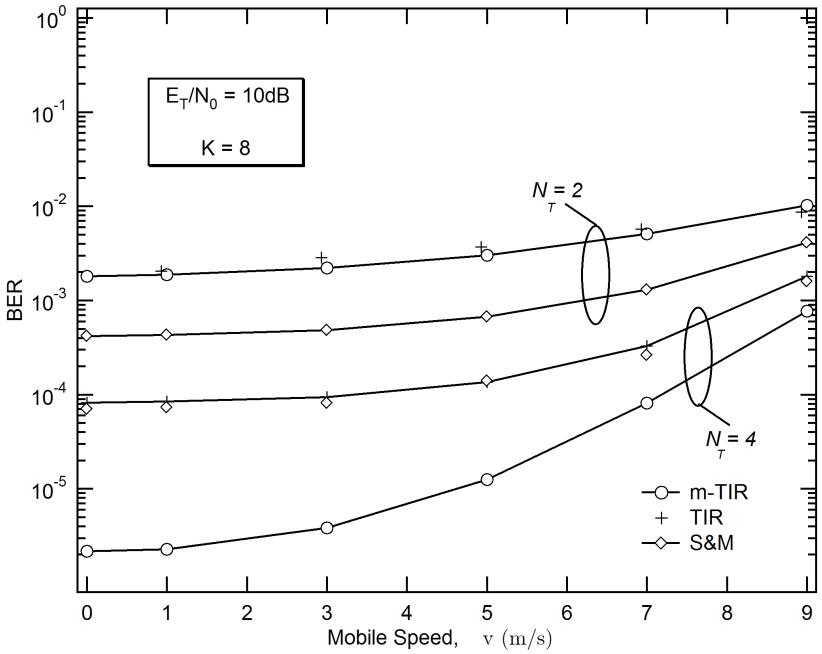

Figure 4.9: BER performance of modified TIR vs. the mobile speed v with $K = 8$ using $N_T = 2$ or 4.

Chapter 5

Non-Linear Pre-Filtering

As seen in previous Chapter, linear pre-filtering based on ZF and/or MMSE criterion is a viable method for mitigating MAI and channel distortions in multiuser downlink transmissions. As shown, the idea behind pre-filtering is to vary the complex gain assigned to each subcarrier so that interference is reduced and the signal at the receiver appears undistorted. In this way, low complex single-user detection schemes can be employed at the MT to detect the transmitted data symbols. The main drawback of linear pre-filtering is represented by the power boosting effect Brandt-Pearce & Dharap (2000). This phenomenon occurs in the presence of deeply faded subcarriers and leads to high power consumption at the transmit side. Scaling the pre-filtering coefficients by a fixed normalization factor is a viable method to reduce the transmit power to a given pre-assigned value. In this way, however, the SNR at the receiver decreases and the system performance is correspondingly degraded.

5.1 Tomlinson-Harashima Pre-coding

An effective solution to mitigate the power boosting effect is represented by non-linear pre-filtering based on THP. This technique employs modulo arithmetic and was originally proposed to combat intersymbol interference in single-user transmissions over highly dispersive channels Tomlinson (1971),H.Harashima & H.Miyakawa (1972). Recently, it has been employed in multi-user systems to counteract the detrimental effect of MAI Windpassinger *et al.* (2004) -Cosovic *et al.* (2005). In these applications, THP can be viewed as the transmit counterpart of the V-BLAST architecture Wolniansky *et al.* (1998). The main difference is that the latter operates at the receive side and can only exploit data decisions for interference cancelation while the former is employed at the transmitter, where the true data symbols are available for signal pre-coding. This means that THP does not suffer from the error propagation phenomena and, accordingly, it is expected to outperform V-BLAST.

So far, non-linear pre-filtering based on THP has been extensively studied for multiple-input multiple-output systems with decentralized receivers (MIMO multi-user) Windpassinger *et al.* (2004)-Kusume *et al.* (2005). In particular, the scheme discussed in Windpassinger *et al.* (2004) is based on a ZF approach and exploits the QR decomposition of the MIMO channel matrix. In Hunger *et al.* (2005), the processing matrices are designed according to an MMSE criterion under a constraint on the overall transmit power. Unfortunately, the solution to this problem requires a large number of matrix inversions (equal to the number of active users) and its application to heavy-loaded systems may be difficult. An efficient implementation of the above algorithm is discussed in Kusume *et al.* (2005), where all matrix inversions are replaced by a single Cholesky factorization. Non-linear pre-filtering for MC-CDMA downlink transmissions has only been investigated in Cosovic *et al.* (2005) under a ZF constraint. Albeit reasonable,

this scheme produces significant differences in the error rate performance of the receive terminals, which may be undesirable in commercial applications requiring a fair treatment of the active users.

In this work we employ a unified framework to investigate non-linear pre-filtering in both MC-CDMA and OFDMA systems equipped with multiple transmit antennas. In designing the processing matrices we aim at minimizing the sum of the MSEs at all MTs under a constraint on the overall transmit power. The resulting scheme mitigates MAI and channel distortions and outperforms the method in Cosovic *et al.* (2005). In addition, it allows the MTs to employ conventional SUD techniques, thereby moving most of the computational burden to the BS, where power consumption and computational resources are not critical issues. It is worth nothing that we do not attempt to perform any joint transmit-receive optimization since, although powerful, this approach would require fully cooperative receivers, which is difficult to achieve in multiuser downlink transmissions W.B. Jang & Pickholtz (1998). To keep the complexity of the MT at a reasonable level, conventional SUD strategies are employed for data detection and the goal is to derive the pre-filtering matrices for a given pre-assigned receive structure.

The use of a unified framework comprising both MC-CDMA and OFDMA allows a comparison between these multiple-access technologies under the same operating conditions. As we shall see, in OFDMA applications the proposed THP-based scheme reduces to MRT over the available transmit antennas T. K. Y. Lo (1999), thereby leading to a significant reduction of complexity as compared to MC-CDMA. Moreover, it is found that OFDMA outperforms MC-CDMA when the system resources are optimally assigned to the active users according to the actual channel realization.

5.2 System model

5.2.1 Transmitter structure

We consider the downlink of a multi-carrier system in which the BS is equipped with N_T antennas and the total number of subcarriers, N, is divided into smaller groups of Q elements. Without loss of generality, we concentrate on a single group and denote $\{i_n\,;\ 1 \leq n \leq Q\}$ the corresponding subcarrier indexes. The BS employs the Q subcarriers to communicate with K active users ($K \leq Q$), which can be separated in various ways. In this work we consider code- and frequency-division-multiplexing, corresponding to MC-CDMA and OFDMA, respectively. The symbol transmitted to the kth user is denoted a_k and belongs to an M-QAM constellation with average energy $\sigma_a^2 = 2(M-1)/3$. This amounts to saying that both the real and imaginary parts of a_k are taken from the set $A = \{\pm 1, \pm 3, ..., \pm\sqrt{M}-1\}$. For convenience, we collect the users' data into a K-dimensional vector $\mathbf{a} = [a_1, a_2, ..., a_K]^T$.

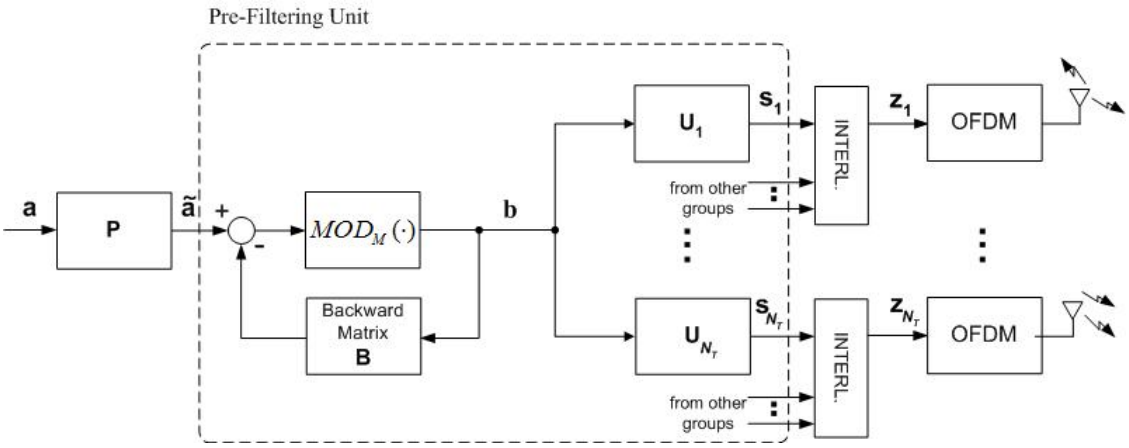

Figure 5.1: Block diagram of transmitter employing non-linear pre-filtering.

Figure 5.1 illustrates the transmit side of the system under investigation. Since

the pre-coding order is expected to have a significant impact on the performance of THP Hunger *et al.* (2005), the input symbols $\mathbf{a}$ are first re-ordered using a permutation matrix

$$\mathbf{P} = \sum_{k=1}^{K} \mathbf{e}_k \mathbf{e}_{\mu_k}^T \tag{5.1}$$

where $\mathbf{e}_k$ is the kth column of $\mathbf{I}_K$ (we denote $\mathbf{I}_K$ the identity matrix of order K) while $\{\mu_1, \mu_2, \ldots, \mu_K\}$ is a suitable permutation of $\{1, 2, \ldots, K\}$ which determines the pre-coding order. The re-ordered vector $\tilde{\mathbf{a}} = \mathbf{Pa}$ is fed to the pre-filtering unit, which consists of a *backward* matrix $\mathbf{B}$ of order K, a non-linear operator $MOD_M(\cdot)$ and N_T *forward* matrices $\mathbf{U}_i$ $(i = 1, 2, \ldots, N_T)$, each with dimensions $Q \times K$. As discussed in Fisher (2002), $\mathbf{B}$ must be strictly lower triangular to allow data pre-coding in a recursive fashion.

To explain the rationale behind the proposed pre-coding structure, we temporarily neglect the non-linear operator in Figure 5.1. In these circumstances, the pre-coded symbols $\mathbf{b} = [b_1, b_2, \ldots, b_K]^T$ are recursively computed from $\tilde{\mathbf{a}}$ as follows

$$b_k = \tilde{a}_k - \sum_{\ell=1}^{k-1} [\mathbf{B}]_{k,\ell}\, b_\ell \tag{5.2}$$

where $[\cdot]_{k,\ell}$ denotes the (k, ℓ)th entry of the enclosed matrix. The above equation may be written in matrix form as

$$\mathbf{b} = \mathbf{C}^{-1}\mathbf{Pa} \tag{5.3}$$

where $\mathbf{C} = \mathbf{B} + \mathbf{I}_K$ is a *unit-diagonal* and lower triangular matrix, i.e., $[\mathbf{C}]_{k,k} = 1$ and $[\mathbf{C}]_{k,\ell} = 0$ for $k < \ell$. Inspection of (5.3) reveals that the energy of the pre-coded symbols depends on $\mathbf{C}$ and it may be very large in the presence of deep fades, thereby leading to a significant increase of the transmit power Wind-

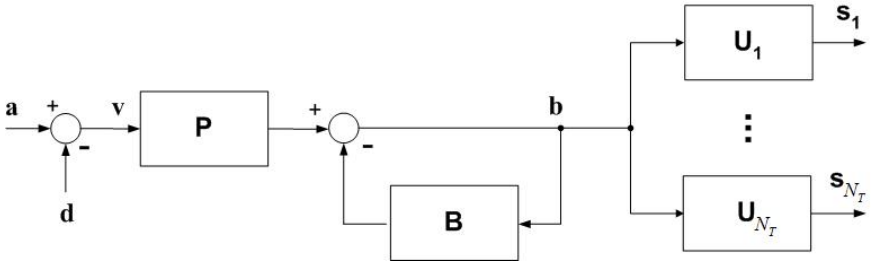

Figure 5.2: Equivalent block diagram of the non-linear pre-filtering unit.

passinger *et al.* (2004). This is a manifestation of the power boosting effect, which occurs in conjunction with linear pre-filtering. To overcome this problem, we adopt a THP approach and introduce a modulo operator that acts independently over the real and imaginary parts of its input according to the following rule

$$MOD_M(x) = x - 2\sqrt{M} \cdot \left\lfloor \frac{x - \sqrt{M}}{2\sqrt{M}} \right\rfloor \tag{5.4}$$

where the notation $\lfloor z \rfloor$ indicates the *smallest* integer larger than or equal to z. In this way, the pre-coded symbols b_k at the output of the non-linear device are constrained into the square region $\aleph = \{x^{(R)} + jx^{(I)} | x^{(R)}, x^{(I)} \in (-\sqrt{M}, \sqrt{M}]\}$ and the transmit power is consequently reduced compared to conventional linear pre-filtering.

Applying the non-linearity (5.4) to the RHS of (5.2), yields

$$b_k = \tilde{a}_k - \sum_{\ell=1}^{k-1} [\mathbf{B}]_{k,\ell}\, b_\ell + \tilde{d}_k \quad k = 1, 2, \ldots, K \tag{5.5}$$

where $\tilde{d}_k = 2\sqrt{M} \cdot \tilde{p}_k$ and $\tilde{p}_k$ is a complex-valued quantity whose real and imaginary components are suitable integers that constraint b_k within the region $\aleph$ (clearly, a unique $\tilde{p}_k$ exists with such a property). The above equation indicates that the modulo operator in Figure 5.1 is equivalent to adding a vector $\mathbf{d} = \mathbf{P}^T \tilde{\mathbf{d}}$

to the input data $\mathbf{a}$, where $\tilde{\mathbf{d}} = \left[\tilde{d}_1, \tilde{d}_2, \ldots, \tilde{d}_K\right]^T$ and we have borne in mind that $\mathbf{P}\mathbf{P}^T = \mathbf{I}_K$. This results into the equivalent block diagram of Figure 5.2, from which it follows that $\mathbf{b} = \mathbf{P}\mathbf{v} - \mathbf{B}\mathbf{b}$ or, equivalently,

$$\mathbf{b} = \mathbf{C}^{-1}\mathbf{P}\mathbf{v} \tag{5.6}$$

where $\mathbf{v} = \mathbf{a} + \mathbf{d}$ is the *effective* data vector. Comparing (5.6) with (5.3), we see that the only difference between linear pre-filtering and THP is that in the latter the true data vector $\mathbf{a}$ is replaced by $\mathbf{v}$. In this way, the transmit power is reduced without any loss of information since $\mathbf{a}$ can be easily regenerated from $\mathbf{v}$ using the modulo device in (5.4).

After non-linear pre-coding, vector $\mathbf{b}$ is passed to N_T matrices $\mathbf{U}_i$ ($i = 1, 2, \ldots, N_T$), one for each antenna branch. This produces the Q-dimensional vectors

$$\mathbf{s}_i = \mathbf{U}_i\mathbf{b} \qquad i = 1, 2, \ldots, N_T \tag{5.7}$$

which can also be written as

$$\mathbf{s}_i = \sum_{k=1}^{K} b_k \mathbf{u}_{k,i} \qquad i = 1, 2, \ldots, N_T \tag{5.8}$$

where $\mathbf{u}_{k,i}$ is the kth column of $\mathbf{U}_i$. Next, each $\mathbf{s}_i$ is frequency interleaved with the contributions of the other groups and the resulting vector $\mathbf{z}_i$ is finally mapped on N subcarriers using an OFDM modulator. Clearly, in OFDMA transmissions only one entry of $\mathbf{u}_{k,i}$ is expected to be non-zero since b_k is transmitted over a single subcarrier. On the other hand, in MC-CDMA $\mathbf{u}_{k,i}$ can be interpreted as the pre-filtered spreading code of the kth user at the ith transmit branch.

Collecting vectors $\mathbf{s}_i$ into a single $N_T Q$-dimensional vector $\mathbf{s} = \left[\mathbf{s}_1^T\, \mathbf{s}_2^T \cdots\, \mathbf{s}_{N_T}^T\right]^T$ and bearing in mind (5.7), it follows that

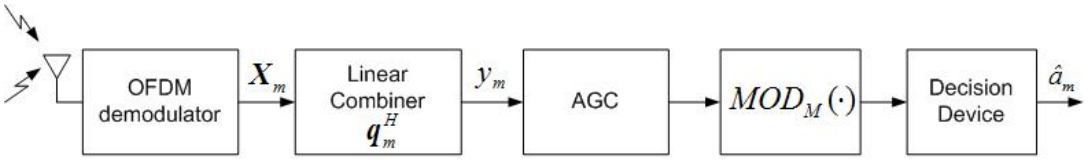

Figure 5.3: Block diagram of the mth receiver in an MC-CDMA network employing non-linear pre-filtering.

$$\mathbf{s} = \mathbf{U}\mathbf{b} \tag{5.9}$$

where $\mathbf{U} = \left[\mathbf{U}_1^H \;\; \mathbf{U}_2^H \cdots \mathbf{U}_{N_T}^H\right]^H$ is a matrix of dimensions $N_T Q \times K$.

5.2.2 Receiver structure

The signals transmitted by the BS array propagate through multipath channels and undergo frequency-selective fading. Without loss of generality, we concentrate on the mth MT and assume that it is equipped with a single-antenna receiver. As shown in Figure 5.3, the incoming waveforms are implicitly combined by the receive antenna and passed to an OFDM demodulator. We denote $\mathbf{X}_m = [X_m(1),\, X_m(2), \ldots, X_m(Q)]^T$ the demodulator outputs corresponding to the Q subcarriers of the considered group. Then, assuming ideal frequency and timing synchronization, we have

$$\mathbf{X}_m = \mathbf{H}_m \mathbf{U}\mathbf{b} + \mathbf{n}_m \tag{5.10}$$

where $\mathbf{H}_m = [\mathbf{H}_{m,1}\mathbf{H}_{m,2}...\mathbf{H}_{m,N_T}]$ is a $Q \times N_T Q$ matrix in which

$$\mathbf{H}_{m,i} = \mathrm{diag}\{H_{m,i}(i_1),\, H_{m,i}(i_2), ..., H_{m,i}(i_Q)\} \tag{5.11}$$

represents the channel frequency response between the ith transmit antenna and the mth MT over the Q subcarriers. Also, $\mathbf{n}_m = [n_m(1),\, n_m(2), \ldots, n_m(Q)]^T$

is thermal noise which is modelled as a Gaussian vector with zero-mean and covariance matrix $\sigma_n^2 \mathbf{I}_Q$.

To keep the complexity of the MT at a reasonable level, data detection is accomplished by means of a simple SUD scheme. For this purpose, the entries of $\mathbf{X}_m$ are linearly combined to form

$$y_m = \sqrt{e} \cdot \mathbf{q}_m^{\mathrm{H}} \mathbf{X}_m \tag{5.12}$$

where $\mathbf{q}_m = [q_m(1), q_m(2), ..., q_m(Q)]^T$ is a unit-norm vector while $e > 0$ is a real parameter that can be thought of as being part of the AGC unit.

Recalling that non-linear pre-coding is equivalent to replacing the true data symbols $\mathbf{a}$ with $\mathbf{v} = \mathbf{a} + \mathbf{d}$, in normal operating conditions we expect that $y_m = v_m + \eta_m$, where $v_m = a_m + d_m$ while η_m is a disturbance term that accounts for residual interference and thermal noise. In order to remove the contribution of d_m, we pass y_m to a non-linear device that operates according to (5.4). The output is finally fed to a threshold unit which delivers an estimate of a_m.

Different selections of $\mathbf{q}_m$ in (5.12) correspond to different SUD schemes and/or multiple-access techniques. For example, in OFDMA systems each sub-carrier in the considered group is exclusively assigned to a different user and the entries of $\mathbf{q}_m$ take the form

$$q_m(n) = \begin{cases} 1 & \text{if } \; n = j_m \\ 0 & \text{otherwise} \end{cases} \tag{5.13}$$

where j_m belongs to the set $\{1, 2, \ldots, Q\}$ and satisfies the constraints $j_m \neq j_k$ for $m \neq k$ since different users must be allocated on different subcarriers. In case of dynamic allocation of the available subcarriers among the active users, the mapping function $m \rightarrow j_m$ $(m = 1, 2, \ldots, K)$ must be properly designed according to the actual channel realizations. Otherwise, we can simply set $j_m = m$

if the system operates in a rigid fashion without performing any dynamic resource allocation. On the other side, in MC-CDMA the users' data are spread over the Q subcarriers using orthogonal WH codes. This is tantamount to setting

$$q_m(n) = c_{j_m}(n)\, t_m(n) \qquad n = 1, 2, \ldots, Q \tag{5.14}$$

where $\mathbf{c}_{j_m} = [c_{j_m}(1),\, c_{j_m}(2), \ldots,\, c_{j_m}(Q)]^T$ is the WH spreading sequence assigned to the mth user, with $c_{j_m}(n) \in \{\pm 1/\sqrt{Q}\}$. Here, the mapping $m \rightarrow j_m$ accounts for possible dynamic allocation of the signature codes while the coefficients $\{t_m(n)\}$ are designed according to the selected SUD strategy. As discussed in Chapter 3, we consider the following well-known SUD techniques.

1) Pure Despreading

$$t_m(n) = 1 \tag{5.15}$$

2) Maximum Ratio Combining

$$t_m(n) = \frac{\sum_{i=1}^{N_T} H_{m,i}(i_n)}{\sqrt{\frac{1}{Q}\sum_{\ell=1}^{Q}\left|\sum_{i=1}^{N_T} H_{m,i}(i_\ell)\right|^2}} \tag{5.16}$$

3) Equal Gain Combining

$$t_m(n) = \frac{\sum_{i=1}^{N_T} H_{m,i}(i_n)}{\left|\sum_{i=1}^{N_T} H_{m,i}(i_n)\right|} \tag{5.17}$$

Substituting (5.10) into (5.12) yields

$$y_m = \sqrt{e} \cdot \mathbf{g}_m^H \mathbf{U}\mathbf{b} + \sqrt{e} \cdot w_m \tag{5.18}$$

where $w_m = \mathbf{q}_m^H \mathbf{n}_m$ is the noise contribution and $\mathbf{g}_m = \mathbf{H}_m^H \mathbf{q}_m$ is a QN_T-dimensional vector that depends on both the channel coefficients and data detection strategy. Finally, letting $\mathbf{y} = [y_1, y_2, ..., y_K]^T$ and $\mathbf{F} = \sqrt{e} \cdot \mathbf{U}$, we obtain

$$\mathbf{y} = \mathbf{G}^H \mathbf{F} \mathbf{b} + \sqrt{e} \cdot \mathbf{w} \tag{5.19}$$

where $\mathbf{G}$ is the following matrix with dimensions $N_T Q \times K$

$$\mathbf{G} = \begin{bmatrix} \mathbf{H}_{1,1}^H \mathbf{q}_1 & \mathbf{H}_{2,1}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,1}^H \mathbf{q}_K \\ \mathbf{H}_{1,2}^H \mathbf{q}_1 & \mathbf{H}_{2,2}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,2}^H \mathbf{q}_K \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{1,N_T}^H \mathbf{q}_1 & \mathbf{H}_{2,N_T}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,N_T}^H \mathbf{q}_K \end{bmatrix} \tag{5.20}$$

while $\mathbf{w} = [w_1, w_2, ..., w_K]^T$ is a Gaussian vector with zero-mean and covariance matrix $\sigma_n^2 \mathbf{I}_Q$.

