



KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT DER TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Aenberg 10, 3001 Leuven (Heverlee)

BLOCK TRANSMISSION TECHNIQUES FOR WIRELESS COMMUNICATIONS

Promotor:
Prof. Dr. ir. M. Moonen

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de toegepaste wetenschappen
door
Olivier ROUSSEAU

December 2004



KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT DER TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)

BLOCK TRANSMISSION TECHNIQUES FOR WIRELESS COMMUNICATIONS

Jury:

Prof. Dr. ir. G. De Roeck, voorzitter
Prof. Dr. ir. M. Moonen, promotor
Prof. Dr. ir. G. Leus (T.U. Delft)
Prof. Dr. ir. B. De Moor
Prof. Dr. ir. E. Van Lil
Prof. Dr. ir. L. Van der Perre (U. Antwerpen)
Dr. ir. M. Engels (F.M.T.C.)
Prof. Dr. ir. B. Ottersen (K.T.H. Stockholm)

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de toegepaste wetenschappen

door

Olivier ROUSSEAU

© Katholieke Universiteit Leuven – Faculteit Toegepaste Wetenschappen
Arenbergkasteel, B-3001 Heverlee (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotocopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm or any other means without written permission from the publisher.

D/2004/7515/98

ISBN 90-5682-561-5

Acknowledgments

The research that is presented in this book is the result of several years of efforts and investigations. Many people have contributed directly or indirectly to the successful outcome of this adventure and I want to thank them before to proceed with the thesis itself.

I first want to thank my promotor Marc Moonen for giving me the opportunity to carry out this PhD research in a stimulating and fertile environment. Marc, thank you for your trust and support throughout the years I have spent in your research group. Your constructive comments and suggestions, your experience and wise advice have been extremely stimulating and largely contributed to keep this research on the right track.

I also want to thank Geert Leus for the outstanding quality of the scientific interactions we have had throughout this research. Geert, thanks for sharing your knowledge and experience, thanks for your daily support, for your helping hand and for always being available, for the numerous and fruitful discussions that we have had throughout this research. Thanks also for having been a nice companion during my time at the KUL.

I would like to thank Prof. Bart De Moor, Prof. Manu Van Lil and Dr. Marc Engels for the time and effort they invested in the proofreading of this book. Their remarks and comments have contributed to improve the quality of the text. I also want to thank the jury members Prof. Liesbet Van Der Perre and Prof. Bjorn Ottersen and the chairman Prof. Guido De Roeck for investing their precious time and experience in this PhD.

Prof. Petre Stoica also deserves a special thanks for his scientific collaboration which resulted in several joint publications. His comments, suggestions and ideas have been of great importance in the development of several key ideas that are presented in this book.

I also want to thank my office mates Raphael Cendrillon, Thomas Klassen, Imad Barhumi, Geert Vanmeerbergen, Sharon Gannot and Benoit Bosschaert for the good memories and nice moments spent during these years at the KUL.

Thanks also to all my colleagues that contributed to make the department a nice and pleasant work environment: Geert Ysebaert, Koen Vanbleu, Gert Cuypers, Toon Van Waterschoot, Hilde Vanhaute, Simon Doclo, An Spriet, Geert Rombouts, Paschalis Tsiaflakis and Jan Vangorp.

Enfin, pour conclure par le primordial, merci a toi Sara pour ta patience ta douceur et ton soutien indefaillible. Tu as porte sur tes epaules la moitie du poids de cette these. Merci aussi a mes parents, mes freres, ma soeur, ma famille et mes amis. Sans vous tous, ceci aurait bien peu de sens.

Abstract

In order to meet the market demand for high data rates, most digital wireless communication systems rely on broadband channels and therefore suffer from *Inter Symbol Interference (ISI)*, a phenomenon that needs to be combatted at the receiver by appropriate equalization techniques in order to restore the transmitted information.

In this context, block transmission techniques based on the use of a *Cyclic-Prefix (CP)* have attracted a lot of attention in the last years for they allow an efficient and computationally cheap ISI cancellation procedure. Historically, *OFDM (Orthogonal Frequency Division Multiplexing)* was the first proposed block transmission scheme and has been adopted in numerous standards for high-speed data transmission in both wired and wireless applications. In the wireless context however, OFDM suffers of several problems, both on an implementational point of view and from a performance perspective.

Some recently proposed block transmission techniques, namely *Single-Carrier Cyclic Prefix (SC-CP)* and *Known Symbol Padding (KSP)* partly solve these implementation problems whilst preserving the possibility of using efficient equalization schemes that have the same low complexity as the OFDM equalizers. These techniques also enable the use of specific equalization schemes that offer higher performance at the cost of an increased computational complexity. In this thesis, we analyze these alternative block transmission techniques in the specific context of wireless communications with a special focus on KSP.

We first detail the equalization schemes that are traditionally used with the different considered block transmission techniques. We then carry out a detailed analysis of the achievable BER performance of these block transmission techniques in the light of existing *Wireless Local Area Networks (WLAN)* standards under the working assumption that the receiver has perfect channel knowledge.

We then analyze the possibilities offered by the different block transmission techniques towards the identification of the transmission channel. KSP appears to be the best suited scheme since the padded sequences can be seen as

short training sequences inserted in the stream of data symbols. We propose a new semi-blind Gaussian Maximum Likelihood (ML) method for the identification of the transmission channel which is specifically suited to this KSP context. This method asymptotically achieves the Cramer-Rao Bound and significantly outperforms existing semi-blind methods and has the advantage of a low computational complexity.

In the light of this discussion on channel identification, we propose a new KSP transmission scheme that simultaneously allows low-complexity equalization and accurate estimation of the channel, namely, *Shifted KSP (S-KSP)*. A detailed analysis of the achievable BER performance of the considered block transmission techniques where realistic channel estimates are used for the design of the equalizers highlights the promising performance of the proposed S-KSP scheme.

Besides this work on channel identification and equalization, we also discuss several approaches for *Direct Sequence Estimation* that are suited to the considered block transmission techniques. We specifically propose a new iterative *Maximum Likelihood Sequence Estimation (MLSE)* technique that is suited for both KSP and SC-CP transmission.

Finally, we consider the situation where the mobility of the users is increased. The channel then becomes time-varying and changes significantly during the transmission of a packet of data symbols. Relying on a newly proposed technique for the modeling of time-varying multipath channels, namely the *Basis Expansion Model (BEM)*, we propose novel methods that allow to estimate the channel relying on the padded sequences of the KSP scheme. This method outperforms existing ones in the considered context with a limited computational complexity. A detailed analysis of the overall system performance shows that KSP is also a suitable candidate for data transmission in the context of time-varying channels.

Dutch Abstract

Om aan de marktvraag naar hoge datasnelheden te kunnen voldoen, baseren de meeste digitale draadloze communicatiesystemen zich op breedbandkanalen. Zulke kanalen veroorzaken echter *Inter Symbol Interferentie (ISI)*, een fenomeen dat aan de ontvanger door aangewezen egalisatietechnieken moet worden bestreden.

In deze context hebben bloktransmissietechnieken, die op het gebruik van een *Cyclische Prefix (CP)* zijn gebaseerd, de laatste jaren heel wat aandacht getrokken, omdat zij een efficiënte en goedkope ISI-onderdrukking toelaten. Historisch gezien is *OFDM (Orthogonal Frequency Division Multiplexing)* één van de eerste bloktransmissietechnieken die voorgesteld werd. Deze techniek is aanwezig in talrijke standaarden voor breedbandtransmissie voor zowel bedrade als draadloze toepassingen. In de draadloze context lijdt OFDM nochtans aan verscheidene problemen, zowel vanuit het standpunt van implementatie als vanuit een performantieperspectief.

Sommige recent voorgestelde bloktransmissietechnieken, zoals *Single-Carrier Cyclic-Prefix (SC-CP)* en *Known Symbol Padding (KSP)*, lossen deze implementatieproblemen gedeeltelijk op, terwijl ze toch de mogelijkheid bewaren om efficiënte egalisatietechnieken te gebruiken met dezelfde lage complexiteit als de OFDM egalisatoren. Anderzijds laten deze technieken ook het gebruik toe van specifieke egalisatietechnieken die een hogere performantie aanbieden ten koste van een verhoogde complexiteit. In deze thesis analyseren wij deze alternatieve bloktransmissietechnieken in de specifieke context van draadloze communicatie. De nadruk zal hoofdzakelijk gelegd worden op KSP.

Wij beschrijven eerst de egalisatietechnieken die traditioneel gepaard gaan met de verschillende onderzochte bloktransmissietechnieken. Daarna voeren we een gedetailleerde analyse uit van de haalbare BER prestaties van deze bloktransmissietechnieken. Dit zal gebeuren in het licht van bestaande *Wireless Local Area Networks (WLAN)* standaarden, in de veronderstelling dat de ontvanger perfecte kanaalkennis heeft.

Wij analyseren dan de mogelijkheden die de verschillende bloktransmissietechnieken bieden voor kanaalidentificatie. KSP blijkt de meest geschikte techniek te zijn, aangezien de toegevoegde sequenties als korte trainingssequenties kunnen worden gebruikt. Wij stellen een nieuwe semi-blinde *Gaussian Maximum Likelihood (GML)* kanaalidentificatiemethode voor die specifiek voor deze KSP context ontwikkeld werd. Deze methode bereikt asymptotisch de *Cramer-Rao Bound (CRB)*, overtreft beduidend bestaande semi-blinde methodes, en heeft het voordeel van een lage rekencomplexiteit.

In het licht van deze bespreking over kanaalidentificatie, stellen wij dan een nieuwe KSP transmissietechniek voor die gelijktijdig een eenvoudige egalisatie en een nauwkeurige kanaalschatting toestaat, namelijk *Shifted KSP (S-KSP)*. Een gedetailleerde analyse van de haalbare BER performantie van de verschillende bloktransmissietechnieken, waarbij realistische kanaalschattingen gebruikt worden voor het berekenen van de egalisatoren, toont de veelbelovende prestaties aan van de voorgestelde S-KSP techniek.

Naast dit werk over kanaalschatting en egalisatie, bespreken wij ook technieken voor *Direct Sequence Estimation* die voor de verschillende bloktransmissietechnieken geschikt zijn. Meer specifiek stellen wij een nieuwe iteratieve *Maximum Likelihood Sequence Estimation (MLSE)* techniek voor die zowel voor KSP als voor SC-CP geschikt is.

Tot slot analyseren wij de situatie waarbij de gebruikers een hoge mobiliteit hebben. Het kanaal wordt dan tijdsvariërend en verandert beduidend tijdens de transmissie van een pakket van datasymbolen. Gebruik makend van het *Basis Expansion Model (BEM)*, een recent voorgestelde modelleringstechniek voor tijdsvariërende breedbandkanalen, stellen wij een aantal nieuwe methodes voor die het kanaal schatten gebruik makend van de korte trainingssequenties aanwezig in KSP. Een gedetailleerde analyse van de verschillende systeempertormanties toont aan dat KSP ook een geschikte kandidaat is voor datatransmissie over tijdsvariërende kanalen.

Glossary

Mathematical Notation

\mathbf{x}	vector \mathbf{x}
\mathbf{X}	matrix \mathbf{X}
\mathbf{X}^T	transpose of matrix \mathbf{X}
\mathbf{X}^H	Hermitian transpose of matrix \mathbf{X}
\mathbf{X}^*	complex conjugate of matrix \mathbf{X}
\mathbf{X}^{-1}	inverse of matrix \mathbf{X}
\mathbf{X}^\dagger	pseudoinverse of matrix \mathbf{X}
$\text{tr}\{\mathbf{X}\}$	trace of matrix \mathbf{X}
$\det\{\mathbf{X}\}$	determinant of matrix \mathbf{X}
$\ \mathbf{X}\ $	Frobenius norm of matrix \mathbf{X}
$\text{diag}\{\mathbf{x}\}$	square diagonal matrix with \mathbf{x} as diagonal.
$\text{diag}\{\mathbf{X}\}$	column vector built with the main diagonal of \mathbf{X}
$\mathbf{X}(k, l)$	element on the k th row and l th column of matrix \mathbf{X}
$x[k]$	k th element of the vector \mathbf{x}
$\mathbf{X}(k : l, :)$	rows k up to l of matrix \mathbf{X}
$\mathbf{X}(:, k : l)$	columns k up to l of matrix \mathbf{X}
$\mathbf{O}_{l,m}$	$l \times m$ all-zero matrix
\mathbf{I}_n	$n \times n$ identity matrix
$\Re\{x\}$	real part of x
$\Im\{x\}$	imaginary part of x
\hat{x}	estimate of x
\check{x}	hard estimate of x
$ x $	absolute value of x
$E\{x\}$	expectation of random variable x
\mathbb{R}	the set of real numbers
\mathbb{C}	the set of complex numbers
$\mathcal{N}(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$\text{Chol}\{\mathbf{X}\}$	Cholesky decomposition of \mathbf{X}
$\frac{\partial}{\partial x}$	partial derivative w.r.t. variable x
$J_0(\cdot)$	0^{th} order Bessel function of the first kind
\mathcal{O}	order