5.3 Design of The Backward and Forward Matrices

The processing matrices $\mathbf{P}$, $\mathbf{C}$ and $\mathbf{U}$ are designed so as to minimize the sum of the mean-square-errors at all mobile terminals. Since the contribution of $\mathbf{d}$ to the received vector $\mathbf{y}$ is removed by the modulo operator employed at each MT, the desired value for $\mathbf{y}$ is $\mathbf{y} = \mathbf{v}$ instead of $\mathbf{y} = \mathbf{a}$. On the other hand, from (5.6) it follows that $\mathbf{v} = \mathbf{P}^T \mathbf{C} \mathbf{b}$ so that our optimality criterion leads to the minimization of the following cost function

$$J = E\left\{ \left\| \mathbf{y} - \mathbf{P}^T \mathbf{C} \mathbf{b} \right\|^2 \right\} \tag{5.21}$$

where the statistical expectation is computed over data symbols and thermal noise.

To maintain the same power as in the case where no pre-filtering is used, we impose a constraint on the overall transmit power. To make the problem mathematically tractable, we assume that the pre-coded symbols b_k are statistically independent with zero mean and the same power σ_a^2 as the user data. Although not rigourously true, this assumption is reasonable for large M-QAM constellations with size $M \geq 16$ Fisher (2002). In the above hypothesis, the power constraint can be formulated as

$$\text{tr}\left\{\mathbf{U}^H\mathbf{U}\right\} = K \tag{5.22}$$

or, equivalently,

$$\text{tr}\left\{\mathbf{F}^H\mathbf{F}\right\} = e \cdot K \tag{5.23}$$

where we have born in mind that $\mathbf{F} = \sqrt{e} \cdot \mathbf{U}$.

5.3.1 Minimum mean square error design

As pointed out in Chapter 4, in minimizing the cost function (5.21) under the power constraint (5.23) we can follow the line of reasoning employed in Choi & Murch (2004) or, alternatively, the more conventional technique presented in Kusume *et al.* (2005). Since in both cases we arrive at the same final result (see Appendix A.2.2), in the sequel we adopt the approach of Choi & Murch (2004) which is mathematically simpler. We begin by substituting (5.19) into (5.21) and computing the statistical expectation over $\mathbf{b}$ and $\mathbf{w}$. This produces

$$J = \sigma_a^2 \cdot \text{tr}\left\{(\mathbf{G}^H\mathbf{F} - \mathbf{P}^T\mathbf{C})(\mathbf{G}^H\mathbf{F} - \mathbf{P}^T\mathbf{C})^H\right\} + e \cdot K\sigma_n^2 \tag{5.24}$$

where we have borne in mind that $\mathbf{b}$ and $\mathbf{w}$ are statistically independent with zero-mean and covariance matrices $\sigma_a^2\mathbf{I}_Q$ and $\sigma_n^2\mathbf{I}_Q$, respectively. Next, substitut-

ing (5.23) into (5.24) and letting $\rho = \sigma_n^2/\sigma_a^2$, we obtain

$$J = \sigma_a^2 \cdot \mathrm{tr}\left\{\left(\mathbf{G}^H\mathbf{F} - \mathbf{P}^T\mathbf{C}\right)\left(\mathbf{G}^H\mathbf{F} - \mathbf{P}^T\mathbf{C}\right)^H + \rho \cdot \mathbf{F}^H\mathbf{F}\right\}. \tag{5.25}$$

Our objective is to determine the matrices $\mathbf{F}$, $\mathbf{C}$ and $\mathbf{P}$ that minimize the RHS of (5.25). For this purpose, we keep $\mathbf{C}$ and $\mathbf{P}$ fixed and set to zero the gradient of J with respect to $\mathbf{F}$. This yields

$$\mathbf{F} = \mathbf{G}(\mathbf{G}^H\mathbf{G} + \rho\mathbf{I}_K)^{-1}\mathbf{P}^T\mathbf{C} \tag{5.26}$$

or, equivalently,

$$\mathbf{F} = \tilde{\mathbf{G}}(\tilde{\mathbf{G}}^H\tilde{\mathbf{G}} + \rho\mathbf{I}_K)^{-1}\mathbf{C} \tag{5.27}$$

where we have defined $\tilde{\mathbf{G}} = \mathbf{G}\mathbf{P}^T$ and we have used the identity $\mathbf{P}^T\mathbf{P} = \mathbf{I}_K$. Next, we substitute (5.27) into (5.25) to obtain

$$J = \sigma_n^2 \cdot \mathrm{tr}\left\{\mathbf{C}^H(\tilde{\mathbf{G}}^H\tilde{\mathbf{G}} + \rho\mathbf{I}_K)^{-1}\mathbf{C}\right\} \tag{5.28}$$

and look for the unit-diagonal and lower-triangular matrix $\mathbf{C}$ minimizing the RHS of (5.28). This problem is addressed in the Appendix B.1 and the solution is given by

$$\mathbf{C} = \tilde{\mathbf{L}}\tilde{\mathbf{D}} \tag{5.29}$$

where $\tilde{\mathbf{L}}\tilde{\mathbf{L}}^H$ is the Cholesky factorization of $\tilde{\mathbf{G}}^H\tilde{\mathbf{G}} + \rho\mathbf{I}_K$, i.e.,

$$\tilde{\mathbf{G}}^H\tilde{\mathbf{G}} + \rho\mathbf{I}_K = \tilde{\mathbf{L}}\tilde{\mathbf{L}}^H \tag{5.30}$$

while $\tilde{\mathbf{D}}$ is a $K \times K$ diagonal matrix which scales the elements on the main diagonal of $\mathbf{C}$ to unity and reads

$$\tilde{\mathbf{D}} = \text{diag}\left\{1/\left[\tilde{\mathbf{L}}\right]_{k,k}; k = 1, 2, \ldots, K\right\}. \tag{5.31}$$

Substituting (5.29) and (5.30) into (5.27) and recalling that $\mathbf{U} = (1/\sqrt{e})\cdot\mathbf{F}$ yields

$$\mathbf{U} = \frac{1}{\sqrt{e}} \cdot \tilde{\mathbf{G}} \left(\tilde{\mathbf{L}}^{-1}\right)^{H} \tilde{\mathbf{D}} \tag{5.32}$$

where e is found from (5.23) after replacing $\mathbf{F}$ with $\tilde{\mathbf{G}} \left(\tilde{\mathbf{L}}^{-1}\right)^{H} \tilde{\mathbf{D}}$. This produces

$$e = \frac{\text{tr}\left\{\tilde{\mathbf{D}}^{\text{H}}\tilde{\mathbf{L}}^{-1}\tilde{\mathbf{G}}^{\text{H}}\tilde{\mathbf{G}} \left(\tilde{\mathbf{L}}^{-1}\right)^{\text{H}} \tilde{\mathbf{D}}\right\}}{K} \tag{5.33}$$

or, equivalently,

$$e = \frac{\text{tr}\left\{\tilde{\mathbf{D}}^{\text{H}}\tilde{\mathbf{D}} - \rho\tilde{\mathbf{D}}^{\text{H}} \left(\tilde{\mathbf{L}}^{\text{H}}\tilde{\mathbf{L}}\right)^{-1} \tilde{\mathbf{D}}\right\}}{K} \tag{5.34}$$

where we have born in mind that $\tilde{\mathbf{G}}^{H}\tilde{\mathbf{G}} = \tilde{\mathbf{L}}\tilde{\mathbf{L}}^{H} - \rho\mathbf{I}_K$ as indicated in (5.30).

Finally, collecting (5.28)-(5.30) it follows that

$$J = \sigma_n^2 \sum_{k=1}^{K} \frac{1}{\left[\tilde{\mathbf{L}}\right]_{k,k}^2}. \tag{5.35}$$

At this stage we are left with the problem of finding the permutation matrix $\mathbf{P}$ that minimizes the RHS of (5.35), where the dependence on $\mathbf{P}$ is hidden in the quantities $\left[\tilde{\mathbf{L}}\right]_{k,k}$. Unfortunately, the optimal solution to this problem can only be found through an exhaustive search over all $K!$ permutations of $\mathbf{a}$, which becomes prohibitive even in the presence of few active users. An interesting alternative to the exhaustive search is the best-first ordering strategy, which was originally proposed in Wolniansky *et al.* (1998) for V-BLAST and has recently been extended to THP in Kusume *et al.* (2005) in the context of MIMO multi-user. This method operates on the rows of the channel matrix and in most

cases achieves the optimal order with a significant reduction of complexity with respect to the exhaustive search. Clearly, the best-first strategy of Kusume *et al.* (2005) can also be applied to our scheme. However, since we are interested in the ultimate performance achievable by the system under investigation, in this work we find the optimal $\mathbf{P}$ through an exhaustive search. Although computationally demanding, the resulting procedure is still affordable since $K \leq Q$ and Q is typically small in practical applications.

In the sequel, the pre-coding scheme based on (5.29) and (5.31)-(5.34) is referred to as NL-TWF Joham *et al.* (2005). Calling $\tilde{\mathbf{L}}_{opt}$ the matrix $\tilde{\mathbf{L}}$ that corresponds to the optimal choice of $\mathbf{P}$, from (5.35) it follows that the minimum of J is given by

$$J_{\min} = \sigma_n^2 \sum_{k=1}^{K} \frac{1}{\left[\tilde{\mathbf{L}}_{opt}\right]_{k,k}^2}. \tag{5.36}$$

The following remarks are of interest.

1) Setting $\mathbf{B} = \mathbf{0}$ (corresponding to $\mathbf{C} = \mathbf{I}_K$) yields the L-TWF discussed in Chapter 4. In these circumstances, from (5.27) we obtain

$$\mathbf{U} = \frac{1}{\sqrt{e}} \cdot \tilde{\mathbf{G}} \left(\tilde{\mathbf{G}}^H \tilde{\mathbf{G}} + \rho \mathbf{I}_K \right)^{-1} \tag{5.37}$$

where we have born in mind that $\mathbf{U} = (1/\sqrt{e}) \cdot \mathbf{F}$. The scalar e is computed from (5.23) and reads

$$e = \frac{\mathrm{tr}\left\{ \tilde{\mathbf{G}} \left(\tilde{\mathbf{G}}^H \tilde{\mathbf{G}} + \rho \mathbf{I}_K \right)^{-2} \tilde{\mathbf{G}}^H \right\}}{K} \tag{5.38}$$

while the cost function in (5.28) becomes

$$J' = \sigma_n^2 \cdot \mathrm{tr}\left\{ \left(\tilde{\mathbf{G}}^H \tilde{\mathbf{G}} + \rho \mathbf{I}_K \right)^{-1} \right\} \tag{5.39}$$

which can also be rewritten as

$$J' = \sigma_n^2 \sum_{k=1}^{K} \frac{1}{\left[\tilde{\mathbf{L}}\right]_{k,k}^2} + \sigma_n^2 \sum_{k=2}^{K} \sum_{i=1}^{k-1} \left| \left[\tilde{\mathbf{L}}^{-1}\right]_{k,i} \right|^2. \tag{5.40}$$

Comparing (5.40) with (5.35) indicates that $J' \geq J$, meaning that L-TWF cannot perform better than NL-TWF. Interestingly, letting $\tilde{\mathbf{G}} = \mathbf{G}\mathbf{P}^T$ into (5.39) and recalling that $\mathbf{P}^T\mathbf{P} = \mathbf{I}_K$ yields

$$J' = \sigma_n^2 \cdot \mathrm{tr}\left\{\mathbf{P}\left(\mathbf{G}^H\mathbf{G} + \rho\mathbf{I}_K\right)^{-1}\mathbf{P}^T\right\} \tag{5.41}$$

from which, using the identity $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{tr}\{\mathbf{BA}\}$ with $\mathbf{A} = \mathbf{P}$ and $\mathbf{B} = \left(\mathbf{G}^H\mathbf{G} + \rho\mathbf{I}_K\right)^{-1}\mathbf{P}^T$, it follows that the performance of L-TWF is independent of the permutation matrix $\mathbf{P}$.

2) In OFDMA systems the combining coefficients $q_m(n)$ $(n = 1, 2, \ldots, Q)$ at the mth MT are selected according to (5.14). Then, from (5.20) it turns out that $\mathbf{G}^H\mathbf{G} + \rho\mathbf{I}_K$ is diagonal and reads

$$\mathbf{G}^H\mathbf{G} + \rho\mathbf{I}_K = \mathrm{diag}\left\{\rho + \sum_{\ell=1}^{N_T} |H_{k,\ell}(i_{j_k})|^2; k = 1, 2, \ldots, K\right\}. \tag{5.42}$$

The above fact, together with the identity

$$\tilde{\mathbf{G}}^H\tilde{\mathbf{G}} + \rho\mathbf{I}_K = \mathbf{P}\left(\mathbf{G}^H\mathbf{G} + \rho\mathbf{I}_K\right)\mathbf{P}^T \tag{5.43}$$

indicates that even $\tilde{\mathbf{G}}^H\tilde{\mathbf{G}} + \rho\mathbf{I}_K$ is diagonal and the same occurs with $\tilde{\mathbf{L}}$. Hence, from (5.30) we see that $\tilde{\mathbf{D}} = \tilde{\mathbf{L}}^{-1}$ so that (5.29) becomes $\mathbf{C} = \mathbf{I}_K$ and $\mathbf{B}$ reduces to the null matrix. This means that NL-TWF boils down to L-TWF when applied to OFDMA transmissions. The reason is that separating different users at a subcarrier level eliminates the MAI, thereby avoiding the need for non-linear pre-coding. As discussed previously, in these circumstances the cost function J

becomes independent of $\mathbf{P}$. Thus, letting $\mathbf{P} = \mathbf{I}_K$ for simplicity, we have $\tilde{\mathbf{G}} = \mathbf{G}$ and the forward matrix $\mathbf{U}$ in (5.37) takes the form

$$\mathbf{U} = \frac{1}{\sqrt{e}} \cdot \mathbf{G} \left(\mathbf{G}^H \mathbf{G} + \rho \mathbf{I}_K \right)^{-1}. \tag{5.44}$$

Substituting (5.44) into (5.9) and recalling that $\mathbf{b} = \mathbf{a}$, yields

$$\mathbf{s} = \frac{1}{\sqrt{e}} \cdot \mathbf{G} \left(\mathbf{G}^H \mathbf{G} + \rho \mathbf{I}_K \right)^{-1} \mathbf{a}. \tag{5.45}$$

Finally, collecting (5.22), (5.42) and (5.45), it follows that the nth entry of $\mathbf{s}_i$ is given by

$$s_i(n) = \frac{H^*_{m,i}(i_n)\, a_m}{\sqrt{e} \cdot \left[\rho + \sum_{\ell=1}^{N_T} |H_{m,\ell}(i_n)|^2 \right]} \qquad i = 1, 2, \ldots, N_T \tag{5.46}$$

where m is the index of the user allocated over the i_nth subcarrier, i.e., $n = j_m$. The above equation indicates that in OFDMA systems the proposed scheme reduces to MRT over the available antennas, leading to a significant reduction of complexity as compared to MC-CDMA.

The cost function is computed from (5.28) letting $\mathbf{C} = \mathbf{I}_K$ and $\tilde{\mathbf{G}} = \mathbf{G}$. Then, from (6.49) we have

$$J = \sigma_n^2 \cdot \sum_{k=1}^{K} \left[\rho + \sum_{\ell=1}^{N_T} |H_{k,\ell}(i_{j_k})|^2 \right]^{-1}. \tag{5.47}$$

5.3.1.1 Simulation results

Computer simulations have been run to assess the performance of the proposed THP-based scheme. The system parameters are as follows.

The investigated multi-carrier system is inspired by the HiperLAN/II standardization project Medbo (1998) and employs a total of $N = 64$ subcarriers.

The latter are divided into 16 groups, each containing $Q = 4$ elements. To better exploit the channel frequency diversity, the subcarriers belonging to the considered group are uniformly distributed over the signal bandwidth with indexes $i_n = 16(n-1)$ for $n = 1, 2, 3$ and 4. The signal bandwidth is $B = 20$ MHz, so that the useful part of each block has duration $T = N/B = 3.2$ s. The sampling period is $T_s = T/N = 5 \cdot 10^{-2}$ μs and a cyclic prefix of length $T_G = 0.8$ μs is employed to eliminate interblock interference. This corresponds to an extended block (including the cyclic prefix) of 4 μs. The channel frequency response between the ith transmit antenna and the mth MT is expressed by

$$H_{m,i}(i_n) = \sum_{\ell=0}^{L-1} h_{m,i}(\ell)\, e^{-j2\pi\, i_n \ell/N} \qquad n = 1, 2, \ldots, Q \tag{5.48}$$

where $\mathbf{h}_{m,i} = [h_{m,i}(0),\, h_{m,i}(1), \ldots, h_{m,i}(L-1)]^T$ represents the corresponding the discrete-time channel impulse response (CIR) of length $L = 8$. The entries of $\mathbf{h}_{m,i}$ are modeled as independent Gaussian random variables with zero-mean and power

$$\mathrm{E}\left\{|h_{m,i}(\ell)|^2\right\} = \lambda \cdot \exp(-\ell/4) \qquad \ell = 0, 1, \ldots, 7 \tag{5.49}$$

where λ is chosen such that the average energy of $\mathbf{h}_{m,i}$ is normalized to unity, i.e., $\mathrm{E}\left\{|\mathbf{h}_{m,i}|^2\right\} = 1$. The CIRs are kept fixed over the downlink time-slot (slow fading) but vary independently from slot to slot. The transmit antennas are adequately separated so as to make vectors $\mathbf{h}_{m,i}$ statistically independent for different values of i and m.

We assume an uncoded transmission in which the information bits are mapped onto 16-QAM symbols using a Gray map. The number of active users in the considered group of subcarriers is $K = 4$ (fully-loaded system) and perfect channel knowledge is assumed at both the transmit and receive ends.

Figure 5.4 illustrates the MSE vs. $1/\sigma_n^2$ for an MC-CDMA system employing

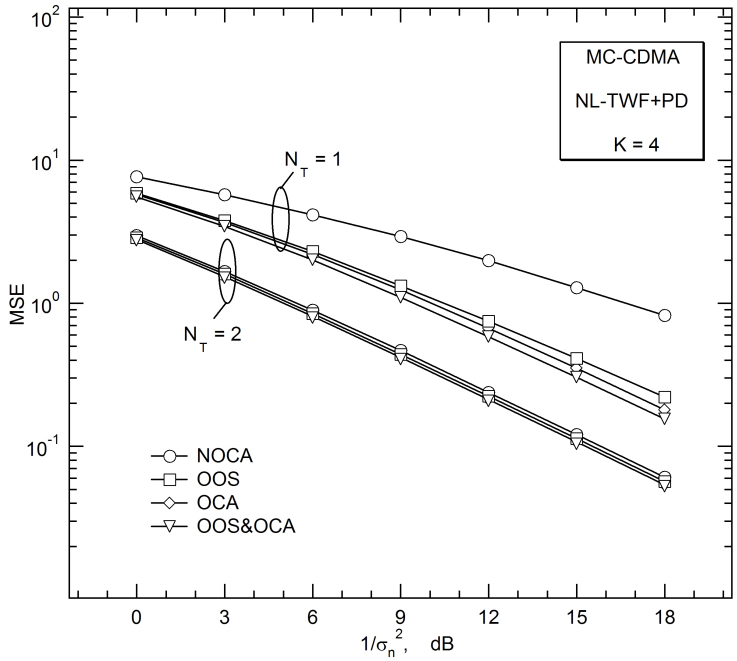

Figure 5.4: MSE vs. $1/\sigma_n^2$ for an MC-CDMA system employing NL-TWF+PD with $N_T = 1$ or 2.

NL-TWF. The results have been obtained through computer simulations by averaging the RHS of (5.35) and (5.36) with respect to the channel statistics. The number of transmit antennas is either $N_T = 1$ or 2 and only a PD operation is accomplished at the receiver. The curve labeled OOS (optimal ordering strategy) has been obtained by selecting the permutation matrix $\mathbf{P}$ that leads to the minimum MSE in (5.36). As mentioned previously, in doing so we have employed an exhaustive search over all possible permutations, although more practical techniques like the best-first strategy Kusume *et al.* (2005) can be used without incurring significant degradations with respect to the exhaustive search. The acronym OCA (optimal code assignment) indicates that the spreading codes are dynamically assigned to the active users at each new simulation run. In practice, calling $\mathbf{c}_{j_m}$ the code assigned to the mth user, we select the indexes $\{j_1, j_2, \ldots, j_K\}$ in

the set $\{1, 2, \ldots, Q\}$ so as to minimize the RHS of (5.36). Clearly, in order to avoid that the same code be assigned to different users, we must ensure that $j_m \neq j_k$ for $m \neq k$. Again, the optimal mapping $m \rightarrow j_m$ is found through an exhaustive search over all $N_a = Q!/(Q-K)!$ possible assignments. The curve labeled OOS&OCA is obtained by jointly looking for the best ordering and code allocation strategies, and requires a search over $N_p \cdot N_a = K!Q!/(Q-K)!$ alternatives at each new channel realization. Albeit too complex for practical implementation, OOS&OCA is useful to assess the ultimate performance achievable by the system under investigation and it is shown here as a benchmark. Finally, the curve labeled NOCA (non-adaptive ordering and code assignment) refers to a system with $\mathbf{P} = \mathbf{I}_k$ and $j_k = k$ for $k = 1, 2, \ldots, K$, i.e., the users' data are not optimally re-ordered before pre-coding nor the spreading codes are adaptively assigned among the active users. The results of Figure 5.4 indicate that both OOS and OCA lead to a significant reduction of the MSE in case of a single transmit antenna while they only provide marginal gains when $N_T = 2$. In any case, it turns out that OCA practically achieves the same performance as OOS&OCA and it is more beneficial than using OOS singularly. As expected, increasing the number of transmit antennas provides the BS with more degrees of freedom to mitigate MAI and channel distortions, thereby leading to a significant improvement of the system performance.

The BER of an MC-CDMA system equipped with NL-TWF is shown in Figure 5.5 as a function of E_T/N_0, where E_T is the *transmitted* energy per bit and N_0 is the two-sided noise power spectral density. The operating conditions are the same as in Figure 5.4, i.e., the receiver employs a simple PD detection strategy and the number of transmit antennas is either $N_T = 1$ or 2. Again, we see that OCA provides better results than OOS. In particular, for $N_T = 2$ OCA has virtually the same performance as OOS&OCA, while for $N_T = 1$ the loss is limited

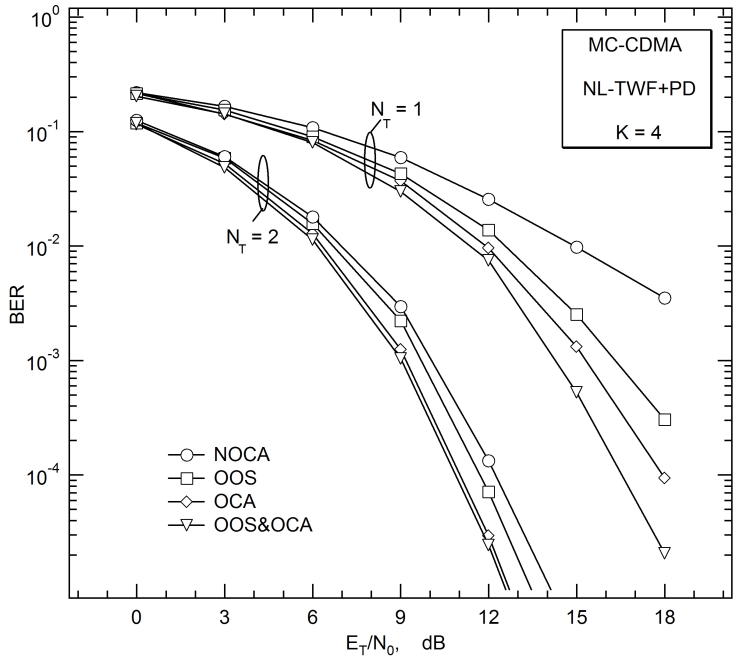

Figure 5.5: BER performance vs. E_T/N_0 for an MC-CDMA system employing NL-TWF+PD with $N_T = 1$ or 2.

to 1dB (at an error rate of 10^{-3}). Note that significant performance degradations occur when NL-TWF operates with NOCA and $N_T = 1$.