\mathcal{P}	bit error probability
\otimes	linear convolution

Fixed Symbols

α	rolloff factor of the pulse shaping filter
c	speed of light
f_c	carrier frequency
f_{max}	Doppler spread
f_{off}	carrier frequency offset
\mathcal{F}_P	$P \times P$ DFT matrix
\mathbf{h}	discrete baseband model of the transmission channel
\mathcal{I}_P	$P \times P$ IDFT matrix
k	block index
K	number of blocks in a packet of data
L	channel order
μ	length of the CP
N	total length of the channel input sequence in a packet
N_{mod}	period of the BEM
N_s	number of data symbols in a block
N_t	padded sequence length in KSP
N_x	total number of symbols in a block of channel input symbols
\mathbf{Q}	correlation matrix of the noise term in the presented GML method
Q	parameter of the BEM, $2Q + 1$ is the total number of complex exponentials
\mathbf{s}_k	k^{th} block of data symbols
\mathbf{t}_k	k^{th} padded sequence
σ^2	noise variance
T_s	sampling time
τ_{max}	maximal delay spread of the channel
τ_{coh}	coherence time of the channel
\mathbf{x}	total channel input sequence
\mathbf{x}_k	k^{th} block of channel input symbols
\mathbf{y}	total sequence of channel output samples
\mathbf{y}_k	k^{th} block of channel output samples
\mathbf{z}_k	k^{th} pre-processed received block

Acronyms and Abbreviations

ADSL	Asymmetric Digital Subscriber Line
AWGN	Additive White Gaussian Noise
BDFE	Block Decision Feedback Equalizer
BEM	Basis Expansion Model

BER	Bit Error Rate
BLE	Block Linear Equalizer
BPSK	Binary Phase Shift Keying
CDMA	Code Division Multiple Access
CFO	Carrier Frequency Offset
CP	Cyclic Prefix
CRB	Cramer Rao Bound
CSI	Channel State Information
DAB	Digital Audio Broadcasting
DFE	Decision Feedback Equalizer
DFT	Discrete Fourier Transform
DMT	Discrete Multi Tone
DVB	Digital Video Broadcasting
EM	Expectation Maximization
FD	Frequency Domain
FFT	Fast Fourier Transform
FIM	Fisher Information Matrix
FIR	Finite Impulse Response
Flop	Floating point operation
GML	Gaussian Maximum Likelihood
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
IBI	Inter Block Interference
ICI	Inter Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
IIR	Infinite Impulse Response
ILSE	Iterative Least Squares with Enumeration
ILSP	Iterative Least Squares with Projections
ISI	Inter Symbol Interference
KSP	Known Symbol Padding
LAN	Local Area Network
LS	Least Squares
MAP	Maximum A Posteriori
MC-CP	Multi Carrier Cyclic Prefix
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MLSE	Maximum Likelihood Sequence Estimation
MMSE	Minimum Mean Squared Error
MSE	Mean Squared Error
NMSE	Normalized Mean Squared Error
NZP-KSP	Non-Zero Padded Known Symbol Padding
OFDM	Orthogonal Frequency Division Multiplexing
PAPR	Peak to Average Power Ratio
PSAM	Pilot Symbol Assisted Modulation

PSD	Power Spectral Density
QAM	Quadrature Amplitude Modulation
QPSK	Quaternary Phase Shift Keying
RLS	Recursive Least Squares
S-KSP	Shifted Known Symbol Padding
SC-CP	Single Carrier Cyclic Prefix
SIMO	Single Input Multiple Output
SISO	Single Input Single Output
SLE	Serial Linear Equalizer
SNR	Signal to Noise Ratio
UMTS	Universal Mobile Telecommunications System
VDSL	Very high rate Digital Subscriber Line
WLAN	Wireless Local Area Network
WLS	Weighted Least Squares
ZF	Zero Forcing
ZP-KSP	Zero-Padded Known Symbol Padding

Contents

Acknowledgments	i
Abstract	iii
Abstract in Dutch	v
Glossary	vii
Contents	xi
1 Introduction	1
1.1 Scope of the Thesis	1
1.2 Context	4
1.3 Problem Statement	6
1.4 State of the Art	7
1.5 Thesis Survey and Contributions	9
2 Background Material	13
2.1 Introduction to Digital Wireless Communications	14
2.1.1 System Setup	14
2.1.2 Channel Model	17
2.2 Block Transmission Techniques: Data Model	23

2.2.1	Cyclic Prefix Transmission	26
2.2.2	KSP Transmission	29
3	Channel Equalization	31
3.1	Block Linear Equalizers	32
3.1.1	Serial Linear Equalizers	33
3.1.2	Block Linear Equalizers for CP Transmission	35
3.1.3	Block Linear Equalizers for KSP Transmission	37
3.2	Block Decision Feedback Equalizers for KSP Transmission	39
3.3	Performance Analysis and Simulation Results	42
3.3.1	BER Analysis	42
3.3.2	Introductory Experiment	47
3.3.3	Equalizers Performance in the Context of the Hiperlan2 and IEEE 802.11a Standards	48
3.4	Complexity Analysis and Implementation Issues	56
3.4.1	Computational Complexity	56
3.4.2	Implementational Issues	58
3.5	Conclusions	62
4	Channel Estimation	65
4.1	Introduction	66
4.2	Data Model	68
4.3	Maximum Likelihood Approach for Channel Identification	70
4.4	Cramer Rao Bounds	71
4.5	Iterative Procedure	72
4.6	Closed Form Solution	74
4.6.1	Constant Training Sequence	74
4.6.2	Changing Training Sequence	76

4.6.3	Noise Variance Estimate	79
4.6.4	Comparison with the Iterative Method	80
4.6.5	Identifiability Conditions	81
4.6.6	Complexity Analysis	81
4.7	Asymptotic Properties of the Closed Form Channel Estimates .	82
4.7.1	Constant Training Sequence	82
4.7.2	Changing Training Sequence	84
4.8	Simulation Results	86
4.8.1	Performance of the Proposed Method	87
4.8.2	Comparison with Existing Methods	93
4.9	Channel Estimation in Block Transmission	95
4.9.1	Channel Estimation in CP Transmission	95
4.9.2	Channel Estimation in KSP Transmission	97
4.10	Shifted Known Symbol Padding (S-KSP)	98
4.11	System Performance with Estimated Channel Models	100
4.12	Conclusions	102
5	Direct Symbol Estimation	111
5.1	Blind Maximum Likelihood Sequence Estimation Techniques .	112
5.2	Data Model and Maximum Likelihood Sequence Estimation . .	114
5.3	Iterative Least Squares with Projection	117
5.4	ILSP for OFDM Systems	119
5.5	Simulation Results	120
5.6	Conclusions	125
6	KSP in Doubly-Selective Channels	129
6.1	Data Model	133

6.1.1	Doubly Selective Channel Model	133
6.1.2	KSP in Doubly Selective Channels: Data Model	138
6.1.3	KSP in Doubly Selective Channels: Equalizers	140
6.2	Estimation of Doubly Selective channels in KSP Transmission	142
6.2.1	Adaptive Implementation of the Gaussian Maximum Likelihood Method	143
6.2.2	Adaptive BEM Method	146
6.2.3	Direct Estimation of the BEM parameters	147
6.3	Simulation Results	150
6.3.1	Channel Estimation	152
6.3.2	Channel Equalization	154
6.4	Conclusions	160
7	Conclusions and Future Research	163
7.1	Main Contributions	163
7.2	Conclusions	168
7.3	Open Issues and Future Research	169
	List of Publications	183
	Biography	185

Chapter 1

Introduction

Wireless telegraphy is all very well but I would rather send a message by a boy on a pony!

Lord Kelvin

1.1 Scope of the Thesis

In a recent past, digital communication systems were mainly devoted to a limited number of computer-to-computer communications. *Local Area Networks (LAN)* using a dedicated cabling infrastructure were used in order to enable the direct communication between computers located in a limited geographical area. Long distance transport of information between different LANs was a costly process that required the use of leased lines or alternative dedicated infrastructure. Remote users could connect to a given LAN through dial-up connections using modems operating at 56 kb/s in the voice band of analog phone lines. Audio or video contents were usually not transported over digital networks but rather transmitted or broadcasted in an analog format.

In the meantime, the evolution of several factors has made digital communications more and more attractive and the amount of digitally transported information has undergone an exponential growth in the last years. Digital networks are now the dominant communication mode and are expected to progressively replace analog formats even for the distribution of audio and video contents in a near future. The key factors that enabled this fast evolution include:

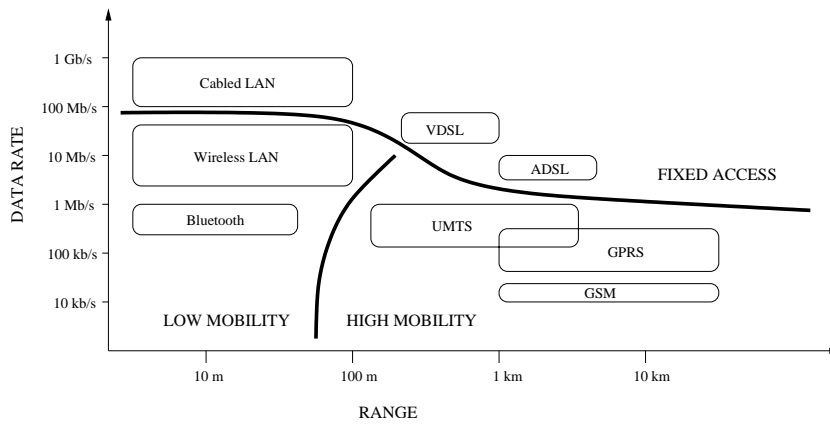


Figure 1.1: Data rate and range of different existing digital communication systems.

- The development of optical fiber, which enabled the deployment of high-speed communication backbones. These backbones now allow to transport large amounts of digital contents over long distances at a significantly reduced financial cost.
- The development and standardization of efficient compression algorithms that allow to represent analog contents such as voice or video with a limited bitrate.
- The development of advanced signal processing techniques and algorithms together with the increased performance of digital signal processing chipsets allow to access the high-speed backbone networks with a limited deployment cost. Thanks to these advanced signal processing techniques, efficient access networks can be deployed over existing infrastructures (phone lines, cable TV, ...) or directly over the air interface. The costly deployment of a dedicated infrastructure for the access network can thus be avoided, making high-speed digital communications available to the mass market.

These elements have triggered the fast development of the Internet and other multimedia applications which in turn results in a constantly increasing demand for high-speed access and local area networks. Improving the performance of these networks in a cost-effective way is today one of the major challenges faced by the telecommunication industry.

In **Fig. 1.1**, we compare several recent and upcoming communication systems that have been developed in this context. The figure compares the data rates

of the different systems as a function of the distance over which they are able to transport the data.

Fixed or wired applications that are located in the upper part of the figure clearly offer the best performance. *Local area networks* have been deployed for a long time and rely on a dedicated cabling infrastructure whose physical characteristics are designed to avoid the distortion of the transmitted signal. The evolution of digital electronic devices that run the switches and routers have allowed to progressively increase their transmission speeds and current LANs offer datarates between 100 Mb/s and 1Gb/s. In order to maintain sufficient propagation quality over the physical link, LANs are limited to a range of less than 100m. They are typically used to exchange data at a local level and to share a high-speed access (typically optical fiber) to the outside world amongst the users that are connected to the LAN. Several applications have been developed in order to allow individual users to remotely access the high-speed backbone network or remote LANs without deploying specific additional physical connections. The long-existing and well-deployed telephone network soon appeared as a straightforward means towards that goal. The first modems (not shown on the graphs) were operating in the same frequency band as the phone conversations for which the cabling infrastructure had been developed. The transmitted signals were not really distorted and mainly suffered from noise and attenuation effects. However the limited bandwidth of the voice band only allowed to establish communications at 56 kb/s in the best scenarios. Recent developments in signal processing now allow to transmit data over telephone lines outside the narrow frequency band for which they were designed. *Asymmetric Digital Subscriber Lines (ADSL)* for instance allows to deliver between 1.5 and 8 Mb/s to remote users up to a distance of 4 km. When the distance between the remote user and the access point to the backbone is less than a kilometer, the emerging *VDSL* technology will allow to deliver up to 55 Mbps over the telephone lines. In the higher frequencies that are used by such systems, the transmitted signals are severely distorted by the transmission channel and recent advances in signal processing and equalization techniques that allow to efficiently compensate for that distortion are the key technological advances that have enabled that evolution. The drawbacks of these systems is that they are all wired and hence do not offer any mobility to the users that wish to establish a connection. Another drawback which is specific to cabled LAN applications is their high deployment cost as they require the installation of a specific cabling infrastructure.