Figure 5.6 compares the BER performance of NL-TWF, NL-TZF, L-TWF and the CSR scheme discussed in Cosovic *et al.* (2005) when applied to an MC-CDMA system employing a PD operation at the receiver side. For a fair comparison with L-TWF, results for NL-TWF, NL-TZF and CSR have been obtained without any OCA or OOS. We see that NL-TWF outperforms the other schemes, especially in case of a single transmit antenna. However, NL-TZF is a promising alternative to NL-TWF when $N_T = 2$, as it achieves similar performance without requiring knowledge of ρ at the transmit side. Extensive simulations indicate that OCA and/or OOS improve the performance of the non-linear pre-filtering schemes, but in any case NL-TWF gives the best results.

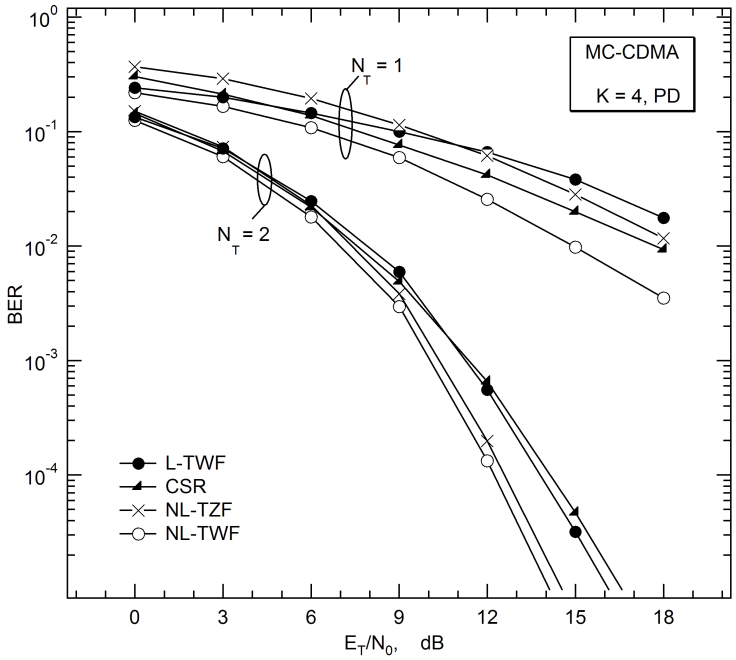

Figure 5.6: BER performance vs. E_T/N_0 for an MC-CDMA system employing various non-linear pre-coding techniques in conjunction with PD.

Figure 5.7 illustrates the impact of various SUD strategies on the BER performance of an MC-CDMA system equipped with NL-TWF. In order to assess the ultimate performance achievable by the considered scheme, we only present results obtained with OOS&OCA. It turns out that MRC is the best strategy, while PD performs poorly since it does not exploit CSI at the receiver side. In particular, at an error rate of 10^{-3} the loss of PD with respect to MRC is approximately 1 dB with $N_T = 2$, and increases to 2 dB with $N_T = 1$. Other simulations (not shown for space limitations) indicate that employing OCA instead of OOS&OCA produces results similar to those in Figure 5.4 when $N_T = 2$, while marginal degradations (smaller than 1 dB) occur with $N_T = 1$.

We now assess the performance of the proposed pre-filtering scheme when used

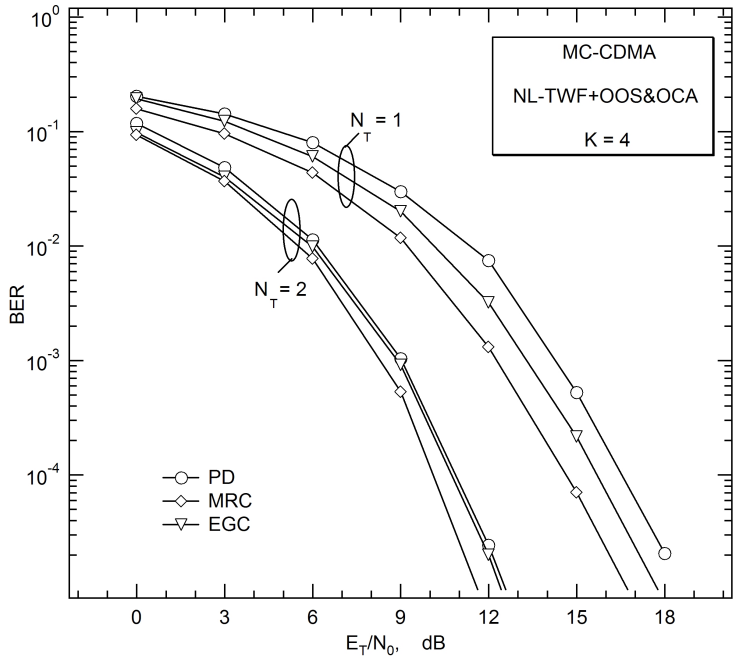

Figure 5.7: BER performance vs. E_T/N_0 for an MC-CDMA system employing NL-TWF in conjunction with OOS&OCA and different SUD strategies.

in conjunction with OFDMA. In this case, the MSE is computed by averaging the RHS of (5.47) with respect to the channel statistics. The results obtained by means of computer simulations are shown in Figure 5.8 as a function of $1/\sigma_n^2$ for either $N_T = 1$ or 2. Here, the acronym OSA (optimal subcarrier assignment) denotes a system in which the available subcarriers are dynamically assigned to the active users according to the actual channel realization. As explained in Sect. II.B, the mth user transmits on the i_{j_m}th subcarrier and the indexes $\{j_1, j_2, \ldots, j_K\}$ are properly selected in the set $\{1, 2, \ldots, Q\}$ so as to minimize the RHS of (5.47). In doing so we resort to an exhaustive search over all possible subcarrier assignments, bearing in mind that $j_m \neq j_k$ for $m \neq k$ since each subcarrier is exclusively assigned to only one user. The curve labeled NSA (non-

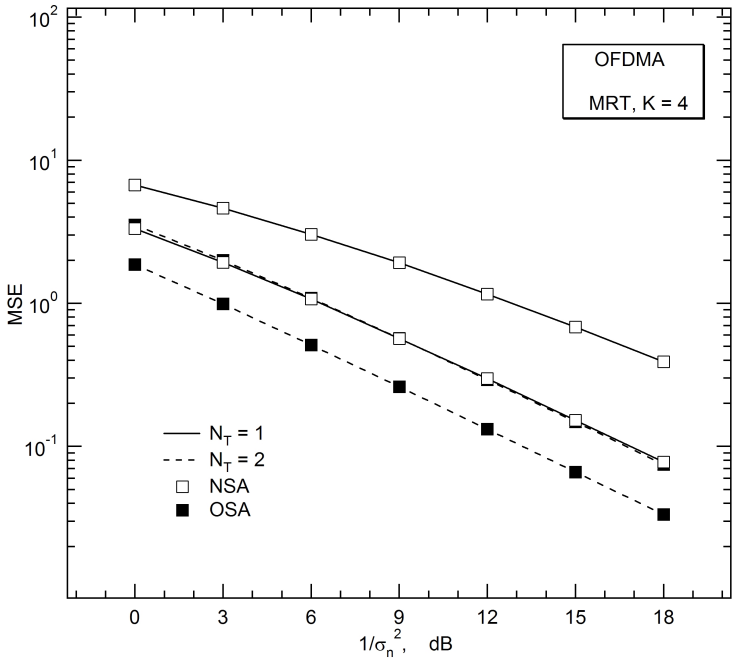

Figure 5.8: MSE vs. $1/\sigma_n^2$ for an OFDMA system employing MRT with $N_T = 1$ or 2.

adaptive subcarrier assignment) refers to a system without OSA where $j_m = m$ for $m = 1, 2, \ldots, K$. As mentioned previously, in OFDMA transmissions NL-TWF reduces to a simple MRT scheme over the available transmit antennas. In these circumstances there is no advantage in adopting a particular ordering strategy for the transmitted data symbols and we may arbitrarily set $\mathbf{P} = \mathbf{I}_k$. The results of Figure 5.5 indicate that OSA has a significant impact on the system performance. In particular, employing OSA with $N_T = 1$ yields the same MSE achieved by NSA with $N_T = 2$.

Figure 5.9 illustrates BER results for an OFDMA system obtained in the same operating conditions of Figure 5.5. Again, we see that OSA dramatically improves the system performance with both $N_T = 1$ and 2. In particular, it is worth noting that curves obtained with OSA are asymptotically steeper than

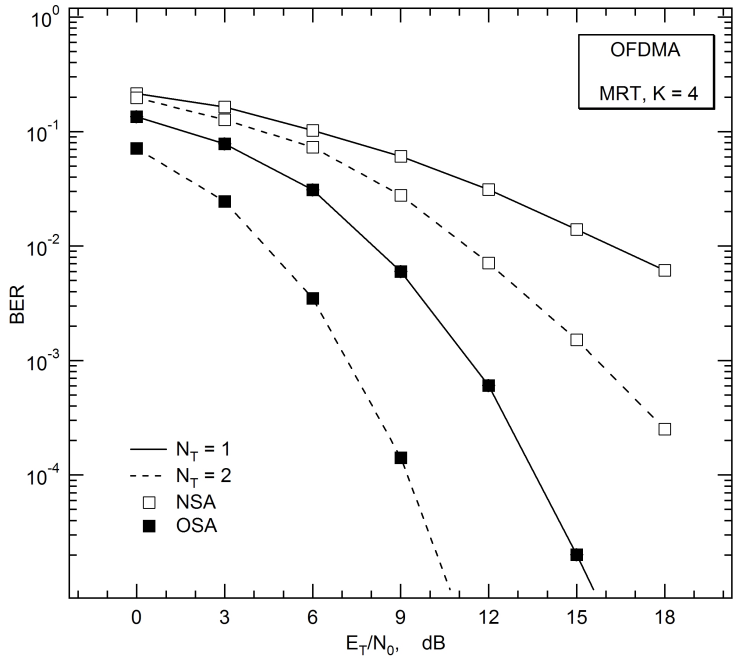

Figure 5.9: BER performance vs. E_T/N_0 for an OFDMA system employing MRT with $N_T = 1$ or 2.

those with NSA. This is a manifestation of the *multiuser diversity gain*, which occurs in OFDMA systems in the presence of dynamic subcarrier assignment Li *et al.* (2003). Interestingly, we see that employing OSA with $N_T = 1$ provides a much lower BER than using NSA with $N_T = 2$, although both configurations are characterized by the same MSE (see Figure 5.5). A possible explanation is that the BER is strictly related to the statistical distribution of the disturbance term that affects the decision statistic and, accordingly, it is not exclusively dictated by the MSE, which only represents the power of the disturbance. This means that different systems characterized by the same MSE may exhibit different BER performance.

Figure 5.10 illustrates the ultimate BER performance achievable by the proposed pre-filtering technique when applied to MC-CDMA and OFDMA. In par-

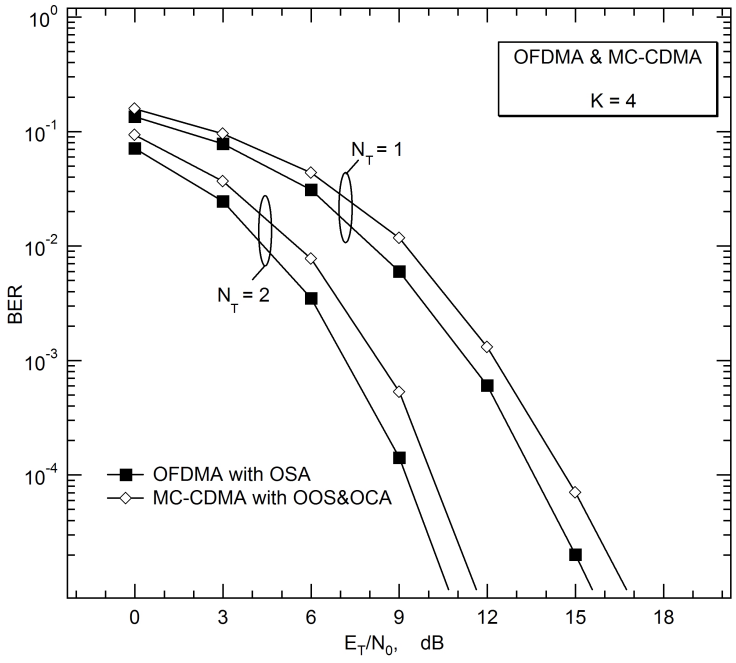

Figure 5.10: BER performance vs. E_T/N_0 for an OFDMA system employing MRT with $N_T = 1$ or 2.

ticular, OFDMA operates with OSA while MC-CDMA employs OOS&OCA in conjunction with MRC as a detection strategy. We see that OFDMA outperforms MC-CDMA and achieves a gain of approximately 1 dB with both $N_T = 1$ and 2. These results indicate that OFDMA has some potential advantages over MC-CDMA, as it achieves better error rate performance with a reduced complexity. Indeed, the computational load involved in OFDMA is rather limited for the following two reasons. First, no data re-ordering is required at the transmit side. Second, the decision statistic at the mth MT is simply given by the i_{j_m}th output of the DFT unit, whereas in MC-CDMA systems it is obtained by linearly combining the DFT outputs and passing the result through a modulo device. As discussed previously, data re-ordering at the transmitter can be avoided even in MC-CDMA without incurring significant performance degradations (provided

that OCA is employed). However, data detection still remains more complex than in OFDMA, thereby leading to higher power consumption at the MTs with a corresponding reduction of the battery life.

Conclusions

We have considered a unified framework for non-linear pre-filtering in MC-CDMA and OFDMA downlink transmissions. The proposed scheme exploits channel knowledge at the transmit side and aims at minimizing the sum of the mean square errors at all MTs. In this way MAI and channel distortions are pre-compensated at the transmitter and low complex SUD schemes can be employed at the receiver for data detection. Specifically, the following results have been found:

1) The performance of MC-CDMA is greatly improved when the available codes are properly assigned to the active users (OCA). Adopting an optimal pre-coding order (OOS) in addition to OCA provides only marginal gains at the price of a significant increase of complexity.

2) Among the considered SUD techniques, MRC gives the best results in a pre-filtered MC-CDMA system operating with OCA and/or OOS.

3) When applied to MC-CDMA, NL-TWF outperforms conventional linear pre-filtering techniques as well as other existing non-linear schemes based on a Zero-Forcing approach.

4) In OFDMA applications the proposed pre-filtering scheme reduces to MRT over the available antennas and the decision statistics at the receiver are directly taken at the DFT output without any further processing. This leads to a significant reduction of complexity with respect to MC-CDMA.

5) OSA among the active users produces dramatic improvements in the error rate performance of an OFDMA system.

6) A comparison between OFDMA (with OSA) and MC-CDMA (with OCA and/or OOS) indicates that the former achieves better performance with reduced complexity.

5.3.2 Zero-forcing design

A possible drawback of NL-TWF is that it requires knowledge of ρ. This parameter is related to the noise power σ_n^2 and it is not available at the BS. Moreover, the afore-mentioned scheme produces excellent results but does not allow the BS to control the relative performance of the active users, which are only dictated by the actual realization of the channel matrix. For this reason, they are not suited for commercial multimedia applications, where several types of information (audio, video, images) with different QoS requirements must be supported simultaneously.

To overcome these problems, in the sequel we propose a THP-based pre-filtering scheme for MC-CDMA downlink transmissions in which the BS is equipped with multiple transmit antennas and can allocate the transmit power according to the QoS requirements of each active user. In doing so we follow a two-step procedure where a ZF approach is first employed to completely remove the MAI from the received samples and the result is then exploited to maximize the SNR at all MTs. The resulting scheme allows the BS to allocate the available power according to the QoS requirements of each user making it well suited for multimedia communication systems where the traffic is a mixture of heterogenous applications, each characterized by possibly different QoS constraints.

From (5.18) we see that the received signals depend on the THP matrices $\mathbf{U}$ and $\mathbf{C}$. The latter must be designed so as to mitigate the MAI while enhancing the desired signal component and providing any desired set of relative SNRs at the MTs. For this purpose, we adopt a two-step procedure which operates as

follows. A ZF approach is first employed to completely remove the MAI from the received samples. The result is then exploited to maximize the SNR at all MTs. To maintain the same power as in the case where no pre-filtering is performed, the pre-filtering matrices are derived so as to meet the power constraint (5.22).

From (5.19) we see that the complete elimination of MAI requires that

$$\mathbf{y} = \sqrt{c} \cdot \mathbf{\Lambda v} + \mathbf{w} \tag{5.50}$$

where c is a non-negative real-valued parameter that must be chosen so as to meet the power constraint (5.22) while $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ is a diagonal matrix normalized so as to satisfy the identity $\text{tr}\{\mathbf{\Lambda}\} = K$. The entries of $\mathbf{\Lambda}$ are non-negative real-valued parameters that allows the BS to meet different QoS constraints.

The sample y_m is fed to the AGC unit and ideally scaled by $\lambda_m\sqrt{c}$. Next, it is passed to the same non-linear device employed at the transmitter. Neglecting the modulo-folding effect on the thermal noise Fisher (2002), the decision statistic for a_m takes the form

$$z_m = a_m + \frac{w_m}{\lambda_m\sqrt{c}} \tag{5.51}$$

and the corresponding SNR reads

$$SNR_m = c \cdot \frac{\lambda_m^2\sigma_a^2}{\sigma_w^2}. \tag{5.52}$$

Comparing (5.19) with (5.50) implies that

$$\mathbf{GU} = \sqrt{c}\mathbf{\Lambda C} \tag{5.53}$$

or, equivalently,

$$\mathbf{C}^{-1}\tilde{\mathbf{G}}\mathbf{F} = \mathbf{I} \tag{5.54}$$

where we have defined $\tilde{\mathbf{G}} = \mathbf{\Lambda}^{-1}\mathbf{G}$ and $\mathbf{F} = (1/\sqrt{c}) \cdot \mathbf{U}$. In these circumstances, the power constraint (5.22) becomes

$$\text{tr}\left\{\mathbf{F}^H\mathbf{F}\right\} = \frac{K}{c}. \tag{5.55}$$

The ZF condition in (5.50) is now exploited to design $\mathbf{F}$. For this purpose, we observe that (5.50) is a system of K^2 linear equations whose unknowns are the $PQ \times K$ entries of $\mathbf{F}$. Since $K \leq Q$, we observe that (5.50) may have more unknowns than equations. In such a case there exists an infinite number of matrix $\mathbf{F}$ satisfying (5.50) and the problem is to find the *best* one. As mentioned earlier, we take the maximization of the SNR at each MT as an optimality criterion.

Solving (5.51) with respect to c and substituting into (5.37) yields

$$SNR_m = \frac{\lambda_m^2\sigma_a^2}{\sigma_w^2} \cdot \frac{K}{\text{tr}\{\mathbf{F}^H\mathbf{F}\}} \tag{5.56}$$

from which it follows that the matrix $\mathbf{F}$ maximizing the right-hand side (RHS) of (5.56) is the minimum-norm solution of (5.50). The latter is found as the pseudo-inverse of $\mathbf{C}^{-1}\tilde{\mathbf{G}}$ Kay (1993) and reads $\mathbf{F} = \tilde{\mathbf{G}}^H(\tilde{\mathbf{G}}\tilde{\mathbf{G}}^H)^{-1}\mathbf{C}$ or, equivalently,

$$\mathbf{F} = \mathbf{G}^H\left(\mathbf{G}\mathbf{G}^H\right)^{-1}\mathbf{\Lambda}\mathbf{C} \tag{5.57}$$

where we have borne in mind that $\tilde{\mathbf{G}} = \mathbf{\Lambda}^{-1}\mathbf{G}$. Substituting (5.57) into (5.56), we obtain

$$SNR_m = \frac{\lambda_m^2\sigma_a^2}{\sigma_w^2} \cdot \frac{K}{\text{tr}\{\mathbf{C}^H\mathbf{\Lambda}^H(\mathbf{G}\mathbf{G}^H)^{-1}\mathbf{\Lambda}\mathbf{C}\}}. \tag{5.58}$$

At this stage we are left with the problem of designing the unit-diagonal and lower triangular matrix $\mathbf{C}$ maximizing the RHS of (5.57). The solution to this problem is provided in Appendix B.2 and reads

$$\mathbf{C} = \mathbf{\Lambda}^{-1}\mathbf{S}\mathbf{D} \tag{5.59}$$

where $\mathbf{S}\mathbf{S}^H$ is the Cholesky factorization of $\mathbf{G}\mathbf{G}^H$ and $\mathbf{D} = \text{diag}\{d_1, d_2, \ldots, d_K\}$ is a $K \times K$ diagonal matrix that scales the elements $[\mathbf{C}]_{k,k}$ to unity, i.e., $d_k = \lambda_k / [\mathbf{S}]_{k,k}$.

Collecting (5.59) and (5.57) and recalling that $\mathbf{U} = \sqrt{c} \cdot \mathbf{F}$ produces

$$\mathbf{U} = \sqrt{c} \cdot \mathbf{G}^H \left(\mathbf{S}^{-1}\right)^H \mathbf{D} \tag{5.60}$$

where c is obtained from (5.55) and reads

$$c = \frac{K}{\sum_{n=1}^{K} \frac{\lambda_n^2}{[\mathbf{S}]_{n,n}^2}}. \tag{5.61}$$

Finally, substituting (5.61) into (5.52) we obtain

$$SNR_m = \frac{\lambda_m^2 \sigma_a^2}{\sigma_w^2} \cdot \frac{K}{\sum_{n=1}^{K} \left(\lambda_n / [\mathbf{S}]_{n,n}\right)^2}. \tag{5.62}$$

Since

$$\frac{SNR_m}{SNR_j} = \frac{\lambda_m^2}{\lambda_j^2} \tag{5.63}$$

it follows that the BS can ensure any set of relative SNRs at the MTs by properly selecting the elements of $\mathbf{\Lambda}$. In the sequel, the scheme that employs the matrices $\mathbf{C}$ and $\mathbf{U}$ given in (5.59) and (5.60) is called ZF-THP algorithm.

1) It is well known that the ordering strategy adopted during the pre-coding process has a significant impact on the performance of THP. Unfortunately, the optimal solution to this problem can only be found through an exhaustive search over all $K!$ permutations of vector $\mathbf{a}$. An interesting alternative is the *best-first* ordering strategy which was originally proposed in Wolniansky *et al.* (1998) for V-BLAST systems and recently extended to THP in Kusume *et al.* (2005) and Liu & Krzymien (2005). This method operates on the rows of the channel matrix

and in most cases achieves the optimal ordering with a significant reduction of complexity with respect to the exhaustive search. In the sequel, we adopt the ordering method proposed in Liu & Krzymien (2005) after replacing $\mathbf{H}$ with $\mathbf{\Lambda}^{-1}\mathbf{G}$.

2) Setting $\mathbf{\Lambda} = \mathbf{I}$ into (5.59) yields $\mathbf{C} = \mathbf{SD}$ while $\mathbf{U}$ is still given in (5.60) with $c = K/\sum_{n=1}^{K} \left(1/\left[\mathbf{S}\right]_{n,n}\right)^2$. It is interesting to compare this result with the TH-based scheme discussed in Cosovic *et al.* (2005) by Cosovic, Sand and Raulefs (CSR), where $\mathbf{C} = \mathbf{DS}$ and $\mathbf{U} = \mathbf{G}^H \left(\mathbf{S}^{-1}\right)^H$. In this case, the SNR at the jth MT takes the form

$$SNR_m = \frac{\sigma_a^2 \left[\mathbf{S}\right]_{m,m}^2}{\sigma_w^2} \tag{5.64}$$

and it is only dictated by the actual channel realization since $\mathbf{S}$ is taken from the QR decomposition of $\mathbf{G}$. This means that CSR cannot allocate the transmit power according to the QoS requirements of the active users.