Simultaneously to this evolution of the fixed digital communication techniques, *wireless communications* have undergone a spectacular evolution in the last years which also relies on the evolution of the available signal processing techniques and algorithms. The first commercially succesfull wireless digital communication systems were the second generation mobile communication systems (e.g. GSM) that offer 13.6 kb/s services within cells whose size ranges from 1 to

more than 30 km. The transmission speeds offered by these 2G networks have recently been improved by the introduction of the GPRS standard that typically offers 50 kb/s speeds (up to 171 kb/s under some circumstances). Third generation mobile systems typically deliver 0.2 to 0.4 Mb/s (up to 2Mb/s) within cells of reduced size (max. 4 km). The main advantage of these 2G and 3G systems is that they allow the users to move freely (possibly at relatively high speeds) whilst continuously transmitting digital information, which is a major advantage over their wired counterparts. Their main drawback though is their relatively low transmission speeds which currently limits the applications that are available for such services. Besides these high-mobility 2G and 3G applications, *Wireless Local Area Networks (WLAN)* offering significantly higher transmission speeds have been developed and standardized recently (Hiperlan 2 and IEEE 802.11a, also known as Wi-Fi). These systems deliver between 6 and 54 Mb/s to users located up to an approximate distance of 100 m around a fixed base station (the highest datarates though are only available within a fraction of that area). They are currently becoming an alternative to their traditional wired counterparts as they have a significantly lower deployment and reconfiguration cost and allow much more flexible connectivity to the LAN users. Their main drawback is again that their transmission speeds currently remain significantly lower than that of the classical cabled LANs.

In order to become truly attractive and meet the market demand for high speed data transmission, it is crucial to improve the transmission speeds of these different wireless systems. Next generation WLANs will have to deliver datarates that are comparable to those offered by their wired counterpart (i.e. 100 Mb/s or more) and should enable the delivery of the highest rates in a more significant portion of the covered geographical area. Similarly, 4G mobile systems should at least offer the datarates that are offered by current WLAN systems (between 10 and 50 Mb/s) to users experiencing a possibly high mobility.

In this thesis, we aim at addressing these challenges through the development of new signal processing techniques that are suited to the context of broadband wireless communication systems. We study both the context of WLAN applications and mobile communication systems.

1.2 Context

Before to present the techniques that are used for data transmission over the wireless interface in further details, we outline below some important characteristics of the transmission channels that will be considered throughout this research.

a) Inter Symbol Interference Multipath effects are an important characteristic of wireless channels. When broadband communication systems are used over such channels, *Inter Symbol Interference (ISI)* arises. This ISI is a major impediment of digital communication systems that needs to be combated at the receiver. In this research, we discuss several channel equalization techniques that are able to cope efficiently with the ISI.

b) Limited Coherence Time Another important characteristic of wireless communication channels is their *limited coherence time*. When both the emitter and the receiver are at fixed positions, the presence of moving objects in their environment causes the channel to change and its coherence time cannot be assumed to be larger than 10 ms. This coherence time is significantly reduced when either the transmitter or the receiver is moving.

Assuming that the different users of the considered communication system sporadically transmit packets of data symbols, this limited coherence time has a considerable impact on the overall system design. In low mobility conditions, it is often reasonable to assume a constant channel model for the transmission of the whole packet, as the typical duration of such a packet is lower than the coherence time of the channel. However, as the latency time between two consecutive packets can be significantly longer than this coherence time, one has to assume independent channel realizations for different packets of data symbols. When the mobility increases, the coherence time of the channel can become smaller than the packet duration, in which case the evolution of the channel parameters has to be tracked within the packet.

c) Channel Unknown to the Transmitter Finally, given the small coherence time of the communication channel and the large number of parameters that characterize the communication channel, feeding back the channel state information to the transmitter is often problematic. Hence, we will focus on the situation where *the transmitter has no information on the channel* over which it is transmitting, a situation which is also relevant to the broadcast channel case.

d) SISO Setup In the presented research, we investigate how the transmitted signal should be designed in order to allow high performance equalization of the signal with a limited computational complexity in this specific context. We review several channel equalization techniques that are suited to the proposed structure of the transmitted signals. We then propose novel algorithms that allow optimal channel identification and direct symbol estimation at the receiver for both stationary and time-varying channels. Our focus is on *Single Input Single Output (SISO)* systems, i.e. communication systems where users having

one antenna each are separated in time and frequency. Given the high symbol rate of the considered systems and in order to possibly allow the real-time implementation of the proposed solutions, we aim at finding low-complexity solutions throughout our study whilst keeping an eye on the optimality of the proposed solutions.

1.3 Problem Statement

In order to offer competitive data transmission speeds, many recently developed digital wireless communication systems rely on broadband communication channels. A major impediment of such systems is that the symbol period can become smaller than the delay spread of the physical channel, especially in multipath scenarios. This results in ISI, a phenomenon that needs to be combatted at the receiver in order to restore the transmitted information. Serial linear equalization schemes are traditionally used towards that goal but they offer sub-optimal performance and suffer from a relatively high computational complexity.

In this context, block transmission techniques based on the use of a *Cyclic-Prefix (CP)* have attracted a lot of attention in the last years for they allow an efficient and computationally cheap ISI cancellation procedure [1], [2]. In CP transmission, the transmitted data symbols are organized in blocks and a CP, which is simply a repetition of the last data symbols of the block, is appended at the beginning of each block. When the CP is longer than the channel order, the effects of the channel can be described with a circulant convolution rather than a linear convolution. ISI can then be suppressed by a single-tap “frequency-domain” equalization on blocks of data symbols using FFT and IFFT operations. This simplified channel equalization procedure has a cost however: for a given symbol rate, CP transmission systems offer a lower throughput than their classical counterpart as the redundant symbols of the CP are transmitted on top of the useful message.

Historically, *OFDM (Orthogonal Frequency Division Multiplexing)* was the first proposed block transmission scheme and has been adopted in numerous standards for high-speed data transmission such as ADSL, Digital Audio and Video Broadcasting and in the Hiperlan2 [3] and IEEE802.11a and IEEE802.11g standards for Wireless Local Area Networks. OFDM belongs to the class of *Multi-Carrier block transmission techniques based on the use of a CP (MC-CP)* also known as *Discrete Multi-Tone (DMT)*. These techniques perform an IFFT at the transmitter after which the CP is added; the receiver performs an FFT followed by a one-tap frequency domain equalization step [1], [2]. This equalization scheme has a very limited computational complexity whilst it effectively cancels the effects of the multipath transmission channel. OFDM transmission

can be seen as the parallel transmission of independent data streams on orthogonal frequency-domain flat-fading channels, also called tones. The knowledge of the transmission channel at the transmitter allows to effectively achieve near-capacity transmission by optimizing the transmitted signals using power- and bit-loading techniques across the tones [4], [1, p 7], which makes OFDM an attractive transmission technique for applications where the transmission channel is known to the transmitter (ADSL or VDSL for instance).

In the considered context however, the channel is unknown to the transmitter. Hence, this loading is not applicable and the performance of the system becomes very sensitive to deep channel fades in the frequency-domain. A zero on the DFT grid of the channel will even result in the systematic loss of the data symbols transmitted on this tone. A solution that is often used to reduce the sensitivity to channel fades in the frequency domain consists in encoding the data across the tones, with the drawback of an increased complexity yielded by the encoding and decoding operations. Besides this, OFDM suffers from several other drawbacks that are discussed in *Ch. 3*. One of these drawbacks is the occurrence of large peaks in the transmitted time-domain signals, i.e. the transmitted signals exhibit a high *Peak to Average Power Ratio (PAPR)*. Other problems include the high sensitivity to incorrect carrier frequency estimation, out-of band noise, sensitivity to radio frequency interference or to time-varying channels.

Note that *Channel State Information (CSI)* is needed at the receiver for the design of the equalizers. Several techniques have been proposed to blindly or semi-blindly exploit the CP-induced signal structure towards channel identification. However, as the CP-induced signal properties are not very well suited for channel identification, the channel models obtained with these techniques generally lack accuracy. Hence, most CP transmission systems use long training sequences and/or pilot tones that further harm the throughput of the system.

1.4 State of the Art

An alternative block transmission technique known as *Single-Carrier Cyclic Prefix (SC-CP)* [5, pp 103-104], [2, p 36] has been proposed recently, where the transmitter simply adds a CP to every block of data symbols. As we will see in *CH. 3*, this technique solves many of the implementation problems encountered by OFDM, while keeping the advantage of computationally cheap equalizers relying on FFT and IFFT operations. Moreover, SC-CP can be seen as a classical OFDM transmission scheme where each block of data is spread across the tones by means of a linear encoder, I.E. a DFT. The sensitivity to channel fades in the frequency-domain is therefore reduced by this technique but channel-irrespective symbol recovery is still not guaranteed (if there is a

zero on the DFT grid of the channel).

An interesting block transmission scheme that has been proposed recently is referred to as *Known Symbol Padding (KSP)* [6] [7] [8]. In KSP transmission, the CP is replaced by a sequence of known symbols that acts as a guard interval between blocks of data symbols. This transmission scheme is called Zero-Padded KSP (ZP-KSP) [2] when all padded sequences are equal to zero, and is called Non-Zero Padded KSP (NZP-KSP) otherwise. Specific equalizers that we call *optimal KSP equalizers* in the rest of the text have been developed in this context. When these optimal equalizers are used, KSP transmission guarantees channel-irrespective symbol recovery and, as opposed to CP-based techniques, fully exploits the delay diversity introduced by the frequency-selective transmission channel [9]. However, this improved performance results in a significantly increased computational complexity of the equalization process. When all the padded sequences are the same, the low-complexity equalizers for CP-based transmission relying on FFT and IFFT operations can be used in the KSP context as well, but full diversity and channel-irrespective symbol recovery are not guaranteed anymore. When these equalizers are used, KSP is strictly equivalent to SC-CP transmission. A third type of equalizer can be used in KSP as well, the *Block Decision Feedback Equalizer (BDFE)* which has a slightly higher computational complexity than the optimal linear KSP equalizer but offers improved performance.

Another important advantage of NZP-KSP is that the known symbols can be exploited as training sequences for time and frequency synchronization (see e.g. [8], [10] or [11]), direct equalizer design [6] or channel estimation [12] [7]. Finally, when the channel becomes time-varying and changes significantly during the transmission of a packet of data symbols, the placement of the known symbols of the KSP scheme is optimal for the tracking of the channel variations [13].

These alternative block transmission techniques have been proposed recently and all their possibilities have not yet been thoroughly investigated. Their compared performance in practical system setups has not been fully investigated. Besides, channel estimation algorithms yielding acceptable performance with a limited computational complexity in the specific framework of KSP still have to be investigated. The impact of the padded sequences composition on the accuracy of the channel estimates has not been studied. The impact of channel modeling errors on the overall system performance is still an open issue.

This list of open questions is by far not exhaustive and shows the need for the development of signal processing techniques and algorithms that are specifically suited to these alternative block transmission techniques. In this thesis, we aim at investigating these alternative block transmission techniques in more details to gain a better understanding of their compared advantages and drawbacks. We also investigate the specific possibilities of KSP transmission scheme and

proposed novel techniques and algorithms that allow to exploit these possibilities keeping an eye on the computational complexity of the proposed methods.

1.5 Thesis Survey and Contributions

In this thesis, we first concentrate on low mobility scenarios where the channel stays constant during the transmission of a whole data packet.

In **Ch. 2**, we present the background material that is going to be used throughout the thesis, i.e. we introduce wireless setups and show how the channel models and input-output relationships are derived. We then introduce the different block transmission techniques that will be investigated and give their corresponding data models.

In **Ch. 3**, we detail the equalization schemes that are traditionally used with the different considered block transmission techniques. We present the specific ZF and MMSE block linear equalizers that are suited for each of the block transmission techniques and introduce the block decision-feedback equalizer in the context of KSP. We then carry out a detailed analysis of the achievable BER performance of these block transmission techniques and their associated equalization schemes in the light of existing WLAN standards (IEEE 802.11a and Hiperlan2).

As mentioned above, the knowledge of the short sequences appended by the KSP transmission scheme can be exploited towards channel identification by the receiver. In **Ch. 4**, we propose a new semi-blind Gaussian Maximum Likelihood (ML) method for the identification of the transmission channel which is suited to the context of KSP transmission. This method outperforms existing semi-blind methods in this context and has the advantage of a low computational complexity. A Cramer-Rao Bound (CRB) study, which we present in this chapter, indicates that when the padded sequence is changed after each KSP block, this results in significantly improved channel estimates. However, when these changing sequences are used, it is not possible to use the low-complexity FD equalizers anymore. To solve this problem, we propose a new KSP transmission scheme, namely *Shifted KSP (S-KSP)*. At the cost of an extra redundant symbol, this scheme allows to use the low-complexity FD equalizers whilst accurate channel estimates can still be obtained as the padded sequence changes from block to block.

The publications that are associated to this chapter are the following:

- O. Rousseaux, G. Leus, P. Stoica and M. Moonen, “Gaussian Maximum Likelihood Sequence Estimation with Short Training Sequences,” to appear in *IEEE Transactions on Wireless Communications*, 2004.