3) Letting $\mathbf{B} = \mathbf{0}$ (corresponding to $\mathbf{C} = \mathbf{I}_K$) reduces ZF-THP to ZF-L precoding scheme. In these circumstances $\mathbf{U}$ becomes

$$\mathbf{U} = \sqrt{c'} \cdot \mathbf{G}^H \left(\mathbf{G}\mathbf{G}^H\right)^{-1} \mathbf{\Lambda} \tag{5.65}$$

with $c' = K/\mathrm{tr}\{\mathbf{\Lambda}^H(\mathbf{G}\mathbf{G}^H)^{-1}\mathbf{\Lambda}\}$, while the SNR at the kth MT reads

$$SNR'_m = \frac{\lambda_m^2 \sigma_a^2}{\sigma_w^2} \cdot \frac{K}{\sum_{n=1}^{K} \left(\lambda_n/\left[\mathbf{S}\right]_{n,n}\right)^2 + \sum_{n=2}^{K} \sum_{\ell=1}^{n-1} \left|\left[\mathbf{S}^{-1}\mathbf{\Lambda}\right]_{n,\ell}\right|^2}. \tag{5.66}$$

Comparing (5.66) with (5.62) indicates that $SNR'_m \leq SNR_m$, meaning that ZF-L cannot perform better than ZF-THP.

5.3.2.1 Numerical results

Computer simulations have been run to assess the performance of the proposed pre-filtering scheme. The total number of subcarriers is $N = 64$ and WH codes of length $Q = 4$ are used for spreading purposes. To better exploit the channel frequency diversity, the subcarriers of the considered group are uniformly distributed over the signal bandwidth with indexes $i_n = 16(n-1)$ for $n = 1, 2, 3$ and 4. The signal bandwidth is $B = 25$ MHz, so that the useful part of each MC-CDMA block has duration $T = N/B = 2.5$ μs. We consider an uncoded transmission in which the information bits are mapped onto 16-QAM symbols using a Gray map. The channel model is inspired by the HiperLAN/II standardization project and it has been extended to the case of multiple transmit antennas by considering N_T independent realizations for each active user. The channels are kept fixed over the downlink time-slot (slow fading), but vary independently from slot to slot. Their average energy is the same for all users.

The performance of the proposed pre-filtering scheme is evaluated in terms of average BER for *each* active user vs. E_T/N_0, where E_T is the *transmitted* energy per bit and $N_0/2$ is the two-sided noise power spectral density. The number of active users is $K = 4$ (fully-loaded system) and perfect channel knowledge is assumed at both the transmit and receive ends.

Figure 5.11 compares the BER of ZF-THP with that of CSR when $\mathbf{\Lambda} = \mathbf{I}_K$. The number of transmit antennas is either $N_T = 1$ or 2 and only a PD operation is performed at the receiver. In both schemes the information symbols are pre-coded *without* employing any ordering strategy. We see that letting $P = 2$ provides the system with spatial diversity and leads to substantial performance improvement with respect to $N_T = 1$. As mentioned earlier, CSR leads to large differences in the error rate performance of the active users. Viceversa, a single curve is shown for ZF-THP since setting $\lambda_k = 1$ into (5.62) for $1 \leq k \leq K$ ensures the same

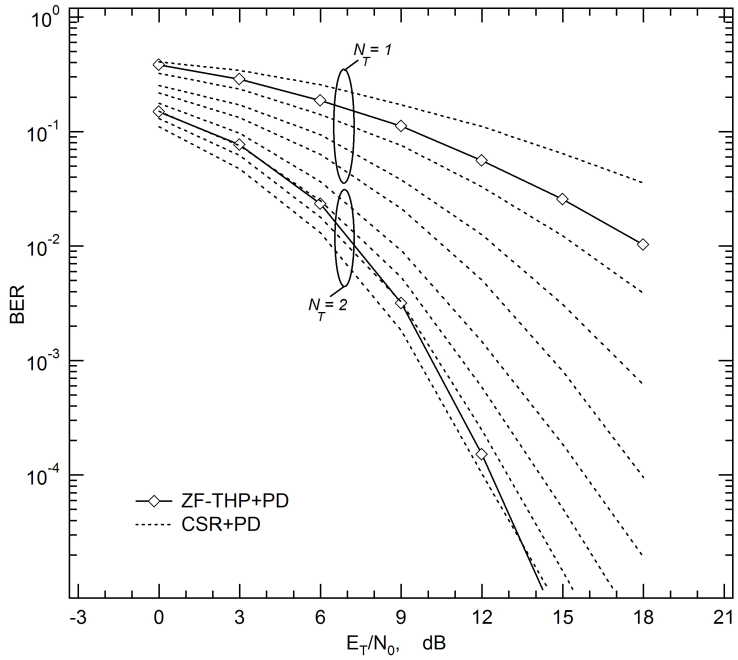

Figure 5.11: BER performance vs. E_T/N_0 with $K = 4$ and $N_T = 1$ or 2 using PD as a detection strategy.

SNR for all active users, thereby leading to a common BER. Compared to CSR, ZF-THP significantly improves the performance of the worst users. In particular, at an error rate of 10^{-3} the gain of ZF-THP with respect to the worst CSR curve is approximately 2 dB for $N_T = 1$, while it reduces to 1.5 dB when $N_T = 2$.

The results of Figure 5.12 have been obtained in the same operating conditions of Figure 5.11, except that now the active users are ordered according to the *best-first* method discussed in Liu & Krzymien (2005). A comparison with Figure 5.11 indicates that the users' ordering improves the BER performance of ZF-THP and dramatically reduces the error rate differences among the active users when CSR is employed. It turns out that both schemes have virtually the same performance for $N_T = 1$, whereas ZF-THP gives the best results when $N_T = 2$. Although the

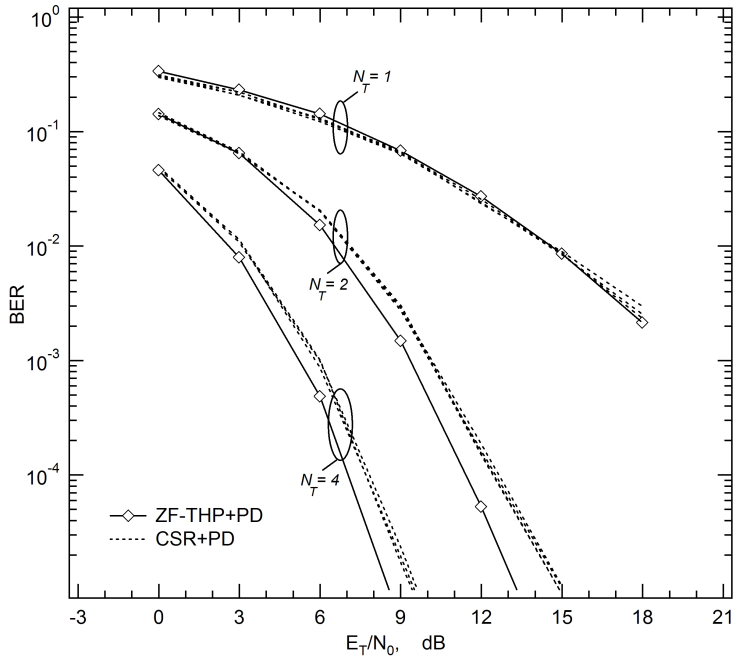

Figure 5.12: BER performance vs. E_T/N_0 with $K = 4$ and $N_T = 1$ or 2 using the best-first ordering strategy in conjunction with PD.

gain with respect to CSR is limited to 1 dB at an error rate of 10^{-3}, it should be noted that it comes for free since both schemes have comparable complexity.

Figure 5.13 shows the BER of ZF-THP when the users are grouped into two separate subsets S_1 and S_2, each containing two elements and characterized by a different QoS requirement. In particular, we set $\lambda_1 = \lambda_2 = (\sqrt{2}-1)/\sqrt{2}$ for the users belonging to S_1, while a 3dB penalty is imposed on the users in S_2 by letting $\lambda_3 = \lambda_4 = (\sqrt{2}-1)/2$. Comparisons are only possible with ZF-LP since CSR cannot ensure pre-assigned SNRs at the MTs. The symbols are pre-coded using the best-first ordering strategy and a PD technique is performed at the receiver. It turns out that ZF-THP outperforms ZF-LP with either $N_T = 1$ or 2. In particular, for an error rate of 10^{-3} the gain of ZF-THP with respect to

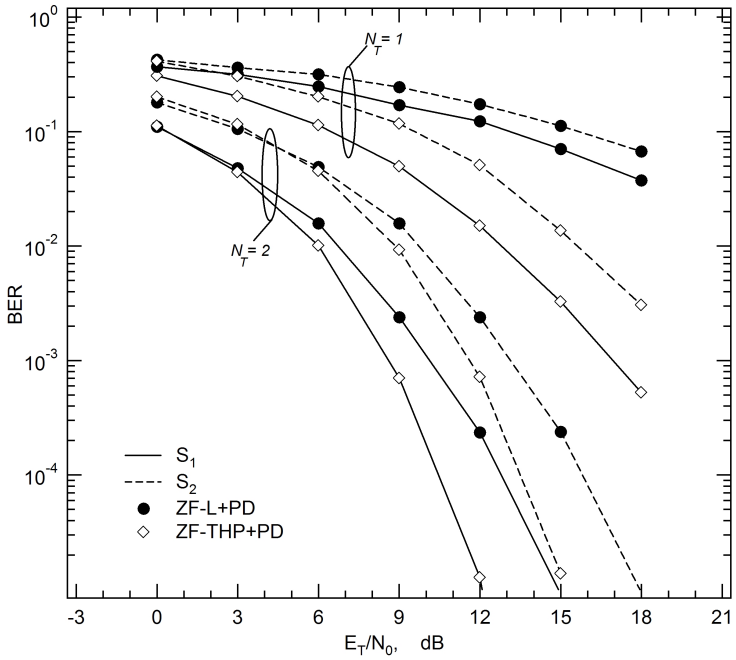

Figure 5.13: BER performance vs. E_T/N_0 with $N_T = 1$ or 2 in the presence of four active users with different QoS requirements and using the best-first ordering strategy in conjunction with PD.

ZF-LP is approximately 1.5 dB when $N_T = 2$.

Figure 5.14 illustrates the performance of ZF-THP in conjunction with different SUD schemes. The users have the same QoS (i.e., $\mathbf{\Lambda} = \mathbf{I}_K$) and are ordered according to the best-first strategy. We see that ZF-THP+EGC is the best scheme for $N_T = 1$ and $E_t/N_0 > 9$ dB, while ZF-THP+MRC takes the lead for $N_T = 2$. At first sight, the fact that ZF-THP+MRC performs worse than ZF-THP+EGC (in case of a single transmit antenna) may appear surprising. A possible explanation is that MRC is effective against thermal noise, but it enhances the MAI. The latter is pre-compensated at the transmit side at the price of some power boosting, which reduces the SNR at the input of the decision device with a cor-

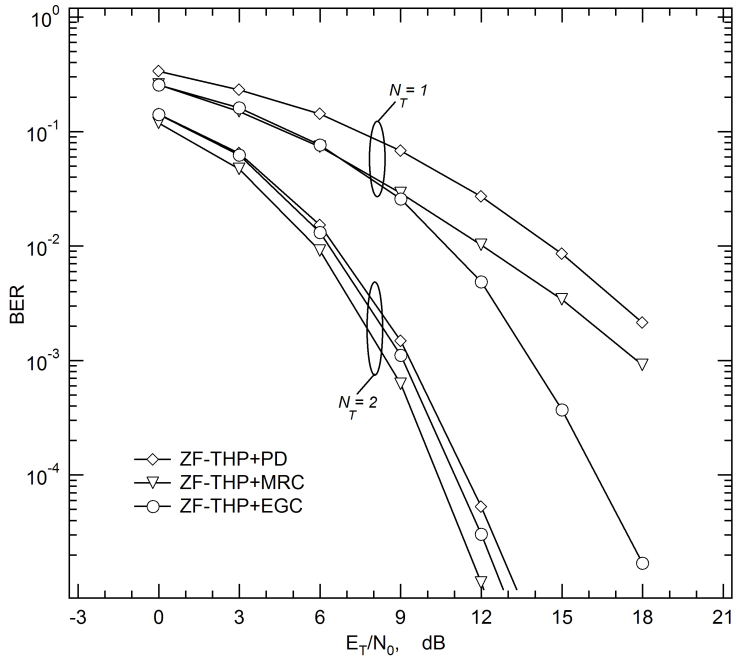

Figure 5.14: BER performance vs. E_T/N_0 with $K = 4$ and $N_T = 1$ or 2 using different SUD schemes and the best-first ordering strategy.

responding degradation of the system performance. On the other hand, letting $N_T = 2$ provides the transmitter with more degrees of freedom, which can be effectively exploited by ZF-THP to mitigate the MAI without excessive power boosting. In such a case the main impairment at the receive side is represented by thermal noise and MRC becomes the best combining strategy.

Chapter 6

Channel Acquisition and Tracking for MC-CDMA Uplink Transmissions

As discussed throughout the dissertation, all pre-filtering schemes, in order to work properly, require explicit knowledge of the channel parameters of the active users. As already mentioned, in TDD systems this information can be achieved by exploiting the channel reciprocity between alternative uplink and downlink transmissions. If channel variations are sufficiently slow, the channel estimate derived at the BS during an uplink time-slot can be reused for pre-filtering in the subsequent downlink time-slot. In addition to pre-filtering, coherent detection of the received signals during the uplink time-slot requires knowledge of the channel responses. For this reasons, in the next, we investigate channel acquisition and tracking in the uplink of MC-CDMA systems.

The problem of channel estimation for MC-CDMA systems has received great attention in the last few years and several solutions already exist Kaiser & Hoeher (1997)- Sanguinetti *et al.* (2004). In downlink transmissions, all the signals

arriving from the BS to a given MT propagate through the same channel. This facilitates the channel estimation task, which can be accomplished with the same methods employed for OFDM applications, i.e., using known symbols inserted in both the frequency and time dimensions Kaiser & Hoeher (1997) or employing a channel-sounding approach in which a train of pulses is periodically transmitted Cacopardi *et al.* (1997). The above schemes are simple to implement but, unfortunately, are not suited for uplink transmissions. The reason is that the active users transmit from different locations and the uplink signals arrive at the BS after passing through different channels. This means that the BS must estimate a large number of parameters, which is expected to degrade the quality of the channel estimates.

A popular approach for estimating the uplink channel responses of the active users is based on subspace methods Tureli *et al.* (2000)- Deng *et al.* (2004). The main advantage of these algorithms is that they operate in a blind fashion and do not need any training overhead. However, they are computationally intense due to complicated matrix manipulations. As an alternative to subspace-based methods, pilot-aided schemes have been proposed in the literature Zheng *et al.* (2004)- Sanguinetti *et al.* (2004). In particular, the algorithm in Zheng *et al.* (2004) exploits properly designed training sequences to separate the signals of different users in the time-domain. A drawback of this scheme is that each user transmits a *single* time-domain sample during the training block and this results in an excessive PAPR. Users' separation can also be achieved in the frequency-domain through FDMA techniques. This approach is adopted in Marques *et al.* (2003), where each user transmits pilot tones on exclusively assigned subcarriers and channel estimation is performed by interpolating between pilots. The main advantage of this solution is that the pilot tones are MAI-free. However, the MTs must continuously switch between two different technologies (i.e., OFDMA

during the training period and MC-CDMA during the data section of the frame), which may be difficult in practical applications where the complexity of the MT must be kept as low as possible to reduce the power consumption. Also, the schemes in Zheng *et al.* (2004) and Marques *et al.* (2003) do not lend themselves to an iterative implementation and, accordingly, they cannot be used to track the channel variations. The only possibility is to recompute the channel estimates periodically by exploiting pilot blocks inserted into the frame structure (with a separation distance smaller than the channel-coherence time). This solution, however, comes at the price of an increased overhead.

In the sequel we address the problem of channel acquisition and tracking in the uplink of an MC-CDMA system equipped with multiple receive antennas. The noise power at each receive branch is also estimated and exploited for data detection purposes. We propose two channel acquisition schemes based on the ML criterion. The first is an extension of the method discussed in Sanguinetti *et al.* (2004) to include the estimation of the noise power. It assumes independently faded subcarriers and, for this reason, it is called UCE. The second is referred to as SCE and effectively exploits the fading correlation among adjacent subcarriers to improve the estimation accuracy. Both schemes require knowledge of the transmitted data symbols. To this purpose, we assume that some blocks carrying known symbols (training blocks) are placed at the beginning of each frame. Channel tracking is pursued by means of LMS or RLS techniques using data decisions from the PPIC receiver proposed in Divsalar *et al.* (1998) and revised in Chapter 3.

The contribution of this part is twofold. First, we derive channel and noise power estimation schemes for MC-CDMA uplink transmissions in the presence of rapidly-varying fading channels. Second, we show how the fading correlation among adjacent subcarriers can be exploited to improve the estimation accuracy

or, alternatively, to reduce the training overhead. Note that the problem of channel tracking in the uplink of an MC-CDMA system has only been discussed in Sanguinetti *et al.* (2004) and no other solution is available in the technical literature.

6.1 MC-CDMA uplink transmissions : System Model

Transmitter

We consider the uplink of an MC-CDMA system in which the total number of subcarriers, N, is divided into smaller groups of Q elements Fazel & Kaiser (2003). We denote $I = N/Q$ the number of groups and $\{j_n^{(i)};\ 1 \leq n \leq Q\}$ the subcarrier indexes within the ith group. In order to exploit the channel frequency diversity, the Q subcarriers are uniformly spread over the signal bandwidth, i.e., we let $j_n^{(i)} = (n-1)I + i$ for $i = 0, 1, \ldots, I-1$. We denote K the number of simultaneously active users ($K \leq Q$) and assume that *each* of them transmits one data symbol in *each* group using its specific spreading code. The users are time aligned to the BS reference in a way similar to that discussed in Morelli (2004).

We call $\{a_k^{(i)}(m); i = 0, 1, \ldots, I-1\}$ the symbols transmitted by the kth user during the mth block. As shown in Figure 6.1, each $a_k^{(i)}(m)$ is spread over Q chips using a WH spreading code $\mathbf{c}_k = [c_k(1), c_k(2), \ldots, c_k(Q)]^T$, where $c_k(n) \in \{\pm 1/\sqrt{Q}\}$ and $(\cdot)^T$ means transpose operation. The entries of the resulting vectors $\{a_k^{(i)}(m)\mathbf{c}_k;\ i = 0, 1, \ldots, I-1\}$ are frequency interleaved and mapped on N subcarriers using an OFDM modulator which comprises an IDFT unit and the insertion of an N_G-point CP to avoid interference between adjacent blocks. The

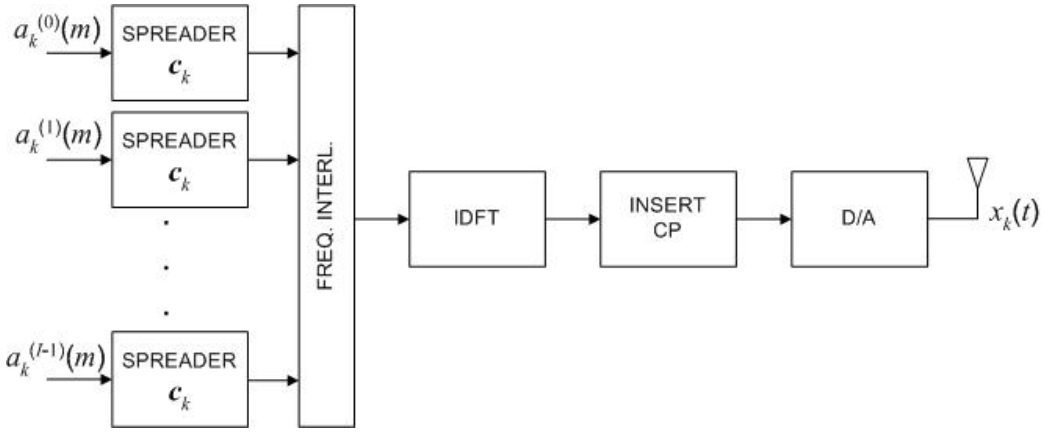

Figure 6.1: Block diagram of the kth transmitter in the uplink of an MC-CDMA network.

resulting time-domain samples $\{s_k(m,n); -N_G \leq n \leq N-1\}$ are finally fed to a D/A converter with impulse response $p(t)$ and signaling interval T_S.

The complex envelope of the signal transmitted by the kth user takes the form

$$x_k(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-N_G}^{N-1} s_k(m,n) p(t - nT_S - mT_B) \tag{6.1}$$

where n counts the samples within a block and $T_B = (N + N_G)T_S$ is the duration of the cyclically extended block.

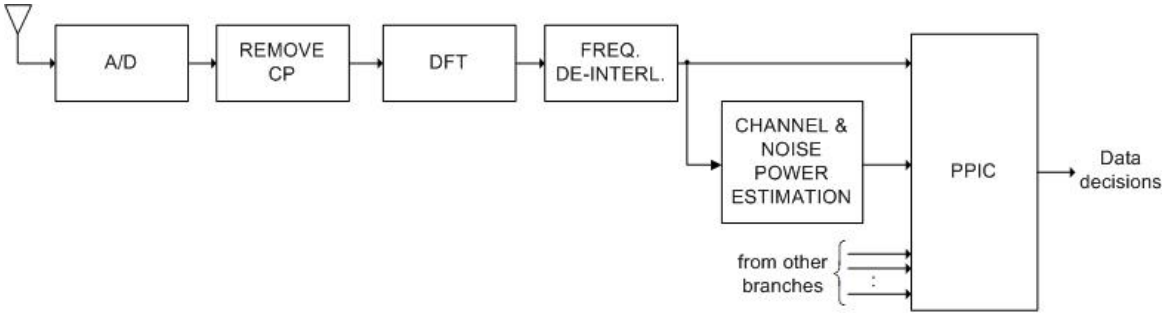

Figure 6.2: Block diagram of the BS receiver.

Receiver

The BS is equipped with P antennas arranged in a uniform linear array. Figure 6.2 illustrates the block diagram of the pth receive branch. The received signal is fed to an A/D converter, which comprises a receive filter and a sampling operation with period T_s. Next, the CP is discarded and the remaining samples are passed to an N-point DFT unit. After frequency de-interleaving, the DFT output is finally passed to a PPIC receiver which also exploits estimates of the channel responses and noise power to compute the final data decisions.

We call $\mathbf{X}_p(m) = [X_p(m,0),\, X_p(m,1), \ldots, X_p(m,N-1)]^T$ the DFT output during the mth block at the pth antenna. Then, denoting $H_{p,k}(m,\ell)$ the kth channel frequency response over the ℓth subcarrier, we may write

$$X_p(m, j_n^{(i)}) = \sum_{k=1}^{K} a_k^{(i)}(m) c_k(n) H_{p,k}(m, j_n^{(i)}) + w_p(m, j_n^{(i)}) \quad 1 \le n \le Q, \quad 0 \le i \le I-1 \tag{6.2}$$

where $w_p(m, j_n^{(i)})$ represents the thermal noise and it is modelled as a complex-valued Gaussian random variable with statistically independent real and imaginary parts, each having zero-mean and variance $\sigma_p^2/2$.

Channel Model

The signals transmitted by the active users propagate through different channels and undergo frequency selective fading. For the sake of clarity, in the next, we revise the underlying channel model.

We consider multipath channels with N_p distinct paths and denote δ the inter-element spacing in the antenna array. Thus, the impulse response of the baseband channel between the kth user and the pth receive antenna during the mth MC-CDMA block takes the form

$$h_{p,k}(m,t) = \sum_{r=1}^{N_p} \alpha_{k,\,r}(m) e^{j(p-1)\omega_{k,\,r}(m)} g(t - \tau_{k,\,r}(m)) \qquad p = 1, 2, \ldots, P \quad (6.3)$$

where $g(t)$ is the convolution between the transmit and receive filters, $\alpha_{k,\,r}(m)$ is the complex gain of the rth path and $\tau_{k,\,r}(m)$ the corresponding delay. The quantities $\{\alpha_{k,\,r}(m)\}$ are modeled as narrow-band independent Gaussian random processes with zero-mean and average power $\sigma^2_{k,\,r} = \mathrm{E}\{|\alpha_{k,\,r}(m)|^2\}$. Also, $\omega_{k,\,r}(m)$ is defined as

$$\omega_{k,\,r}(m) = \frac{2\pi}{\lambda} \delta \sin\left[\varphi_{k,\,r}(m)\right] \quad (6.4)$$

where λ denotes the free-space wavelength and $\varphi_{k,\,r}(m)$ is the DOA of the rth path during the mth MC-CDMA block.

The channel frequency response $H_{p,k}(m,n)$ is computed as the DFT of $h_{p,k}(m,\ell)$ and reads

$$H_{p,k}(m,n) = \sum_{\ell=0}^{L_k-1} h_{p,k}(m,\ell)\, e^{-j2\pi\, n\, \ell/N} \qquad n = 0, 1, \ldots, N-1 \quad (6.5)$$

where $h_{p,k}(m,\ell)$ is the sample of $h_{p,k}(m,t)$ at $t = \ell T_S$ and L_k denotes the length of the kth CIR in sampling periods. Note that $L_k = \mathrm{int}\{(\tau_k^{(\max)} + T_g)/T_s\}$, where T_g is the duration of $g(t)$, $\tau_k^{(\max)} = \max_{\ell}\{\tau_{k,\,\ell}\}$ is the maximum path delay and $\mathrm{int}\{x\}$ denotes the maximum integer not exceeding x. Since $\tau_k^{(\max)}$ is usually unknown, in practice L_k is replaced by $L = \mathrm{int}\{(\tau^{(\max)} + T_g)/T_s\}$, where $\tau^{(\max)}$ is the maximum *expected* path delay among all active users.

Letting $\mathbf{H}_{p,k}(m) = [H_{p,k}(m,0),\, H_{p,k}(m,1), \ldots, H_{p,k}(m,N-1)]^T$, we may rewrite (6.5) in matrix form as

$$\mathbf{H}_{p,k}(m) = \mathbf{F}\mathbf{h}_{p,k}(m) \quad (6.6)$$

where $\mathbf{h}_{p,k}(m) = [h_{p,k}(m,0),\ h_{p,k}(m,1), \ldots, h_{p,k}(m,L-1)]^T$ is the kth CIR vector at the pth antenna and $\mathbf{F}$ is an $N \times L$ matrix with entries

$$[\mathbf{F}]_{n,\ell} = e^{-j2\pi\, n\ell/N} \qquad 0 \le n \le N-1\,, \quad 0 \le \ell \le L-1. \tag{6.7}$$

6.2 Estimation Of The Channel Responses and Noise Power

Coherent detection of the transmitted signals requires knowledge of the channel frequency responses $\{\mathbf{H}_{p,k}(m); 1 \le k \le K\}$ at each receive branch p, with $1 \le p \le P$. In addition to channel state information, the noise powers $\{\sigma_p^2;\ 1 \le p \le P\}$ are also needed for linear MMSE multiuser detection Verdú (1998).

In this Section we present two schemes for jointly estimating the channel responses and noise powers. The first is reminiscent of the method discussed in Sanguinetti *et al.* (2004) and considers the entries of $\mathbf{H}_{p,k}(m)$ as unknown independent parameters. The second scheme takes advantage of the fading correlation between adjacent subcarriers and effectively exploits the structure of $\mathbf{H}_{p,k}(m)$ in (6.6) to improve the system performance. Both schemes are based on the ML criterion and require knowledge of the transmitted data symbols. To this purpose, we assume that the MC-CDMA blocks are organized in frames and each frame is preceded by N_T training blocks. For simplicity, we neglect channel variations during the training sequence (i.e., we let $\mathbf{H}_{p,k}(m) \approx \mathbf{H}_{p,k}$ for $m = 1, 2, \ldots, N_T$) and assume that each σ_p^2 is constant over a frame. Then, the DFT output at the pth antenna during the mth block can be written as

$$\mathbf{X}_p(m) = \sum_{k=1}^{K} \mathbf{D}_k(m)\mathbf{H}_{p,k} + \mathbf{w}_p(m) \tag{6.8}$$

where $\mathbf{w}_p(m)$ is Gaussian noise with zero-mean and covariance matrix $\sigma_p^2\mathbf{I}_N$ while $\mathbf{D}_k(m) = \text{diag}\{d_k(m,0), d_k(m,1),$
$\ldots, d_k(m,N-1)\}$. The quantities $d_k(m,n)$ depend on the spreading code and training symbols $t_k^{(i)}(m)$ $(i = 0,1,\ldots,I-1)$ of the kth user through the following relation

$$d_k(m,n) = c_k(\ell+1)t_k^{(i)}(m) \qquad 0 \le n \le N-1 \tag{6.9}$$

where i represents the remainder of the ratio n/I and ℓ is the integer part of n/I. Without loss of generality, we assume that the training symbols belongs to a PSK constellation with $|\,t_k^{(i)}(m)| = 1$. Stacking the channel frequency responses of the active users into a single KN-dimensional vector $\mathbf{H}_p = \left[\mathbf{H}_{p,1}^T\,\mathbf{H}_{p,2}^T\cdots\mathbf{H}_{p,K}^T\right]^T$, we may rewrite $\mathbf{X}_p(m)$ as

$$\mathbf{X}_p(m) = \mathbf{D}(m)\mathbf{H}_p + \mathbf{w}_p(m) \tag{6.10}$$

where

$$\mathbf{D}(m) = \left[\mathbf{D}_1(m)\,\mathbf{D}_2(m)\cdots\mathbf{D}_K(m)\right]. \tag{6.11}$$

6.2.1 Unstructured Channel Estimation

In the next, we extend the method in Sanguinetti *et al.* (2004) to jointly estimate the channel responses $\mathbf{H}_p$ and the noise power σ_p^2. From (6.10) we see that the joint estimates of σ_p^2 and $\mathbf{H}_p$ based on the observations $\{\mathbf{X}_p(m); m = 1,2,\ldots,N_T\}$ are found looking for the maximum of the log-likelihood function

$$\Lambda(\tilde{\sigma}_p^2, \tilde{\mathbf{H}}_p) = -N_T N \ln(\pi\tilde{\sigma}_p^2) - \frac{1}{\tilde{\sigma}_p^2}\sum_{m=1}^{N_T}\left\|\mathbf{X}_p(m) - \mathbf{D}(m)\tilde{\mathbf{H}}_p\right\|^2 \tag{6.12}$$

with respect to the trial values $\tilde{\sigma}_p^2$ and $\tilde{\mathbf{H}}_p$. Keeping $\tilde{\sigma}_p^2$ fixed and maximizing $\Lambda(\tilde{\sigma}_p^2, \tilde{\mathbf{H}}_p)$ with respect to $\tilde{\mathbf{H}}_p$ produces

$$\hat{\mathbf{H}}_p^{(UCE)} = \mathbf{T}^{-1} \sum_{m=1}^{N_T} \mathbf{D}^H(m)\mathbf{X}_p(m) \tag{6.13}$$

with

$$\mathbf{T} = \sum_{m=1}^{N_T} \mathbf{D}^H(m)\mathbf{D}(m). \tag{6.14}$$

In Sanguinetti *et al.* (2004) it is shown that $\hat{\mathbf{H}}_p^{(UCE)}$ is unbiased and has the following MSEE

$$\frac{1}{KN}\mathrm{E}\left\{\left\| \hat{\mathbf{H}}_p^{(UCE)} - \mathbf{H}_p \right\|^2\right\} = \frac{\sigma_p^2}{KN}\mathrm{tr}\{\mathbf{T}^{-1}\}. \tag{6.15}$$

Substituting (6.13) into (6.12) and maximizing with respect to $\tilde{\sigma}_p^2$ yields

$$\hat{\sigma}_p^2 = \frac{1}{NN_T}\left[\sum_{m=1}^{N_T} \|\mathbf{X}_p(m)\|^2 - \mathbf{y}^H\mathbf{T}^{-1}\mathbf{y}\right] \tag{6.16}$$

with

$$\mathbf{y} = \sum_{m=1}^{N_T} \mathbf{D}^H(m)\mathbf{X}_p(m). \tag{6.17}$$

In Appendix C.1 it is shown that

$$E\left\{\hat{\sigma}_p^2\right\} = \frac{N_T - K}{N_T}\sigma_p^2 \tag{6.18}$$

which means that $\hat{\sigma}_p^2$ is a *biased* estimator. On the other hand, collecting (6.16) and (6.18) it is seen that an *unbiased* estimate of σ_p^2 is given by

$$\hat{\sigma}_p^{2(UCE)} = \frac{1}{N(N_T - K)}\left[\sum_{m=1}^{N_T} \|\mathbf{X}_p(m)\|^2 - \mathbf{y}^H\mathbf{T}^{-1}\mathbf{y}\right]. \tag{6.19}$$

The variance of $\hat{\sigma}_p^{2(UCE)}$ is computed in Appendix C.1 and reads

$$\text{var}\left\{\hat{\sigma}_p^{2(UCE)}\right\} = \frac{\sigma_p^4}{N\left(N_T - K\right)}. \tag{6.20}$$

Inspection of (6.13) and (6.19) indicates that both the channel and noise power estimators require the invertibility of $\mathbf{T}$. As shown in Sanguinetti *et al.* (2004), this condition is met *if and only if* $N_T \geq K$. On the other hand, from (6.19) we see that σ_p^2 cannot be estimated if $N_T = K$. This means that $N_T > K$ is a necessary condition for the *joint* estimation of the channel responses and noise power by means of UCE.

Using the same arguments of Tung *et al.* (2001) it can be shown that, for a given number N_T of training blocks, the MSEE in the RHS of (6.15) is minimized when $\mathbf{T}$ is diagonal. This can be achieved by employing WH training sequences of length $N_T = Q$. In these circumstances $\mathbf{T}$ reduces to $\mathbf{I}_{KN}$ and the estimates (6.13) and (6.19) take the form

$$\hat{\mathbf{H}}_p^{(UCE)} = \sum_{m=1}^{N_T} \mathbf{D}^H(m)\mathbf{X}_p(m), \tag{6.21}$$

$$\hat{\sigma}_p^{2(UCE)} = \frac{1}{N\left(N_T - K\right)}\left[\sum_{m=1}^{N_T} \left\|\mathbf{X}_p(m)\right\|^2 - \left\|\hat{\mathbf{H}}_p^{(UCE)}\right\|^2\right] \tag{6.22}$$

while the MSEE becomes

$$MSEE = \sigma_p^2. \tag{6.23}$$

The overall operations (real additions and products) involved in the computation of $\hat{\mathbf{H}}_p^{(UCE)}$ are shown in the first line of Table 6.1. In writing this line we have taken into account that a complex addition is equivalent to two real additions while $\mathbf{D}^H(m)\mathbf{X}_p(m)$ is obtained from $\mathbf{X}_p(m)$ by means of simple sign inversions.

Table 6.1: Computational complexity (per receive branch) of the channel acquisition schemes using WH training sequences of length $N_T = Q$

Algorithm	Real products	Real Additions
UCE	None	$2NK(N_T - 1)$
SCE	$4NK \log_2 N$	$2NK(N_T - 1 + 3\log_2 N)$
MGF	$4NK$	$6NK$

6.2.2 Structured Channel Estimation

We collect the CIRs of all active users into a KL-dimensional vector $\mathbf{h}_p = [\mathbf{h}_{p,1}^T \; \mathbf{h}_{p,2}^T \; \cdots \; \mathbf{h}_{p,K}^T]^T$. Thus, from (6.6) we see that $\mathbf{H}_p$ may be written as

$$\mathbf{H}_p = \mathbf{\Sigma}\mathbf{h}_p \tag{6.24}$$

where $\mathbf{\Sigma} = \text{diag}\{\mathbf{F}, \mathbf{F}, \ldots, \mathbf{F}\}$ is a block diagonal matrix with K identical blocks $\mathbf{F}$.

The SCE exploits the structure of $\mathbf{H}_p$ shown in (6.24). In particular, it computes an estimate $\hat{\mathbf{h}}_p$ of the CIR vector which is then exploited to obtain the corresponding estimate of $\mathbf{H}_p$ in the form

$$\hat{\mathbf{H}}_p^{(SCE)} = \mathbf{\Sigma}\hat{\mathbf{h}}_p. \tag{6.25}$$

The rationale behind this approach is that it reduces the number of unknowns since in practical applications the channel order L is much smaller than the number N of subcarriers.

Substituting (6.24) into (6.10), we see that the log-likelihood function for σ_p^2 and $\mathbf{h}_p$ takes the form

$$\Lambda(\tilde{\sigma}_p^2, \tilde{\mathbf{h}}_p) = -N_T N \ln(\pi\tilde{\sigma}_p^2) - \frac{1}{\tilde{\sigma}_p^2} \sum_{m=1}^{N_T} \left\| \mathbf{X}_p(m) - \mathbf{D}(m)\boldsymbol{\Sigma}\tilde{\mathbf{h}}_p \right\|^2. \tag{6.26}$$

The ML estimates of σ_p^2 and $\mathbf{h}_p$ are found by maximizing $\Lambda(\tilde{\sigma}_p^2, \tilde{\mathbf{h}}_p)$ with respect to $\tilde{\sigma}_p^2$ and $\tilde{\mathbf{h}}_p$. This produces

$$\hat{\mathbf{h}}_p = \mathbf{R}^{-1}\mathbf{z}, \tag{6.27}$$

$$\hat{\sigma}_p^2 = \frac{1}{NN_T} \left[\sum_{m=1}^{N_T} \|\mathbf{X}_p(m)\|^2 - \mathbf{z}^H \mathbf{R}^{-1} \mathbf{z} \right] \tag{6.28}$$

where

$$\mathbf{z} = \boldsymbol{\Sigma}^H \sum_{m=1}^{N_T} \mathbf{D}^H(m)\mathbf{X}_p(m) \tag{6.29}$$

and

$$\mathbf{R} = \boldsymbol{\Sigma}^H \left[\sum_{m=1}^{N_T} \mathbf{D}^H(m)\mathbf{D}(m) \right] \boldsymbol{\Sigma}. \tag{6.30}$$

Finally, substituting (6.27) into (6.25) and recalling (6.29), we obtain an estimate of $\mathbf{H}_p$ in the form

$$\hat{\mathbf{H}}_p^{(SCE)} = \boldsymbol{\Sigma}\mathbf{R}^{-1}\boldsymbol{\Sigma}^H \sum_{m=1}^{N_T} \mathbf{D}^H(m)\mathbf{X}_p(m). \tag{6.31}$$

The MSEE of $\hat{\mathbf{H}}_p^{(SCE)}$ is computed in Appendix C.2 and reads

$$MSEE = \frac{\sigma_p^2}{K} \mathrm{tr}\{\mathbf{R}^{-1}\}. \tag{6.32}$$

As for $\hat{\sigma}_p^2$, it can be shown that (6.28) is a biased estimate of σ_p^2 and the corresponding unbiased estimator is given by

$$\hat{\sigma}_p^{2(SCE)} = \frac{1}{N(N_T - KL/N)} \left[\sum_{m=1}^{N_T} \|\mathbf{X}_p(m)\|^2 - \mathbf{z}^H \mathbf{R}^{-1} \mathbf{z} \right]. \tag{6.33}$$

The variance of $\hat{\sigma}_p^{2(SCE)}$ can be computed using the same arguments of Appendix C.1 and reads

$$\text{var}\left\{\hat{\sigma}_p^{2(SCE)}\right\} = \frac{\sigma_p^4}{N\left(N_T - KL/N\right)}. \tag{6.34}$$

The following remarks are of interest.

1) Since $\mathbf{D}(m)\boldsymbol{\Sigma}$ has dimensions $N \times KL$, it is straightforward to see that $\text{rank}\{\mathbf{D}(m)\boldsymbol{\Sigma}\} = \text{rank}\{\boldsymbol{\Sigma}^H \mathbf{D}^H(m)\mathbf{D}(m)\boldsymbol{\Sigma}\} \leq \min\{N, KL\}$. On the other hand, using the property $\text{rank}[\mathbf{A} + \mathbf{B}] \leq \text{rank}[\mathbf{A}] + \text{rank}[\mathbf{B}]$, from (6.30) we see that $\text{rank}[\mathbf{R}] \leq N_T \cdot \min\{N, KL\}$. Observing that $\mathbf{R}$ is a matrix of order KL, it follows that a *necessary* condition for the existence of $\mathbf{R}^{-1}$ is $N_T \geq \lfloor KL/N \rfloor$, where $\lfloor x \rfloor$ is the smallest integer greater than or equal to x. Finally, from (6.33) we see that σ_p^2 can be estimated only if $N_T > KL/N$. Bearing in mind the above facts, it turns out that SCE operates correctly provided that

$$N_T > \lfloor KL/N \rfloor. \tag{6.35}$$

Comparing (6.35) with the result $N_T > K$ obtained with UCE, we see that SCE leads to a potential reduction of the training overhead by a factor N/L. In particular, one training block is enough when $N > KL$.

2) Inspection of (6.32) reveals that, for a given N_T, the minimum MSEE is achieved when $\mathbf{R}$ is diagonal Tung *et al.* (2001). Collecting (6.11) and (6.30), we see that this condition implies

$$\mathbf{F}^H \sum\nolimits_{m=1}^{N_T} \mathbf{D}_k^H(m)\mathbf{D}_\ell(m)\,\mathbf{F} = \mathbf{0} \qquad k \neq \ell. \tag{6.36}$$

In the above circumstances, $\mathbf{R}$ reduces to $(N_T N/Q)\,\mathbf{I}_{KL}$ and from (6.32) it follows that the minimum MSEE is given by

$$MSEE_{\min} = \frac{\sigma_p^2 QL}{N_T N}. \tag{6.37}$$

3) Optimal training sequences satisfying the set of constraints (6.36) can be easily designed when $N = QL$. In these circumstances we see from (6.32) that the number of training blocks can be limited to $N_T = 1$ and conditions (6.36) are met by

$$d_k(1,n) = \frac{1}{\sqrt{Q}} \exp[-j2\pi n(k-1)/Q] \qquad k = 1, 2, \ldots, K. \tag{6.38}$$

The corresponding training sequences are next derived substituting (6.38) into (6.9) while the MSEE is computed setting $N = QL$ into (6.37) and we have

$$MSEE_{\min} = \sigma_p^2. \tag{6.39}$$

Comparing with (6.23), we see that SCE achieves the same accuracy of UCE but reduces the training overhead N_T from Q to 1.

An interesting interpretation of the sequences $\{d_k(1,n)\}$ in (6.38) is possible if we consider the corresponding time-domain samples, which are computed as the N-point IDFT of $\{d_k(1,n)\}$ (multiplied by $\sqrt{N}$)

$$t_k(n) = \frac{1}{\sqrt{N}} \sum_{\ell=0}^{N-1} d_k(1,\ell)\, e^{j2\pi n\ell/N}, \qquad 0 \le n \le N-1. \tag{6.40}$$

Substituting (6.38) into (6.40) and setting $N = QL$, yields

$$t_k(n) = \begin{cases} \sqrt{L} & \text{for } n = (k-1)L + L_{GPF} \\ 0 & \text{otherwise}. \end{cases} \tag{6.41}$$

The above equation indicates that each user transmits a *single* time-domain sample during the training block. Also, samples of different users are spaced apart

by multiples of $L' = 2L_{GPF} + 1$. Bearing in mind that the CIR has duration L (i.e., $h(\ell) = 0$ for $\ell > L$), we see that signals transmitted by different users are received at the BS over adjacent and non-overlapping time intervals $\Delta_k = [(k-1)L'T_s, kL'T_s)$, with $k = 1, 2, \ldots, K$. This means that there is no interference among the active users as each received sample at the input of the DFT unit is contributed by one user only. On the other hand, transmitting a single time-domain sample of power L as indicated in (6.41) is disadvantageous since it may produce an excessive PAPR at the mobile unit.

4) Employing WH training sequences of length Q is an alternative method to meet (6.36). The corresponding MSEE is computed setting $N_T = Q$ into (6.37) and reads

$$MSEE = \frac{\sigma_p^2 L}{N} \tag{6.42}$$

while the estimates (6.31) and (6.33) reduce to

$$\hat{\mathbf{H}}_p^{(SCE)} = \boldsymbol{\Sigma}\boldsymbol{\Sigma}^H \hat{\mathbf{H}}_p^{(UCE)}, \tag{6.43}$$

$$\hat{\sigma}_p^{2(SCE)} = \frac{1}{N(N_T - KL/N)} \left[\sum_{m=1}^{N_T} \|\mathbf{X}_p(m)\|^2 - \left\| \hat{\mathbf{H}}_p^{(SCE)} \right\|^2 \right] \tag{6.44}$$

with $\hat{\mathbf{H}}_p^{(UCE)}$ as given in (6.21). Comparing (6.42) with (6.23) we see that, for the same number of training blocks, the MSEE is reduced by a factor N/L with respect to UCE.

5) The overall operations involved in the computation of $\hat{\mathbf{H}}_p^{(SCE)}$ in (6.43) are summarized in the second line of Table 6.1. In writing this line we have born in mind that $\hat{\mathbf{H}}_p^{(SCE)}$ and $\hat{\mathbf{H}}_p^{(UCE)}$ can be segmented into K subvectors $\hat{\mathbf{H}}_{p,k}^{(SCE)}$ and $\hat{\mathbf{H}}_{p,k}^{(UCE)}$ $(k = 1, 2, \ldots, K)$ which are related by $\hat{\mathbf{H}}_{p,k}^{(SCE)} = \mathbf{F}\mathbf{F}^H \hat{\mathbf{H}}_{p,k}^{(UCE)}$. Each $\hat{\mathbf{H}}_{p,k}^{(SCE)}$ is then efficiently computed by feeding $\hat{\mathbf{H}}_{p,k}^{(UCE)}$ to an N-point IDFT unit,

setting to zero the last $N - L$ outputs and finally passing the resulting vector to an N-point DFT unit.

6) A possible shortcoming of SCE is that all groups of subcarriers must be assigned to the same set of users, so that the maximum number of contemporarily active users is limited to $K_{MAX} = Q$. On the other hand, from (6.13) and (6.21) it can be seen that UCE operates independently on each group and, therefore, it is also suited for more versatile systems in which different groups of subcarriers are assigned to different groups of users. In these circumstances, the number of contemporarily active users can be as large as the number N of subcarriers.

6.3 Tracking Time-Varying Channels

The batch estimators derived previuosly are used at the beginning of each frame for initial acquisition of the channel parameters. However, in a time-varying environment the channel estimates must be continuously updated during the data section of the frame. A solution to this problem is now discussed by resorting to the LMS algorithm and exploiting both the unstructured and structured approaches. The resulting schemes operate in a DD mode using data decisions provided by a multiuser detector.