- O. Rousseaux, G. Leus and M. Moonen, “Block Transmission and Shifted Known Symbol Padding for Efficient Data Communication in a WLAN Context”, Dec 2004, submitted to *Wireless Personal Communications*.
- O. Rousseaux, G. Leus, P. Stoica and M. Moonen, “A Stochastic Method for Training Based Channel Identification,” in *Proc. of the Seventh International Symposium on Signal Processing and its Applications (ISSPA 2003)*, July 2003, Paris, France, pp. 657-660.
- O. Rousseaux, G. Leus, P. Stoica and M. Moonen, “Training Based Maximum Likelihood Channel Identification: Constant Training Sequences,” in *Proc. of the IEEE Symposium on Signal Processing Advances in Wireless Communications (SPAWC 2003)*, June 2003, Rome, Italy, pp. 334-338.
- O. Rousseaux, G. Leus, P. Stoica and M. Moonen, “Generalized Training Based Channel Identification,” in *Proc. of the Global Conference on Communications (GLOBECOM 2003)*, December 2003, San Francisco, California, pp. 2432-2436.

In **Ch. 5**, we propose a new method that directly aims at identifying the transmitted data symbols in the context of single-carrier block transmission techniques (i.e. SC-CP or KSP). This method proves to be useful mainly when there are no pilots or training symbols present in the signal and the channel cannot be estimated accurately with classical channel estimation algorithms. This method belongs to the family of Blind Maximum Likelihood Sequence Estimation (MLSE) methods and the approach can be considered as an Iterative Least Squares with Projections (ILSP) approach.

The publications that are associated to this chapter are the following:

- O. Rousseaux, G. Leus, M. Moonen, “A suboptimal Iterative Method for Maximum Likelihood Sequence Estimation in a Multipath Context,” in *EURASIP Journal on Applied Signal Processing (JASP)*, vol. 2002, no. 12, pp. 1437-1447, Dec. 2002.
- O. Rousseaux, G. Leus, M. Moonen, “An Iterative Procedure for Semi-Blind Symbol Estimation in a Multipath SISO Channel Context Exploiting Finite Alphabet Properties,” in *Proc. of the International Zurich Seminar on Broadband Communications (IZS 2002)*, Feb 2002, Zurich, Switzerland, pp. 21.1-21.5.

Finally, in **Ch. 6**, we consider the situation where the mobility of the users is increased. The channel then becomes time-varying and changes significantly during the transmission of a packet of data symbols. Relying on a newly proposed technique for the modeling of time-varying multipath channels, namely

the *Basis Expansion Model (BEM)*, we propose a novel method that allows to identify the BEM coefficients relying on the padded sequences of the KSP scheme. This method seems to outperform existing ones in the considered context but has a slightly higher computational complexity than the one proposed in the context of stationary channels. We adapt existing KSP equalizers to this context of time-varying channels. We then carry out a detailed analysis of the overall system performance when these equalizers are used in conjunction with the proposed channel estimates in a realistic simulation environment. The simulation results show that KSP is indeed a suitable candidate for data transmission in the context of time-varying channels.

The publications that are associated to this chapter are the following:

- O. Rousseaux, G. Leus and M. Moonen, “Estimation and Equalization of Doubly Selective Channels using Known Symbol Padding,” to appear in *IEEE Transactions on Signal Processing*, 2004.
- O. Rousseaux, G. Leus and M. Moonen, “An Iterative Method for Improved Training-Based Estimation of Doubly Selective Channels,” in *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2004)*, May 2004, Montreal, Canada, pp. iv.889-iv.892.

Besides the material presented in the thesis, we also published or contributed to the following papers during our PhD research:

- O. Rousseaux, G. Leus and M. Moonen, “A Blind Receiver for Block Transmission in a Multi-User MIMO Context with Multipath,” in *Proc. of the Signal Processing Symposium (SPS2002)*, Mar. 2002, Leuven, Belgium, pp. 33-36.
- O. Rousseaux, G. Leus and M. Moonen, “A Reduced Complexity Deterministic Blind Transceiver with Space- Only Block Coding in a Multi-User MIMO Context with Severe Multipath,” in *Proc. of the Benelux Meeting on Systems and Controls*, March 2001, Houffalize, Belgium.
- O. Rousseaux, G. Leus and M. Moonen, “A Blind Multi-User MIMO Transceiver Using Code Modulation in a Multipath Context,” in *Proc. of the 14th Conference on Digital Signal Processing 2002 (DSP 2002)*, July 2002, Santorini, Greece, pp.67-270.
- R. Cendrillon, O. Rousseaux, M. Moonen, E. Van den Bogaert, J. Verlinden, “Simplified Power Allocation for the DSL Multi-Access Channel through Column-Wise Diagonal Dominance,” in *Proc. of 24th Symposium on Information Theory in the Benelux*, May 2003, Netherlands.
- G. Leus, I. Barhumi, O. Rousseaux and M. Moonen, “Direct Semi-Blind Design of Serial Linear Equalizers for Doubly-Selective Channels,” in

Proc. of the International Conference on Communications (ICC 2004),
June 2004, Paris, France, pp. 2626-2630.

Chapter 2

Background Material

Before presenting the main contributions of our research, we present in this chapter existing concepts and ideas upon which the rest of the text relies.

We first briefly present how wireless transceivers are organized, describing the *functional architecture* of the transmitter and the receiver. We then show how the mathematical model that relates the channel output samples to the channel input sequence is derived from this functional architecture. We also show how the design parameters of the transmission system impact the resulting *channel model*, with a special attention to broadband communication systems where the multipath effects of the physical channel result in a frequency-selective channel model that causes *Inter Symbol Interference (ISI)* in the transmission scheme.

This ISI is a major impediment of broadband communication systems and needs to be combatted at the receiver by appropriate means. Several block transmission techniques have been proposed recently in order to allow the receiver to cope with this ISI in an efficient way within a limited computational complexity. In the second part of this chapter, we review these *block transmission techniques* in the general framework of Affine Precoding [14] and show how they design the channel input sequence. We also describe the organization of the receiver for these different block transmission techniques and, relying on the earlier derived channel models, describe their specific data models.

The organization and functional architecture of digital wireless transceivers is described in *sec. 2.1.1* and the resulting channel model is derived in *sec. 2.1.2*. Block transmission techniques and their corresponding data models are introduced in *sec. 2.2*.

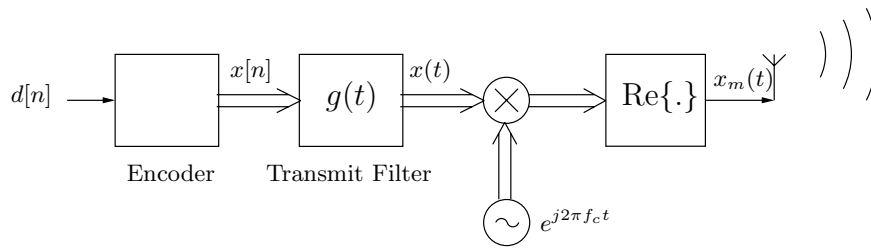


Figure 2.1: Structure of a typical wireless transmitter. Single arrows indicate real scalars, double arrows indicate complex numbers.