6.3.1 Least Mean Squares UCE

The LMS-UCE employs the LMS algorithm to minimize the following mean square error

$$J_{LMS-UCE}(\tilde{\mathbf{H}}_p(m)) = \mathrm{E}\left\{\left\|\mathbf{X}_p(m) - \hat{\mathbf{D}}(m)\tilde{\mathbf{H}}_p(m)\right\|^2\right\} \qquad p = 1, 2, \ldots, P \tag{6.45}$$

with respect to $\tilde{\mathbf{H}}_p(m)$, where $\hat{\mathbf{D}}(m)$ is an estimate of $\mathbf{D}(m)$ obtained by replacing $\{t_k^{(i)}(m)\}$ in (6.9) with the decisions $\{\hat{a}_k^{(i)}(m)\}$. The minimum in (6.45) is achieved for $\tilde{\mathbf{H}}_p(m) = \hat{\mathbf{H}}_p(m)$, where $\hat{\mathbf{H}}_p(m)$ is the estimate of $\mathbf{H}_p(m)$ and satisfies the following recursion

$$\hat{\mathbf{H}}_p(m+1) = \hat{\mathbf{H}}_p(m) + \mu \mathbf{e}_p(m) \qquad p = 1, 2, \ldots, P. \tag{6.46}$$

In the above equation, $\mathbf{e}_p(m)$ is defined as

$$\mathbf{e}_p(m) = \hat{\mathbf{D}}^H(m) \left[\mathbf{X}_p(m) - \hat{\mathbf{D}}(m) \hat{\mathbf{H}}_p(m) \right] \tag{6.47}$$

while μ is the *step-size*, which controls the convergence properties of the algorithm and it is chosen so as to achieve a reasonable trade-off between steady-state performance and tracking capabilities.

In Appendix C.3 we evaluate the performance of LMS-UCE over a static channel (i.e., $\mathbf{H}_p(m) = \mathbf{H}_p$) assuming statistically independent data symbols with zero-mean and variance A_2. It is found that $\hat{\mathbf{H}}_p(m)$ is unbiased and has the following MSEE

$$\frac{1}{KN} \mathrm{E} \left\{ \left\| \hat{\mathbf{H}}_p(m) - \mathbf{H}_p \right\|^2 \right\} = \frac{2Q B_L T_B}{A_2} \sigma_p^2 \tag{6.48}$$

where $B_L T_B = \mu A_2 / [2Q(2 - \mu A_2 / Q)]$ is the noise equivalent bandwidth Mengali & A.N.D'Andrea (1997) of the recursion (6.46), normalized to $1/T_B$.

The complexity of LMS-UCE can be assessed as follows. Evaluating $\mathbf{e}_p(m)$ in (6.47) requires $2KN$ complex products plus NK additions. Finally, updating the channel estimates in the RHS of (6.46) needs KN more complex additions. The overall operations *per receive branch* are summarized in the first line of Table 6.2, where we have born in mind that a complex multiplication corresponds to four real products plus two real additions. Note that a significant reduction of complexity is achieved when the data symbols are taken from a BPSK or QPSK

Table 6.2: Computational complexity (per receive branch) of the channel tracking schemes

Algorithm	Real products	Real Additions
LMS-UCE	$8NK$	$8NK$
LMS-UCE (BPSK or QPSK)	None	$4NK$
LMS-SCE	$8NK(1+\frac{1}{2}\log_2 N)$	$8NK(1+\frac{3}{4}\log_2 N)$
LMS-SCE (BPSK or QPSK)	$4NK\log_2 N)$	$4NK(1+\frac{3}{2}\log_2 N)$

constellations. In this case the multiplications by $\hat{\mathbf{D}}(m)$ and $\hat{\mathbf{D}}^H(m)$ in (6.47) reduce to simple sign inversions and/or swaps between the real and imaginary parts of the involved quantities. The resulting operations are shown in the second line of Table 6.2.

6.3.2 Least Mean Squares SCE

The LMS-SCE provides an estimate of $\mathbf{h}_p(m)$ by minimizing the quantity

$$J_{LMS-SCE}(\tilde{\mathbf{h}}_p(m)) = \mathrm{E}\left\{\left\|\mathbf{X}_p(m) - \hat{\mathbf{D}}(m)\boldsymbol{\Sigma}\tilde{\mathbf{h}}_p(m)\right\|^2\right\} \qquad p = 1, 2, \ldots, P \tag{6.49}$$

with respect to $\tilde{\mathbf{h}}_p(m)$. This leads to the following recursion

$$\hat{\mathbf{h}}_p(m+1) = \hat{\mathbf{h}}_p(m) + \mu\boldsymbol{\Sigma}^H\mathbf{e}_p(m) \qquad p = 1, 2, \ldots, P \tag{6.50}$$

where $\mathbf{e}_p(m)$ is still given in (6.47) and $\hat{\mathbf{h}}_p(m)$ is the CIR estimate at the mth step. Pre-multiplying both sides of (6.50) by $\boldsymbol{\Sigma}$ and bearing in mind (6.25) yields

$$\hat{\mathbf{H}}_p(m+1) = \hat{\mathbf{H}}_p(m) + \mu\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H\mathbf{e}_p(m) \qquad p = 1, 2, \ldots, P. \tag{6.51}$$

The performance of LMS-SCE over a static channel can be computed with the same arguments of Appendix C.3. It is found that $\hat{\mathbf{H}}_p(m)$ is unbiased and has the following MSEE

$$\frac{1}{KN}\mathrm{E}\left\{\left\|\hat{\mathbf{H}}_p(m)-\mathbf{H}_p\right\|^2\right\}=\frac{2LQB_LT_B}{NA_2}\sigma_p^2. \tag{6.52}$$

Note that the only difference between LMS-SCE and LMS-UCE is the presence of the matrix $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H$ in (6.51), which performs a better noise filtering by taking into account that $\mathbf{h}_p(m)$ has duration $L < N$. This translates into a reduction of the MSEE by a factor N/L, as is seen by comparing (6.52) with (6.48). For $L = N$ the LMS-SCE boils down to LMS-UCE since in this case $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H = \mathbf{I}_N$.

In assessing the complexity of LMS-SCE, we observe that $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H\mathbf{e}_p(m)$ can be segmented into K subvectors $\mathbf{F}\mathbf{F}^H\mathbf{e}_{p,k}(m)$ $(k = 1, 2, \ldots, K)$, with $\mathbf{e}_{p,k}(m) = \hat{\mathbf{D}}_k^H(m)\left[\mathbf{X}_p(m) - \hat{\mathbf{D}}(m)\hat{\mathbf{H}}_p(m)\right]$. Each subvector may be efficiently computed by feeding $\mathbf{e}_{p,k}(m)$ to an N-point IDFT unit, setting to zero the last $N - L$ outputs and passing the resulting vector to an N-point DFT unit. The overall operations per receive branch are shown in the third line of Table 6.2 for general PSK/QAM constellations and in the fourth line for BPSK/QPSK data symbols. Compared with LMS-UCE, we see that the system complexity increases from $O(NK)$ to $O(NK\log_2 N)$.

6.4 Numerical Results

System Parameters

The transmitted symbols belong to a QPSK constellation and are obtained from the information bits through a Gray map. The total number of subcarriers is N=128 and Walsh-Hadamard codes of length Q=8 are used for spreading purposes. The signal bandwidth is $B = 8$ MHz, so that the sampling period is equal

to $T_S = 0.125$ μs. The carrier frequency is $f_0 = 5$ GHz (corresponding to a wavelength $\lambda = 15$ cm) and the inter-element spacing in the antenna array is $\delta = 2\lambda$.

The channel impulse responses $h_{p,k}(m,t)$ are generated as indicated in (6.3) with eight paths ($N_p = 8$). The modulation pulse $g(t)$ has a raised cosine Fourier transform with roll-off 0.35 and duration $T_g = 8T_S$. We assume that all the CIRs $\mathbf{h}_{p,k}(m)$ have the same length L and the maximum path delay $\tau^{(\max)}$ is fixed to $8T_S = 1$ μs. Recalling that $L = \text{int}\{(\tau^{(\max)} + T_g)/T_S\}$, this corresponds to a CIR length of $L = 16$. To avoid inter-block interference, the length N_G of the cyclic prefix is set equal to L. This means that each MC-CDMA block has length $T_B = (N + N_G)T_S = 18$ μs (including the cyclic prefix). Each frame has 64 blocks, corresponding to a duration of 1.15 ms and is preceded by N_T training blocks which are exploited for channel acquisition and noise power estimation.

A new set of path delays, complex gains and DOAs are randomly generated at the start of each frame. The path delays and DOAs are uniformly distributed within $[0,1]$ μs and $[-60°, 60°]$, respectively, and are kept fixed within the frame. The path gains $\{\alpha_{k,r}(m)\}$ have powers

$$\sigma_r^2 = \beta \cdot \exp(-r/5) \qquad r = 0, 1, \ldots, 7 \tag{6.53}$$

where β is chosen such that the average energy of the CIR is normalized to unity, i.e., $E\{\|\mathbf{h}_{p,k}\|^2\} = 1$. Each $\alpha_{k,r}(m)$ is generated by filtering statistically independent white Gaussian processes in a third-order low-pass Butterworth filter. The 3-dB bandwidth of the filter is taken as a measure of the Doppler rate $f_D = f_0 v/c$, where v denotes the mobile speed and $c = 3 \times 10^8$ m/s is the speed of light. The mobile velocity and the number of diversity branches are varied throughout simulations to assess their impact on the system performance.

A simulation run begins with channel acquisition and noise power estimation, which are performed by means of UCE or SCE exploiting WH training sequences

of length $N_T = 8$. For the sake of simplicity, we assume that σ_p^2 is the same at each diversity branch (i.e., we let $\sigma_p^2 = \sigma^2$ for $p = 1, 2, \ldots, P$). The LMS algorithm is next employed to update the channel estimates during the data section of the frame. The optimal selection of the step-size μ depends on the mobile speed v and the number of users K. Simulations indicate that a good choice for $30 \leq v \leq 90$ km/h and $1 \leq K \leq 8$ is $\mu = 0.9$ for LMS-UCE and $\mu = 8 \times 10^{-3}$ for LMS-SCE. Data detection is performed by means of the PPIC receiver. As already discussed, the latter is a multistage detector in which MAI is estimated using tentative data decisions and *partially* subtracted from the DFT output. As the tentative decisions of the earlier stages are less reliable than those of the later stages, the fraction of MAI that is canceled out increases for successive stages. This results in better performance with respect to the conventional (brute-force) PIC receiver, which tries to subtract all the MAI at each iteration stage. As it occurs with all IC-based schemes, the PPIC needs an initial estimate of the transmitted data. Unless otherwise specified, the latter is taken as the output of a linear MMSE multiuser detector Verdú (1998).

Performance Assessment

We begin by assessing the estimation accuracy of the channel responses and noise power during the acquisition phase. Figure 6.3 illustrates the channel MSEE vs. $1/\sigma^2$ as obtained with UCE, SCE and a third scheme proposed by Marques, Gameiro and Fernandes (MGF) in Marques *et al.* (2003). Marks indicate simulation while solid lines represent theoretical analysis only for UCE and SCE. The system is fully-loaded ($K = 8$) and the receiver is equipped with a single antenna ($P = 1$). Similar results are obtained with different values of K and P since, as predicted by (6.23) and (6.42), the channel MSEE is independent of these parameters. In the MGF scheme, each user transmits a given number of pilot tones,

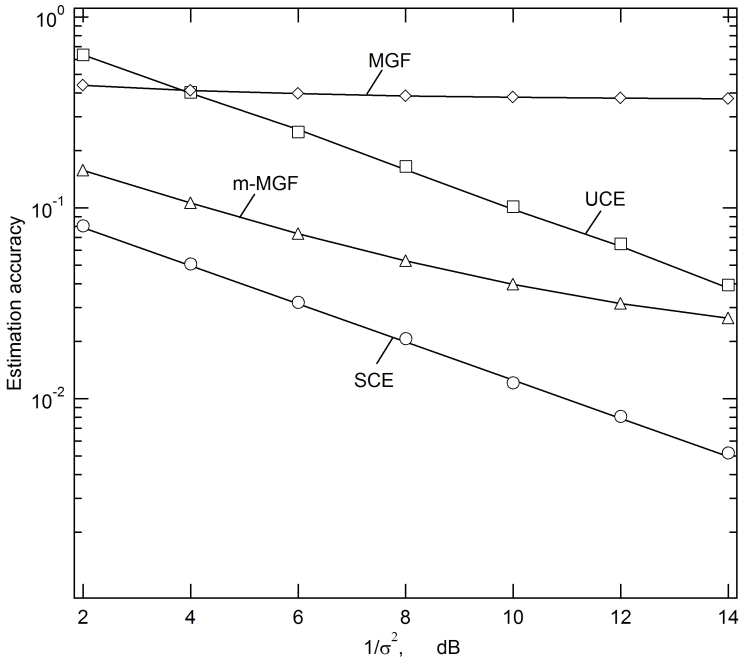

Figure 6.3: Accuracy of the channel acquisition schemes vs. $1/\sigma^2$ with $P = 1$ and $K = 8$.

say N_{pilot} , over each training block using exclusively assigned subcarriers so as to eliminate the MAI. Channel estimation is then performed through a second-order interpolation among pilots. Note that the maximum number of different channels that MGF can estimate (or, equivalently, the maximum number of users the system can support) is limited to $K_{MAX} = N/N_{pilot}$. Therefore, N_{pilot} results from a trade-off between contrasting requirements. It must be large enough to reduce interpolation errors (in practice, the pilot spacing should not exceed the channel coherence bandwidth), but it must be sufficiently small to guarantee adequate multiple-access capabilities. In our simulation we set $N_{pilot} = 16$, so that $K_{MAX} = 8$. We see that SCE performs remarkably better than UCE and achieves a gain of 9 dB in terms of estimation accuracy. As for MGF, its accuracy is rather poor at all SNR values. The reason is that the pilot spacing with $N_{pilot} = 16$ is

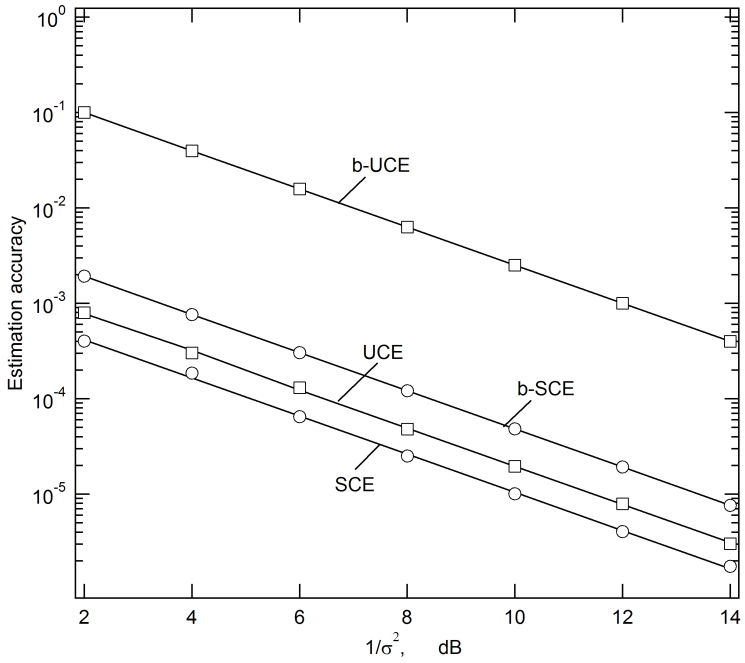

Figure 6.4: Accuracy of the noise power estimators vs. $1/\sigma^2$ with $P = 1$ and $K = 4$.

comparable to the channel coherence bandwidth and results in large interpolation errors. To mitigate this problem, we resort to a slight modification of the MGF scheme (m-MGF) in which the training period is divided into two adjacent segments, each comprising four training blocks. The first segment is allocated to a group of four users, each of them transmitting 32 pilots per block. The second segment is used by the remaining four users in a similar way. Although a total of 128 pilots per user are available with both MGF and m-MGF, the latter is expected to be more effective against interpolation errors thanks to its reduced pilot spacing (in the frequency domain). The results in Figure 6.3 confirm this intuition and show that m-MGF achieves significant improvements with respect to both MGF and UCE. However, it is still inferior to SCE.

Figure 6.4 shows the MSEE of the noise power estimate vs. $1/\sigma^2$ as obtained

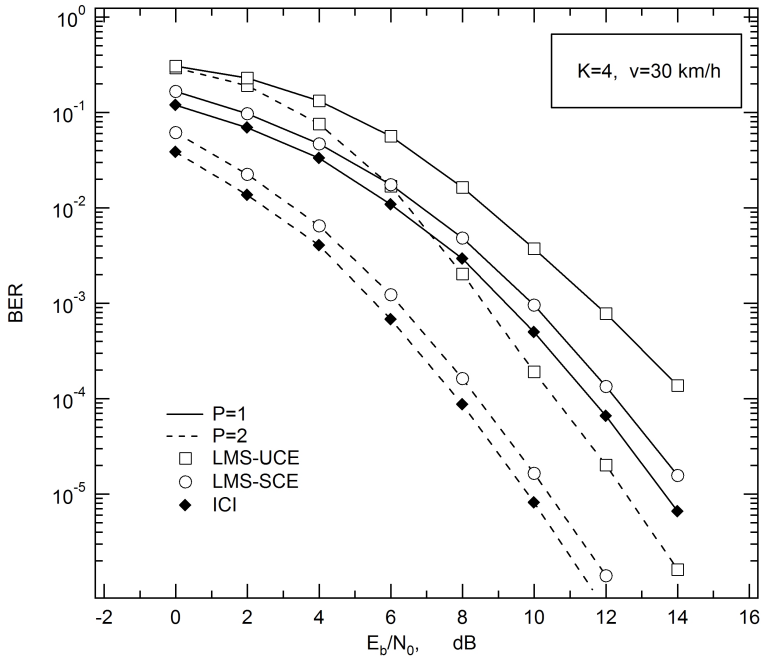

Figure 6.5: BER vs. E_b/N_0 with $K = 4$, $v = 30$ km/h and $P = 1$.

with UCE, SCE and the biased estimators (6.16) and (6.28) (denoted b-UCE and b-SCE, respectively) in the presence of four users. Marks indicate simulation results while solid lines represent theoretical analysis. Good agreement is observed between simulations and theory. As expected, SCE gives the best results while the biased estimators perform rather poorly. For this reason, they are not considered further in the sequel.

Figure 6.5 illustrates the BER of a PPIC receiver endowed with the proposed channel and noise power estimators vs. E_b/N_0, where E_b is the average received energy per bit and $N_0/2$ is the two-sided noise power spectral density. Marks indicate simulation results while solid lines are only drawn to improve the reading of the graphs (the same occurs with all the BER curves shown in the subsequent figures). The number of iterations of the PPIC has been fixed to three while the

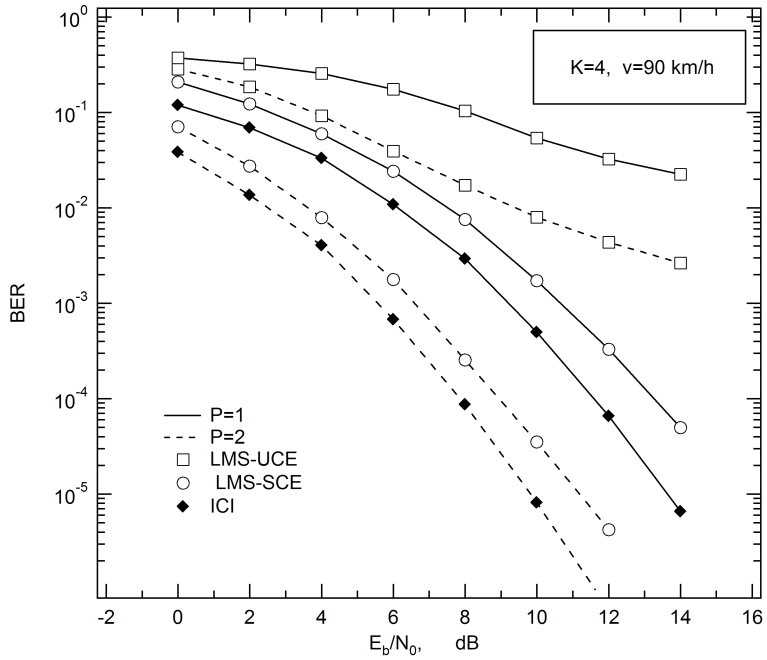

Figure 6.6: BER vs. E_b/N_0 with $K = 4$, $v = 90$ km/h and $P = 1$.

number of receive antennas is either $P = 1$ or $P = 2$. The system is half-loaded ($K = 4$) and the mobile velocity is $v = 30$ km/h (corresponding to $f_D = 140$ Hz). The curves labelled ICI correspond to perfect knowledge of the channel responses and noise power, and serve as benchmarks. We see that LMS-SCE has the best performance. In particular, for an error probability of 10^{-3}, the gain with respect to LMS-UCE is approximately 2 dB for $P = 1$ and 2.5 dB for $P = 2$. Compared to the ICI curves, LMS-SCE looses less than 1 dB when $P = 1$ and 0.6 dB in case of a two-branch receiver. Significant performance improvements are observed with both schemes as the number of receive antennas increases.

The results of Figure 6.6 have been obtained in the same operating conditions of Figure 6.5, except that the mobile speed is now $v = 90$ km/h. Note that the simulation results with ICI do not depend on the fading rate as the channel

is assumed static over each block. It is seen that LMS-SCE is still superior to LMS-UCE, which exhibits a floor at high SNR values due to its poor tracking capabilities.

Figure 6.7 illustrates the BER of the system vs. K for $E_b/N_0 = 10$ dB. The number of receive antennas is either $P = 1$ or $P = 2$ and the mobile speed is 30 km/h. As discussed previously, UCE cannot estimate the noise power when $K = 8$ since in that case we have $N_T = K$. Increasing N_T represents a possible solution to this problem but it may entail an excessive overhead. For this reason, when $K = 8$ the PPIC receiver (endowed with LMS-UCE) is initialized with data decisions provided by a decorrelating detector Verdú (1998). The latter has worse performance than the linear MMSE detector but dispenses from knowledge of the noise power. As it is seen, the system performance deteriorates as K increases and LMS-SCE is always superior to LMS-UCE.

The impact of channel variations on the system performance is addressed in Figure 6.8, where the BER of a half-loaded system is shown as a function of the mobile speed v for E_b/N_0=10 dB and $P = 1$ or 2. We see that the BER degrades as v increases and the best results are obtained with LMS-SCE. In this case, the performance loss is negligible for mobile speeds up to 90 km/h.

Figure 6.9 illustrates the total operations (additions plus multiplications) of the various channel acquisition schemes as a function of the number N of subcarriers. The results are normalized to the number of receive antennas and active users, and are taken from Table 6.1 by dividing the corresponding quantities by K. It turns out that MGF and UCE have comparable complexity, while SCE entails a significant increase of the computational burden. However, it should be borne in mind that the extra-complexity of SCE is mainly due to the IDFT and DFT operations required in (6.43) to compute $\hat{\mathbf{H}}_p^{(SCE)}$ starting from $\hat{\mathbf{H}}_p^{(UCE)}$. This may not represent a serious problem in practice since IDFT and DFT units are

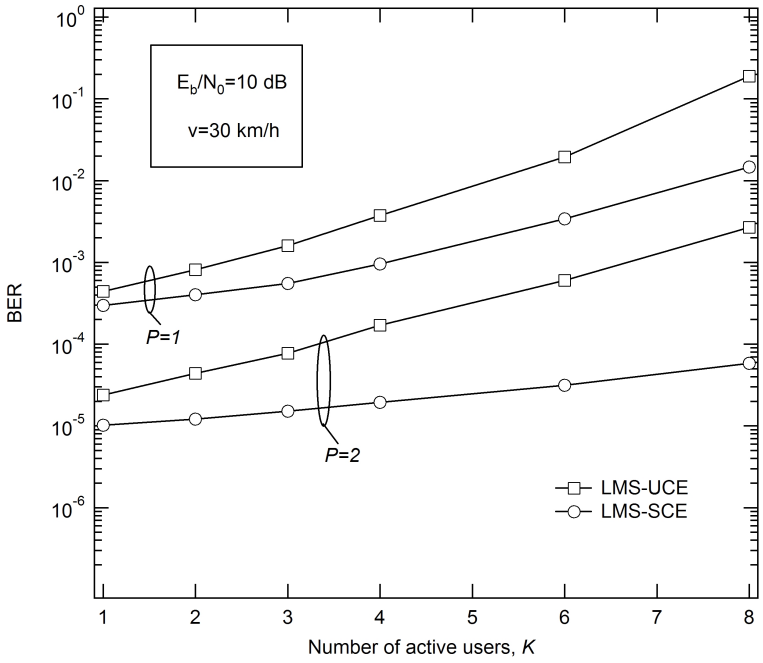

Figure 6.7: BER vs. K with E_b/N_0=10 dB , $v = 30$ km/h and P=1 or 2.

anyhow present in the receiver and, therefore, no extra circuitry is needed with respect to UCE. Vice versa, channel estimation by means of MGF requires a second-order interpolation between pilots, which may need a dedicated circuitry.

Conclusions

In this Chapter, we have addressed the problem of channel acquisition and tracking in MC-CDMA uplink transmissions. Channel tracking is pursued by means of the LMS algorithm while channel acquisition and noise power estimation are performed using two different schemes based on the ML criterion. The first (UCE) assumes independently faded subcarriers while the second (SCE) exploits a structured approach to reduce the number of unknown parameters. It is shown that SCE has some potential benefits over UCE. For a given overhead, it provides a

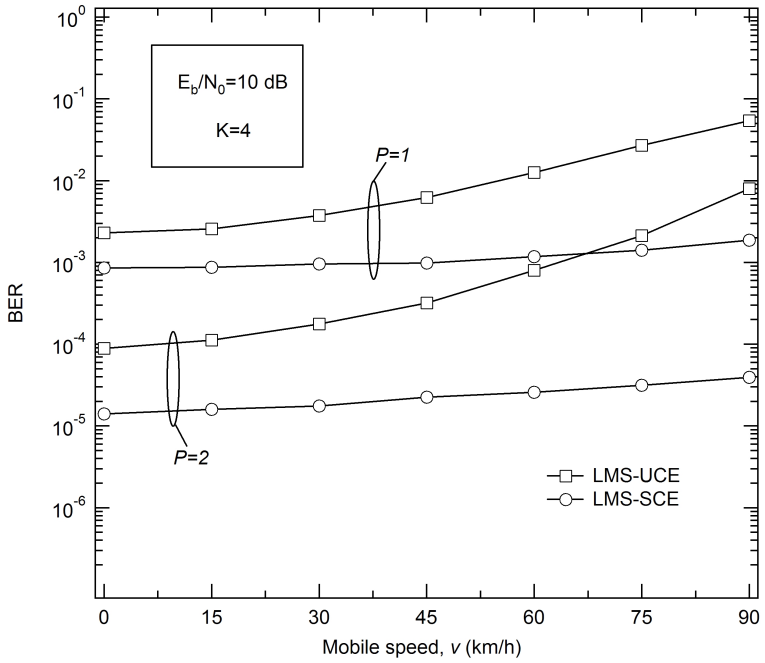

Figure 6.8: BER vs. v with E_b/N_0=10 dB, $K = 4$ and P=1 or 2.

better estimation accuracy whereas for a given MSEE it leads to a reduction of the training overhead. On the other hand, UCE is simpler to implement and can also be used in applications where different groups of subcarriers are assigned to different groups of users so as to guarantee multimedia services with different data rates.

The performance of a PPIC receiver endowed with the proposed channel estimation schemes has been investigated by simulation. The best results are provided by LMS-SCE, which can handle mobile speeds up to 90 km/h with a negligible performance loss. Comparisons have been made with other existing schemes in terms of both estimation accuracy and system complexity.

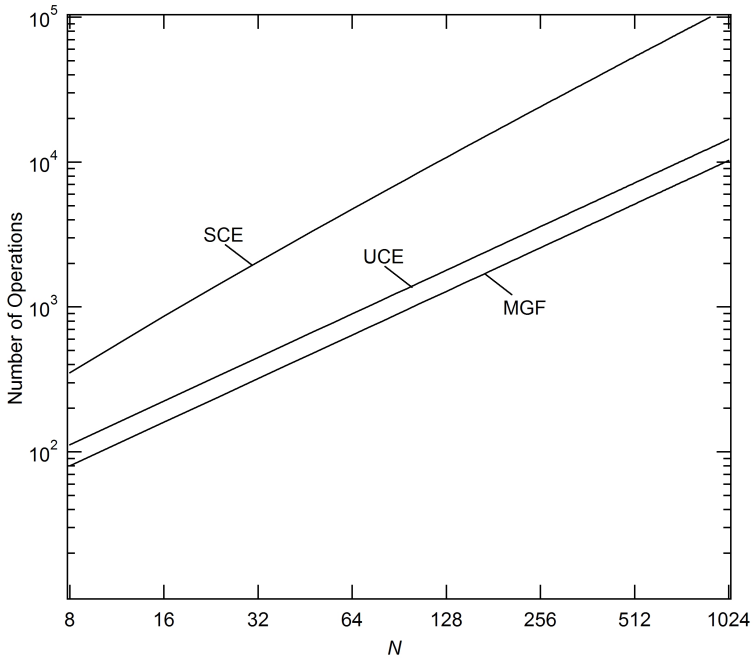

Figure 6.9: Complexity of the channel acquisition schemes per user and receive branch.

Appendix A

Designing pre-filtering schemes according to the MMSE criterion

A.1 Problem Formulation

In this Appendix we discuss three different methods that can be used in designing pre-filtering algorithms according to the MMSE criterion. To this purpose, we assume that the decision statistic of the generic mth MT takes the form of (4.8), i.e.,

$$y_m = e \cdot \mathbf{g}_m^H \mathbf{U}\mathbf{a} + e \cdot w_m \tag{A.1}$$

where $\mathbf{g}_m = \mathbf{H}_m^H \mathbf{q}_m$ is an QN_T-dimensional vector that depends on the channel coefficients and data detection strategy, $\mathbf{a}$ is the data vector and $\mathbf{U}$ is the forward matrix of dimensions $N_T Q \times Q$ which must be properly designed. In addition, w_m is a Gaussian random variable with zero-mean and variance σ^2 while the scalar e can be thought of as being part of an automatic gain control which does not impair the signal-to-noise ratio at the receiver.

Stacking the decision statistics of all users into a single vector $\mathbf{y} = [y_1, y_2, ..., y_K]^T$,

we get

$$\mathbf{y} = e \cdot \mathbf{G}^H \mathbf{U}\mathbf{a} + e \cdot \mathbf{w} \tag{A.2}$$

where $\mathbf{w} = [w_1, w_2, ..., w_K]^T$ is a Gaussian vector with zero-mean and covariance matrix $\sigma^2 \mathbf{I}_Q$ whereas $\mathbf{Q}$ is the following matrix with dimensions $N_T Q \times K$

$$\mathbf{G} = \begin{bmatrix} \mathbf{H}_{1,1}^H \mathbf{q}_1 & \mathbf{H}_{2,1}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,1}^H \mathbf{q}_K \\ \mathbf{H}_{1,2}^H \mathbf{q}_1 & \mathbf{H}_{2,2}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,2}^H \mathbf{q}_K \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{1,N_T}^H \mathbf{q}_1 & \mathbf{H}_{2,N_T}^H \mathbf{q}_2 & \cdots & \mathbf{H}_{K,N_T}^H \mathbf{q}_K \end{bmatrix}. \tag{A.3}$$

The objective of this Appendix is to provide some viable methods to find the forward matrix $\mathbf{U}$ that minimizes the sum of the mean square errors at all MTs, i.e.,

$$J = E\left\{ \ \|\mathbf{y} - \mathbf{a}\|^2 \right\}. \tag{A.4}$$

Assuming that the data vector $\mathbf{a}$ has zero mean and covariance matrix $\sigma_a^2 \mathbf{I}_K$ and is statistically independent from $\mathbf{w}$, the above equation can also be rewritten as

$$J = \sigma_a^2 \cdot \mathrm{tr}\left\{ \left(e\mathbf{G}^H \mathbf{U} - \mathbf{I}_K \right) \left(e\mathbf{G}^H \mathbf{U} - \mathbf{I}_K \right)^H \right\} + \sigma^2 e^2 \cdot K. \tag{A.5}$$

To maintain the same power as in the case where no pre-filtering is used, in minimizing the RHS of (A.5) $\mathbf{U}$ must fulfill the following identity

$$\mathrm{tr}\left\{ \mathbf{U}^H \mathbf{U} \right\} = K. \tag{A.6}$$

A.2 Optimization strategies

A.2.1 Conventional Method

The first method that we propose follows the line of reasoning of other conventional approaches based on the method of Lagrange multipliers. To this end, we assume $e = 1$ and look for the minimum of the following augmented function

$$J_1 = \sigma_a^2 \cdot \mathrm{tr}\left\{\left(\mathbf{G}^H\mathbf{U} - \mathbf{I}_K\right)\left(\mathbf{G}^H\mathbf{U} - \mathbf{I}_K\right)^H\right\} + \sigma^2 \cdot K + \mu' \cdot \mathrm{tr}\left\{\mathbf{U}^H\mathbf{U}\right\} \tag{A.7}$$

where μ' is a Lagrangian multiplier. Taking the derivative of J_1 with respect to $\mathbf{U}$ and setting it to zero yields

$$\left(\sigma_a^2\mathbf{G}\mathbf{G}^H + \mu'\mathbf{I}_{QN_T}\right)\mathbf{U} = \sigma_a^2\mathbf{G} \tag{A.8}$$

or, equivalently,

$$\mathbf{U} = \left(\mathbf{G}\mathbf{G}^H + \mu\mathbf{I}_{QN_T}\right)^{-1}\mathbf{G} \tag{A.9}$$

where we have defined $\mu = \mu'/\sigma_a^2$. Using the MIL, equation (A.9) can be also written as

$$\mathbf{U} = \mathbf{G}^H\left(\mathbf{G}^H\mathbf{G} + \mu\mathbf{I}_K\right)^{-1}. \tag{A.10}$$

At this stage we must define the parameter μ so as to meet the power constraint (A.6). To this end, we substitute (A.10) into (A.6) to obtain

$$\mathrm{tr}\left\{\left(\mathbf{G}^H\mathbf{G} + \mu\mathbf{I}_K\right)^{-2}\mathbf{G}^H\mathbf{G}\right\} = K. \tag{A.11}$$

Unfortunately, the problem of finding μ from the identity (A.11) is too complex and, to the author knowledge, it has no closed-form solution. For this reason, in the next we discuss two interesting alternatives.

A.2.2 Alternative Methods

Inspections of (A.5) reveals that the cost function J depends on the pre-coding matrix $\mathbf{U}$ and the real parameter e. In the next, we aim at designing both rather than $\mathbf{U}$ only.

The following method was proposed in Kusume *et al.* (2005) for MIMO single-user transmissions and is still based on the method of Lagrange multipliers. The solution to the problem is found looking for the minimum of the following augmented cost function

$$J_2 = \sigma_a^2 \cdot \mathrm{tr}\left\{\left(e\mathbf{G}^H\mathbf{U} - \mathbf{I}_K\right)\left(e\mathbf{G}^H\mathbf{U} - \mathbf{I}_K\right)^H\right\} + \sigma^2 e^2 \cdot K + \mu' \cdot \mathrm{tr}\left\{\mathbf{U}^H\mathbf{U}\right\} \quad \text{(A.12)}$$

where μ' is still the Lagrange multiplier. Taking the derivatives with respect to $\mathbf{U}$ and e and setting them to zero yields

$$e\left(\mathbf{G}\mathbf{G}^H + \xi\mathbf{I}_{QN_T}\right)\mathbf{U} = \mathbf{G} \quad \text{(A.13)}$$

and

$$2e \cdot \mathrm{tr}\left\{\mathbf{G}^H\mathbf{U}\mathbf{U}^H\mathbf{G} + \rho\mathbf{I}_K\right\} = \mathrm{tr}\left\{\mathbf{G}^H\mathbf{U} + \mathbf{U}^H\mathbf{G}\right\} \quad \text{(A.14)}$$

where we have defined $\xi = \mu'/(e^2\sigma_a^2)$ and $\rho = \sigma^2/\sigma_a^2$. From (A.13) we obtain

$$\mathbf{U} = \frac{1}{e}\left(\mathbf{G}\mathbf{G}^H + \xi\mathbf{I}_{QN_T}\right)^{-1}\mathbf{G} \quad \text{(A.15)}$$

or equivalently, applying the MIL,

$$\mathbf{U} = \frac{1}{e}\mathbf{G}\left(\mathbf{G}^H\mathbf{G} + \xi\mathbf{I}_K\right)^{-1}. \quad \text{(A.16)}$$

From (A.15) it turns out that

$$\mathbf{G}^H\mathbf{U} = \frac{1}{e}\mathbf{G}^H\left(\mathbf{G}\mathbf{G}^H + \xi\mathbf{I}_{QN_T}\right)^{-1}\mathbf{G} = \mathbf{U}^H\mathbf{G} \tag{A.17}$$

from which it follows that equation (A.15) can be rewritten as follows

$$e \cdot \mathrm{tr}\left\{\mathbf{G}^H\mathbf{U}\mathbf{U}^H\mathbf{G} + \rho\mathbf{I}_K\right\} = \mathrm{tr}\left\{\mathbf{U}^H\mathbf{G}\right\}. \tag{A.18}$$

On the other hand, from (A.15) it also follows that

$$\mathbf{U}^H\mathbf{G} = e \cdot \mathbf{U}^H\left(\mathbf{G}\mathbf{G}^H + \xi\mathbf{I}_{QN_T}\right)\mathbf{U} \tag{A.19}$$

from which we have

$$\mathrm{tr}\left\{\mathbf{U}^H\mathbf{G}\right\} = e \cdot \mathrm{tr}\left\{\mathbf{U}^H\mathbf{G}\mathbf{G}^H\mathbf{U}\right\} + e\xi \cdot \mathrm{tr}\left\{\mathbf{U}\mathbf{U}^H\right\}. \tag{A.20}$$

Substituting (A.20) into (A.18) produces

$$e\rho K = e\xi \cdot \mathrm{tr}\left\{\mathbf{U}\mathbf{U}^H\right\} \tag{A.21}$$

from which, bearing in mind (A.6), we get

$$\xi = \rho. \tag{A.22}$$

Finally, it follows that

$$\mathbf{U} = \frac{1}{e}\mathbf{G}\left(\mathbf{G}^H\mathbf{G} + \rho\mathbf{I}_K\right)^{-1} \tag{A.23}$$

where e must be designed so as to meet (A.6) and reads

$$e = \sqrt{\frac{\mathrm{tr}\left\{\left(\mathbf{G}^H\mathbf{G} + I_K\right)^{-2}\mathbf{G}^H\mathbf{G}\right\}}{K}}. \tag{A.24}$$

Using (A.5) and recalling that $e \cdot \mathbf{G}^H\mathbf{U} - \mathbf{I}_K = -\rho \cdot \left(\mathbf{G}^H\mathbf{G} + \rho\mathbf{I}_K\right)^{-1}$ yields

$$J_{2,\min} = \sigma_w^2 \cdot \mathrm{tr}\left\{\left(\mathbf{G}^H\mathbf{G} + \mathbf{I}_K\right)^{-1}\right\}. \tag{A.25}$$

In the next we introduce an alternative approach that follows the line of reasoning of the method proposed in Choi & Murch (2004). It leads to the same solution discussed in Kusume *et al.* (2005) but it is mathematically simpler. To proceed, we rewrite (A.5) as follows

$$J_3 = \sigma_a^2 \cdot \mathrm{tr}\left\{\left(e\mathbf{G}^H\mathbf{U} - I_K\right)\left(e\mathbf{G}^H\mathbf{U} - I_K\right)^H\right\} + \sigma^2\mathrm{tr}\left\{e^2\mathbf{U}\mathbf{U}^H\right\} \tag{A.26}$$

where we have used the constraint (A.6). Equation indicates that J_3 *does not depend* on $\mathbf{U}$ and e separately but only on their product. This means that J_3 is only a function of the matrix $\mathbf{F}$ given by

$$\mathbf{F} = e \cdot \mathbf{U}. \tag{A.27}$$

From the above discussion, it follows that the solution to the problem can be find minimizing first J_3 with respect $\mathbf{F}$ and then using the result to properly define the parameter e, which must be defined so as to meet (A.6). Substituting (A.27) into (A.26) produces

$$J_3 = \sigma_a^2 \cdot \mathrm{tr}\left\{\left(\mathbf{G}^H\mathbf{F} - I_K\right)\left(\mathbf{G}^H\mathbf{F} - I_K\right)^H\right\} + \sigma^2\mathrm{tr}\left\{\mathbf{F}\mathbf{F}^H\right\}. \tag{A.28}$$

Taking the derivative of J_3 with respect to $\mathbf{F}$ and setting the result to zero, we obtain

$$\mathbf{F} = \mathbf{G}\left(\mathbf{G}\mathbf{G}^H + \rho\mathbf{I}_K\right)^{-1}. \tag{A.29}$$

Recalling that $\mathbf{U} = \frac{1}{e} \cdot \mathbf{F}$, from (A.6) we have

$$\mathrm{tr}\left\{\mathbf{F}\mathbf{F}^H\right\} = Ke^2 \tag{A.30}$$

from which it follows that

$$e = \sqrt{\frac{\mathrm{tr}\left\{\left(\mathbf{G}^H\mathbf{G} + I_K\right)^{-2}\mathbf{G}^H\mathbf{G}\right\}}{K}}. \tag{A.31}$$

Collecting (A.29) and (A.31) and bearing in mind that $\mathbf{U} = \frac{1}{e} \cdot \mathbf{F}$ we obtain the same results as given by (A.23) and (A.24).

Appendix B

Design of the backward matrix in THP-based schemes

B.1 Without QoS constraints

In the next we highlight the major steps leading to (5.29). Our goal is to find the *unit-diagonal* and *lower triangular* matrix $\mathbf{C}$ that minimizes the RHS of (5.28). For this purpose, we consider the Cholesky factorization of the matrix

$$\tilde{\mathbf{G}}^H\tilde{\mathbf{G}} + \rho\mathbf{I}_K. \tag{B.1}$$

This leads to the following result

$$\tilde{\mathbf{G}}^H\tilde{\mathbf{G}} + \rho\mathbf{I}_K = \tilde{\mathbf{L}}\tilde{\mathbf{L}}^H \tag{B.2}$$

where $\tilde{\mathbf{L}}$ is a $K \times K$ *lower triangular* matrix with real positive elements $\left[\tilde{\mathbf{L}}\right]_{k,k}$ $(k = 1, 2, \ldots, K)$ on its main diagonal. Substituting (B.2) into (5.28) produces

$$J = \sigma_n^2 \cdot \mathrm{tr}\left\{\left(\tilde{\mathbf{L}}^{-1}\mathbf{C}\right)^{\mathrm{H}}\tilde{\mathbf{L}}^{-1}\mathbf{C}\right\}. \tag{B.3}$$

Next, we observe that $\tilde{\mathbf{L}}^{-1}\mathbf{C}$ is still lower triangular and, in consequence, we may rewrite (B.3) in the equivalent form

$$J = \sigma_n^2 \sum_{k=1}^{K} \frac{1}{\left[\tilde{\mathbf{L}}\right]_{k,k}^2} + \sigma_n^2 \sum_{k=2}^{K} \sum_{i=1}^{k-1} \left| \left[\tilde{\mathbf{L}}^{-1}\mathbf{C}\right]_{k,i} \right|^2 \tag{B.4}$$

where we have born in mind that $[\mathbf{C}]_{k,k}$=1 for $k = 1, 2, \ldots, K$. From (B.4) it follows that the minimum of J is achieved when $\tilde{\mathbf{L}}^{-1}\mathbf{C}$ is diagonal, i.e.,

$$\tilde{\mathbf{L}}^{-1}\mathbf{C} = \tilde{\mathbf{D}}. \tag{B.5}$$

Finally, premultiplying both sides of (B.5) by $\tilde{\mathbf{L}}$ produces the result (5.29) in the text.

B.2 With QoS constraints

In this Appendix we highlight the major steps leading to (5.59). Our goal is to find the matrix $\mathbf{C}$ that maximizes the RHS of (5.60), which is equivalent to minimizing

$$J = \mathrm{tr}\{\mathbf{C}^H \mathbf{\Lambda}^H (\mathbf{G}\mathbf{G}^H)^{-1} \mathbf{\Lambda} \mathbf{C}\}. \tag{B.6}$$

We begin by considering the Cholesky factorization of the matrix $\mathbf{G}\mathbf{G}^H$ in (B.6), i.e.,

$$\mathbf{G}\mathbf{G}^H = \mathbf{S}\mathbf{S}^H \tag{B.7}$$

where $\mathbf{S}$ is a $K \times K$ lower triangular matrix with real positive elements on the main diagonal. Substituting (B.7) into (B.6) produces

$$J = \mathrm{tr}\left\{\mathbf{C}^H \mathbf{\Lambda}^H \left(\mathbf{S}^H\right)^{-1} \mathbf{S}^{-1} \mathbf{\Lambda} \mathbf{C}\right\} \tag{B.8}$$

or, equivalently,

$$J = \mathrm{tr}\left\{ \left(\mathbf{S}^{-1}\mathbf{\Lambda C}\right)^H \mathbf{S}^{-1}\mathbf{\Lambda C} \right\}. \tag{B.9}$$

Next, we observe that $\mathbf{S}^{-1}\mathbf{\Lambda C}$ is still a lower triangular matrix and, in consequence, equation (B.9) can be rewritten as

$$J = \sum_{k=1}^{K} \frac{\lambda_k^2}{[\mathbf{S}]_{k,k}^2} + \sum_{k=2}^{K}\sum_{\ell=1}^{k-1} \left| \left[\mathbf{S}^{-1}\mathbf{\Lambda C}\right]_{k,\ell} \right|^2 \tag{B.10}$$

where we have born in mind that $\mathbf{C}$ is unit-diagonal and lower triangular, i.e., $[\mathbf{C}]_{k,k}$=1 for $k = 1, 2, \ldots, K$.

From (B.10) it follows that the minimum of (B.6) is achieved when $\mathbf{S}^{-1}\mathbf{\Lambda C}$ is diagonal, i.e.,

$$\mathbf{S}^{-1}\mathbf{\Lambda C} = \mathbf{D}. \tag{B.11}$$

Finally, premultiplying both sides of (B.11) by $\mathbf{\Lambda}^{-1}\mathbf{S}$ produces the result (5.59).

Appendix C

Estimation accuracy of the noise power and channel estimators

C.1 Computing the variance of the unstructured noise power estimate

In this Appendix we compute the expectation of the noise power estimator in (6.16) and the variance of $\hat{\sigma}_p^{2(UCE)}$. Substituting (6.17) into (6.16) and bearing in mind (6.10) produces

$$\hat{\sigma}_p^2 = \frac{1}{NN_T} \cdot \left\{ \sum_{m=1}^{N_T} \mathbf{w}_p^H(m)\mathbf{w}_p(m) - \sum_{m=1}^{N_T}\sum_{\ell=1}^{N_T} \mathbf{w}_p^H(m)\mathbf{D}(m)\mathbf{T}^{-1}\mathbf{D}^H(\ell)\mathbf{w}_p(\ell) \right\} \tag{C.1}$$

or, alternatively,

$$\hat{\sigma}_p^2 = \frac{1}{NN_T} \cdot \left\{ \mathbf{w}_p^H \left(\mathbf{I}_{NN_T} - \mathbf{L} \right) \mathbf{w}_p \right\} \tag{C.2}$$

where $\mathbf{w}_p = \left[\mathbf{w}_p^T(1)\mathbf{w}_p^T(2) \cdots \mathbf{w}_p^T(N_T) \right]^T$ is a Gaussian vector with zero-mean and covariance matrix $\mathbf{C}_{\mathbf{w}p} = \sigma_p^2 \mathbf{I}_{NN_T}$ while $\mathbf{L}$ is a matrix composed by N_T^2 blocks

$[\mathbf{L}]_{m,\ell}$ $(1 \leq m, \ell \leq N_T)$, each of dimensions $N \times N$ and expressed by

$$[\mathbf{L}]_{m,\ell} = \mathbf{D}(m)\mathbf{T}^{-1}\mathbf{D}^H(\ell). \tag{C.3}$$

Recalling that Kay (1993)

$$\mathrm{E}\left\{\mathbf{w}_p^H\left(\mathbf{I}_{NN_T} - \mathbf{L}\right)\mathbf{w}_p\right\} = \mathrm{tr}\left\{\left(\mathbf{I}_{NN_T} - \mathbf{L}\right)\mathbf{C}_{\mathbf{w}_p}\right\} \tag{C.4}$$

from (C.2) we have

$$\mathrm{E}\left\{\hat{\sigma}_p^2\right\} = \frac{\sigma_p^2}{NN_T} \cdot \mathrm{tr}\left\{\left(\mathbf{I}_{NN_T} - \mathbf{L}\right)\right\}. \tag{C.5}$$

On the other hand, bearing in mind (C.3) yields

$$\mathrm{tr}\left\{\mathbf{L}\right\} = \sum_{m=1}^{N_T} \mathrm{tr}\left\{\mathbf{D}(m)\mathbf{T}^{-1}\mathbf{D}^H(m)\right\}. \tag{C.6}$$

At this stage we use the identity $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{tr}\{\mathbf{BA}\}$ in the RHS of (C.6) with $\mathbf{A} = \mathbf{D}(m)$ and $\mathbf{B} = \mathbf{T}^{-1}\mathbf{D}^H(m)$. This produces

$$\mathrm{tr}\left\{\mathbf{L}\right\} = \mathrm{tr}\left\{\mathbf{I}_{KN}\right\} = KN \tag{C.7}$$

where we have also taken (6.14) into account. Finally, substituting (C.7) into (C.5) yields the result (6.18).

Now, we compute the variance of $\hat{\sigma}_p^{2(UCE)}$ in (6.19). To this purpose, we rewrite $\hat{\sigma}_p^{2(UCE)}$ in the equivalent form

$$\hat{\sigma}_p^{2(UCE)} = \frac{1}{N\left(N_T - K\right)} \cdot \left\{\mathbf{w}_p^H\left(\mathbf{I}_{NN_T} - \mathbf{L}\right)\mathbf{w}_p\right\} \tag{C.8}$$

and exploit the result Kay (1993)

$$\mathrm{var}\left\{\mathbf{w}_p^H\mathbf{A}\mathbf{w}_p\right\} = \mathrm{tr}\left\{\mathbf{A}\mathbf{C}_{\mathbf{w}_p}\mathbf{A}\mathbf{C}_{\mathbf{w}_p}\right\} \tag{C.9}$$

which holds true for any Hermitian matrix $\mathbf{A}$. In particular, letting $\mathbf{A} = \mathbf{I}_{NN_T} - \mathbf{L}$ and recalling that $\mathbf{C}_{\mathbf{w}p} = \sigma_p^2 \mathbf{I}_{NN_T}$, it follows that

$$\text{var}\left\{\hat{\sigma}_p^{2(UCE)}\right\} = \frac{\sigma_p^4}{N^2\left(N_T - K\right)^2}\text{tr}\left\{\left(\mathbf{I}_{NN_T} - \mathbf{L}\right)^2\right\} \tag{C.10}$$

which can also be rewritten as

$$\text{var}\left\{\hat{\sigma}_p^{2(UCE)}\right\} = \frac{\sigma_p^4}{N^2\left(N_T - K\right)^2}\text{tr}\left\{\left(\mathbf{I}_{NN_T} - \mathbf{L}\right)\right\} \tag{C.11}$$

since $\mathbf{L}$ is idempotent (i.e., $\mathbf{L}^2 = \mathbf{L}$). Finally, substituting (C.7) into (C.11) leads to (6.20).

C.2 Computing the MSEE of SCE

In this Appendix we compute the expectation and the MSEE of $\hat{\mathbf{H}}_p^{(SCE)}$. Substituting (6.10) into (6.31) and bearing in mind (6.30) we get

$$\hat{\mathbf{H}}_p^{(SCE)} = \mathbf{H}_p + \mathbf{\Sigma}\mathbf{R}^{-1}\mathbf{\Sigma}^H \sum_{m=1}^{N_T} \hat{\mathbf{D}}^H(m)\mathbf{w}_p(m) \tag{C.12}$$

from which it follows that $\hat{\mathbf{H}}_p^{(SCE)}$ is unbiased since $\mathbf{w}_p(m)$ has zero mean. On the other hand, recalling that $\{\mathbf{w}_p(m); m = 1, 2, \ldots, N_T\}$ are statistically independent (i.e., $E\{\mathbf{w}_p^H(m)\mathbf{w}_p(\ell)\} = \mathbf{0}$ with $\ell \neq m$ and $E\{\mathbf{w}_p^H(m)\mathbf{w}_p(m)\} = \sigma_p^2\mathbf{I}_N$), from (C.12) we see that the correlation matrix of $\hat{\mathbf{H}}_p^{(SCE)}$ is given by

$$\mathbf{C}_{\hat{\mathbf{H}}_p^{(SCE)}} = \sigma_p^2 \ \mathbf{\Sigma}\mathbf{R}^{-1}\mathbf{\Sigma}^H. \tag{C.13}$$

Using the identity

$$\text{E}\{||\ \hat{\mathbf{H}}_p^{(SCE)} - \mathbf{H}_p||^2\} = \text{tr}\{\mathbf{C}_{\hat{\mathbf{H}}_p^{(SCE)}}\} \tag{C.14}$$

and recalling that

$$MSEE = \frac{1}{KN}\mathrm{E}\left\{ \left\| \hat{\mathbf{H}}_p^{(SCE)} - \mathbf{H}_p \right\|^2 \right\} \tag{C.15}$$

from (C.13) we have

$$MSEE = \frac{\sigma_p^2}{KN}\mathrm{tr}\{\boldsymbol{\Sigma}\mathbf{R}^{-1}\boldsymbol{\Sigma}^H\}. \tag{C.16}$$

Finally, applying the property $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{tr}\{\mathbf{BA}\}$ in the RHS of (C.16) with $\mathbf{A} = \boldsymbol{\Sigma}$ and $\mathbf{B} = \mathbf{R}^{-1}\boldsymbol{\Sigma}^H$ produces

$$MSEE = \frac{\sigma_p^2}{KN}\mathrm{tr}\{\mathbf{R}^{-1}\boldsymbol{\Sigma}^H\boldsymbol{\Sigma}\} \tag{C.17}$$

which is equivalent to the result (6.32) since $\boldsymbol{\Sigma}^H\boldsymbol{\Sigma} = N \cdot \mathbf{I}_{KL}$.

C.3 Evaluating the performance of LMS-UCE over a static channel

In this Appendix we highlight the major steps leading to the performance of the LMS-UCE. For the sake of simplicity, we drop the subscript $(\cdot)_p$ designating the antenna branch and assume that the channel is static (i.e., $\mathbf{H}_p(m) = \mathbf{H}_p$). We begin by computing the conditional expectation $\mathrm{E}\{\mathbf{e}(m) \,\big|\, \hat{\mathbf{H}}(m)\}$. To this purpose we substitute (6.10) into (6.47) to get

$$\mathbf{e}(m) = \hat{\mathbf{D}}^H(m)\hat{\mathbf{D}}(m)\Delta\hat{\mathbf{H}}(m) + \hat{\mathbf{D}}^H(m)\mathbf{w}(m) \tag{C.18}$$

where $\Delta\hat{\mathbf{H}}(m) = \mathbf{H} - \hat{\mathbf{H}}(m)$ is the estimation error at the mth step. Then, using the identity $\mathrm{E}\{\hat{\mathbf{D}}^H(m)\hat{\mathbf{D}}(m)\} = (A_2/Q) \times \mathbf{I}_{KN}$ (which is valid for independent data symbols with zero mean and variance A_2), we have

$$\mathrm{E}\left\{ \mathbf{e}(m) \,\middle|\, \hat{\mathbf{H}}(m) \right\} = \frac{A_2}{Q} \times \Delta\hat{\mathbf{H}}(m). \tag{C.19}$$

The above result indicates that $\mathbf{e}(m)$ may be thought of as the sum of $(A_2/Q) \times \Delta\hat{\mathbf{H}}_m$ plus a zero-mean disturbance term $\mathbf{n}(m)$. Accordingly, recursion (6.46) may be rewritten as

$$\Delta\hat{\mathbf{H}}(m+1) = (1 - \mu\frac{A_2}{Q}) \times \Delta\hat{\mathbf{H}}(m) - \mu\mathbf{n}(m) \tag{C.20}$$

with $\mathbf{n}(m) = [\hat{\mathbf{D}}^H(m)\hat{\mathbf{D}}(m) - (A_2/Q) \times \mathbf{I}_{KN}]\Delta\hat{\mathbf{H}}(m) + \hat{\mathbf{D}}^H(m)\mathbf{w}(m)$. Since in the steady-state $\hat{\mathbf{H}}(m) \approx \mathbf{H}$ (i.e., $\Delta\hat{\mathbf{H}}(m) = \mathbf{0}$), it is reasonable to approximate $\mathbf{n}(m)$ as

$$\mathbf{n}(m) \approx \hat{\mathbf{D}}^H(m)\mathbf{w}(m). \tag{C.21}$$

Inspection of (C.20) reveals that $\Delta\hat{\mathbf{H}}(m)$ may be viewed as the response to $\mathbf{n}(m)$ of a digital filter with impulse response

$$p_k = \begin{cases} -\mu(1 - \mu\frac{A_2}{Q})^{k-1} & k \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{C.22}$$

so that (C.20) becomes

$$\Delta\hat{\mathbf{H}}(m) = \sum_i p_i\, \mathbf{n}(m-i). \tag{C.23}$$

Recalling that $\mathbf{n}(m)$ has zero-mean, from (C.23) we see that $\mathrm{E}\{\Delta\hat{\mathbf{H}}(m)\} = \mathbf{0}$, meaning that $\hat{\mathbf{H}}(m)$ is an unbiased estimate of $\mathbf{H}(m)$.

Returning to (C.21), we observe that vectors $\{\mathbf{n}(m)\}$ are statistically independent for different values of m and have covariance matrix $\mathbf{C}_\mathbf{n} = \sigma^2 A_2/Q \times \mathbf{I}_{KN}$. Putting these facts together, from (C.23) we have

$$\mathrm{E}\left\{\Delta\hat{\mathbf{H}}_m \Delta\hat{\mathbf{H}}_m^H\right\} = \left[\sigma^2 \frac{A_2}{Q} \sum_i p_i^2\right] \times \mathbf{I}_{KN}. \tag{C.24}$$

Next, substituting (C.22) into (C.24) and using the identity

$$||\Delta\hat{\mathbf{H}}(m)||^2 = \mathrm{tr}\{\Delta\hat{\mathbf{H}}(m)\Delta\hat{\mathbf{H}}^H(m)\} \tag{C.25}$$

produces

$$\mathrm{E}\left\{\left\|\Delta\hat{\mathbf{H}}_m\right\|^2\right\} = \frac{\mu KN}{2-\mu\frac{A_2}{Q}}\sigma^2. \tag{C.26}$$

At this stage we introduce the noise equivalent bandwidth of the filter p_k Mengali & A.N.D'Andrea (1997)

$$B_L = \frac{\mu A_2}{2Q(2-\mu\frac{A_2}{Q})T_B}. \tag{C.27}$$

Finally, collecting (C.26) and (C.27) yields (6.48).

Bibliography

Bello, P.A. (1963). Characterization of randomly time-varying linear channels. *IEEE Trans. Commun. Syst.*, **11**, 360–393. 15

Bowers, R. (1978). *Communications for a Mobile Society*. Sage Publications, Cornell University. 3

Brandt-Pearce, M. & Dharap, A. (2000). Transmitter-based multiuser interference rejection for the down-link of a wireless CDMA system in a multipath environment. *IEEE Trans. on Signal Processing*, **18**, 607–617. 42, 59

Brittain, J.E. (1992). Scanning the past: Guglielmo marconi. *Proc. of IEEE*, **80**, 1341–1342. 3

Cacopardi, S., Frescura, F. & Reali, G. (1997). Performance comparison of multicarrier DS-SS radio access schemes for WLAN using measured channel delay profiles. In *Proc. of IEEE Vehicular Technology Conf. (VTC'97)*, vol. 3, 1877–1881. 39, 98

Choi, L. & Murch, R.D. (2004). A transmit MIMO scheme with frequency domain pre-equalization for wireless frequency selective channels. *IEEE Trans. on Communications*, **3**, 929–938. 41, 70, 132

Cooper, M. (Granted on September, 1975). *US patent number 3,906,166*. 4

Cosovic, I., Sand, S. & Raulefs, R. (2005). A non-linear precoding technique for downlink MC-CDMA. In *Proc. of IEEE Vehicular Technology Conference (VTC'05, Spring)*, Stockholm, Sweden. 60, 61, 79, 90

Deng, K., Yin, Q., Luo, M. & Zeng, Y. (2004). Blind uplink channel and DOA estimator for space-time block coded MC-CDMA systems with uniform linear array. In *Proc. of IEEE Vehicular Technology Conference (VTC'04)*, vol. 1, 69–73. 98

Divsalar, D., Simon, M. & Raphaeli, D. (1998). Improved parallel interference cancellation for CDMA. *IEEE Trans. on Communications*, **46**, 258–268. 6, 28, 29, 99

Fazel, K. (1993). Performance of CDMA/OFDM for mobile communications systems. In *Proc. of IEEE Int. Conf. on Universal Personal Commun. (ICUPC'93)*, vol. 2, 975–979. 18

Fazel, K. & Kaiser, S. (2003). *Multi-Carrier and Spread Spectrum Systems*. John Wiley and Sons. 7, 18, 27, 28, 35, 100

Fisher, R.F.H. (2002). *Precoding and Signal Shaping for Digital Transmission*. Wiley-IEEE Press. 63, 70, 87

H.Harashima & H.Miyakawa (1972). Matched-transmission technique for channels with intersymbol interference. *IEEE Trans. on Commununications*, **20**, 774–780. 60

Horn, R.A. & Johnson, C.R. (1985). *Matrix Analysis*. Cambridge University Press, New York. 16

Hunger, R., Joham, M. & Utschick, W. (2005). Extension of linear and non-linear transmit filters for decentralized receivers. In *Proc. of European Wireless Conference*, vol. 1, 40–46. 60, 63

Irmer, R., Rave, W. & Fettweis, F. (2003). Minimum BER transmission for TDD CDMA in frequency selective channels. In *Proc. of IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'03)*, 1260–1264, Beijing, China. 34

Joham, M., Utschick, W. & Nossek, J.A. (2005). Linear transmit processing in MIMO communications systems. *IEEE Trans. on Signal Processing*, **53**, 2700–2711. 73

Kaiser, S. & Hoeher, P. (1997). Performance of multi-carrier CDMA systems with channel estimation in two dimensions. In *Proc. of IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications PIMRC'97*, vol. 1, 115–119. 39, 97, 98

Kay, S. (1993). *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall. 88, 140

Keusgen, W., Walke, C. & Rembold, B. (2001). A system model considering the influence of front-end imperfections on the reciprocity of up- and down-link system impulse responses. In *Proc. of Aachen Symposium on Signal Theory (ASST'01)*, 243–248, Aachen, Germany. 33

Kusume, K., Joham, M., Utschick, W. & Bauch, G. (2005). Efficient Tomlinson-Harashima precoding for spatial multiplexing on flat MIMO channel. In *Proc. of IEEE Int. Conf. on Communications (ICC'05)*, Seoul, Korea. 41, 60, 70, 72, 73, 77, 89, 130, 132

Lancaster, P. (1969). *Theory of matrices*. Academic Press, New York. 16

Li, J., Kim, H. & Kim, Y. (2003). A novel broadband wireless OFDMA scheme for downlink in cellurar communications. *IEEE WCNC 2003*, **3**, 1907–1911. 83

Liu, J. & Krzymien, W.A. (2005). A novel non-linear precoding algorithm for the downlink of multiple antenna multi-user systems. In *Proc. of IEEE Vehicular Technology Conference (VTC'05, Spring)*, Stockholm, Sweden. 89, 90, 92

Lupas, R. & Verdú, S. (1990). Near-far resistance of multiuser detectors in asynchronous channels. *IEEE Trans. Comm.*, **38**, 496–508. 6

Lutkepohl, H. (1996). *Handbook of Matrices*. John Wiley and Sons, NY. 41

Madhow, U. & Honig, M.L. (1990). MMSE interference suppression for direct-sequence spread-spreactrum cdma. *IEEE Trans. Comm.*, **42**, 3178–3188. 6

Marques, P., Gameiro, A. & Fernandes, J. (2003). Vectorial channel estimation for uplink MC-CDMA in beyond 3G wireless systems. In *Proc. of IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'03)*, vol. 1, 984–988. 98, 99, 118

Medbo, J. (1998). *Channels Models for Hiperlan/2 in different indoor scenarious*. ETSI BRAN doc. 44, 75

Mengali, U. & A.N.D'Andrea (1997). *Synchronization Techniques for Digital Receiver*. Plenum Press, New York. 114, 144

Morelli, M. (2004). Timing and frequency synchronization for the uplink of an OFDMA system. *IEEE Trans. on Communications*, **52**, 296–306. 100

Moshavi, S. (1996). Multiuser detection for DS-CDMA communications. *IEEE Comm. Magazine*, 124–136. 7

Patel, P. & Holtzman, J. (1990). Analysis of a simple successive interference cancellation scheme in a DS-CDMA system. *IEEE Journ. Select. Areas Comm.*, **12**, 796–807. 6

Peijun, S. & Rappaport, T.S. (1998). Parallel interference cancellation PIC improvements for CDMA multiuser receveirs using partial cancellation of MAI estimates. In *Proc. of IEEE Global Telecommunications Conf. (GlobeCom'98)*, vol. 6, 3282–3287. 29

Popovic, B.V. (1999). Spreading sequences for multicarrier CDMA systems. *IEEE Trans. Commun.*, **47**, 918–926. 18

Proakis, J. (1995). *Digital Communications*. New York: McGraw-Hill. 15, 26

Ring, D.H. (1947). *Mobile Telephony - wide area coverage*. Technical Report, Bell Laboratories. 3

Sälzer, T. & Mottier, D. (2003). Downlink strategies using antenna arrays for interference mitigation in Multi-Carrier CDMA. In *Proc. of the Int. Workshop on Multi-Carrier Spread-Spectrum (MC-SS'03)*, 315–326, Oberpfaffenhofen, Germany. 33, 34, 35, 40, 42, 55

Sanguinetti, L., Morelli, M. & Mengali, U. (2004). Channel estimation and tracking for MC-CDMA signals. *European Trans. on Telecommunications*, **15**, 249–258. 97, 98, 99, 100, 104, 105, 106, 107

Silva, A. & Gameiro, A. (2003). Pre-filtering antenna array for downlink TDD MC-CDMA systems. In *Proc. of IEEE Vehicular Technology Conf. (VTC'03, Spring)*, vol. 1, 641–645. 33, 34, 40, 43, 54, 55

Sun, W., Li, H. & Amin, M. (2003). MMSE detection for space-time coded MC-CDMA. In *Proc. of IEEE Int. Conf. on Communications (ICC'03)*, vol. 5, 3452–3456. 48

T. K. Y. Lo (1999). Maximum ratio transmission. *IEEE Trans. on Communications*, **47**, 1458–1461. 61

Tomlinson, M. (1971). New automatic equalizer employing modulo arithmetic. *Electronics Letters*, **7**, 138–139. 60

Tse, D. & Viswanath, P. (2004). *Fundamentals of Wireless Communication*. Cambridge. 26

Tung, T.L., Yao, K. & Hudson, R.E. (2001). Channel estimation and adaptive power allocation for perfomance and capacity improvement of multiple-antenna OFDM systems. In *Proc. of IEEE Third Workshop on Signal Processing Advances in Wireless Commun.*, vol. 5, 82–85. 107, 110

Tureli, U., Kivanc, D. & Liu, H. (2000). Channel estimation for multi-carrier CDMA. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2000*, vol. 5, 2909–2912, Instanbul, Turkey. 98

Verdú, S. (1998). *Multiuser Detection*. Cambridge University Press. 6, 104, 118, 123

W.B. Jang, B.R.V. & Pickholtz, R.L. (1998). Joint transmitter-receiver optimization in synchronous multiuser communications over multipath channels. *IEEE Trans. on Communications*, **46**, 269–278. 35, 61

Windpassinger, C., Fisher, R.F.H., Vencel, T. & Huber, J.B. (2004). Precoding in multiantenna and multiuser communications. *IEEE Trans. on Wireless Communications*, **3**, 1305–1316. 60, 63

Wolniansky, P., Foschini, G., Golden, G. & Valenzuela, R. (1998). V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel. In *Proc. of Int. Symposium on Signals, Systems, and Electronics (ISSSE'98)*, 295 – 300. 60, 72, 89

Wu, X. & Yin, Q. (2002). Uplink vector channel estimation in ISI-corrupted MC-CDMA systems with multiple antennas. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, 2765–2768.

Yee, N., Linnartz, J.P. & Fettweis, G. (1993). Multi-carrier CDMA in indoor wireless radio networks. In *Proc. of IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'93)*, vol. 3, 109–113, Yokohama, Japan. 18

Zheng, K., Zeng, G. & Wang, W. (2004). DFT-based uplink channel estimation for uplink MC-CDMA systems. In *Proc. of IEEE Int. Symposium on Spread Spectrum Techniques & Applications (ISSSTA'04)*, 570–574, Sydney, Australia. 98, 99

Biography

Sanguinetti Luca was born in Empoli, Italy, on February 19nd, 1977.

In June 1996 he received the *diploma di maturità scientifica* from Liceo Scientifico Statale "Giosué Carducci" in Piombino, Italy.

Since 1997 he joined the Facoltà di Ingegneria of the Università degli Studi di Pisa, from which he received the Doctor Engineer degree (cum laude) in information engineering in 2002.

Since 2002 he was with the Department of Information Engineering of the University of Pisa, where he worked toward the Ph.D. degree in information engineering under the supervision of Prof. Umberto Mengali and Prof. Michele Morelli.

In 2004, he was a visiting Ph.D. student at the German Aerospace Center (DLR), Oberpfaffenhofen, Germany.

Research Interests

His expertise and general interests span the areas of communications and signal processing, estimation and detection theory. Current research topics focus on transmitter and receiver diversity techniques for single- and multi-user fading communication channels, antenna array processing, channel estimation and equalization, MIMO systems, multicarrier systems, linear and non-linear prefiltering for interference mitigation in multiuser environment, space-time coding for block

transmissions.

Editorial Work

Reviewer for the journals: IEEE Trans. Commun.; IEEE Trans. Signal Processing; IEEE Trans. Wireless Commun.; IEEE Comm. Letters; IEEE Trans. on Vehicular Technology.