2.1 Introduction to Digital Wireless Communications

Digital wireless communication systems aim at transmitting binary data from a transmitter to a receiver using electromagnetic waves that are propagated over the air interface. Whilst the transmitted binary information is inherently discrete-time, the transmission channel is continuous-time in nature. Hence, digital communication systems need to represent the stream of data bits with a continuous-time signal to allow the transmission of information over the air interface. Moreover, as wireless systems are usually allowed to use a limited portion of the spectrum of the air interface, the transmitted signal must be designed to fit within the allowed spectrum, which is characterized by its bandwidth W and its central frequency, or *carrier frequency* f_c . We outline below how classical wireless transmission systems meet these requirements.

2.1.1 System Setup

In this section, we give a functional overview of the organization of traditional digital wireless transmitters and receivers. We further describe the signals that are present at the different stages of the transceiver.

Transmitter Structure The architecture of a typical wireless transmitter is depicted in **Fig. 2.1**. Let $d[n]$ be the sequence of information bits that will undergo the transmission. This sequence is fed to an encoder that maps the logical bits onto an information-bearing sequence of complex numbers $x[n]$, the *channel input sequence*. This (discrete-time) sequence of information symbols is fed to a *pulse-shaping transmit filter* with impulse response $g(t)$, at the rate $1/T_s$. This results in the (continuous-time) complex baseband signal $x(t)$. The bandwidth occupied by $x(t)$ depends on the specific choice of the pulse $g(t)$. In

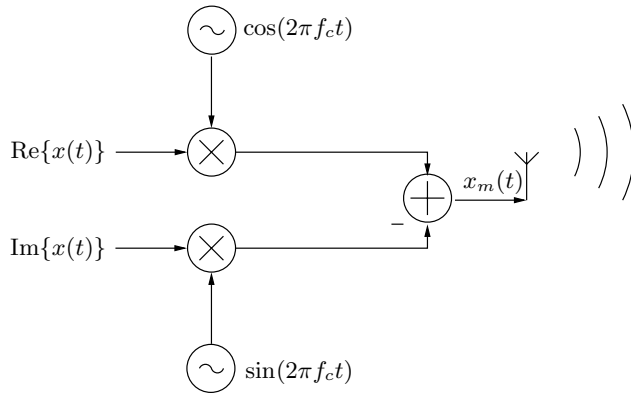


Figure 2.2: Equivalent transmitter architecture. The in-phase and in-quadrature components of the transmitted signal appear to be respectively the real and imaginary parts of the baseband signal.

order to satisfy the Nyquist criterion and allow the reconstruction of $x[n]$ from the signal $x(t)$, $g(t)$ must at least have a total bandwidth $W = 1/T_s$. Most practical systems use real-valued pulses. As the bandwidth is a limited resource, transmit pulses are designed in order to use the smallest possible bandwidth for a given symbol rate. The minimum bandwidth is obtained with a *sinc* pulse with a central lobe of duration $2T_s$. However, as the practical implementation of such a pulse can be problematic, real-life systems often rely on alternative pulse shapes that have some excess bandwidth. The *raised-cosine* pulse is a typical example that is used in many practical systems:

$$g(t) = \frac{\sin(\pi t/T_s)}{\pi t/T_s} \frac{\cos(\alpha \pi t/T_s)}{1 - (2\alpha t/T_s)^2}, \quad (2.1)$$

where the parameter α ranges from 0 to 1. The total bandwidth of this pulse is $W = (1 + \alpha)/T_s$ and the parameter α , called the rolloff factor, represents the excess bandwidth of that pulse. The time- and frequency-domain description of such raised-cosine pulses is shown in **Fig. 2.3**. Larger values of α yield a faster decay of the time-domain pulse (and hence a less complex implementation).

The complex baseband signal is then modulated on the carrier frequency, i.e. it is multiplied by a complex exponential $e^{j2\pi f_c t}$. This multiplication shifts the spectrum of the complex baseband signal, which becomes centered around the carrier frequency f_c rather than around 0. The real part of the resulting signal, $x_m(t)$, is transmitted by the antenna over the air interface and is called the passband signal. The effect of these different steps on the signal spectrum is depicted in **Fig. 2.4**.

This modulation scheme can equivalently be described by the scheme presented

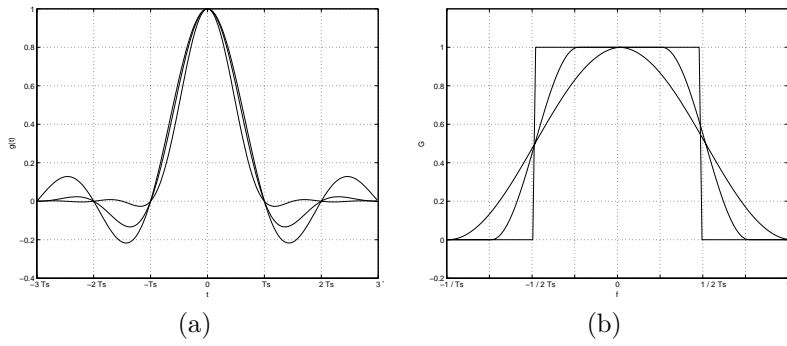


Figure 2.3: Raised cosine pulses for different values of α (0, 0.5 and 1). The time-domain pulse is shown in (a) and its spectrum is displayed in (b).

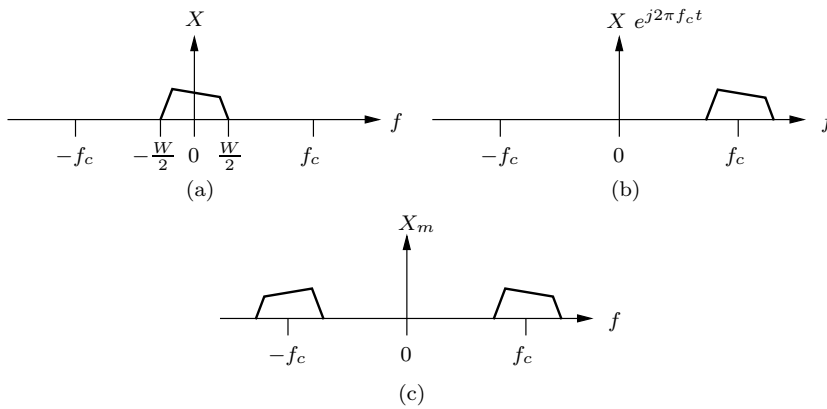


Figure 2.4: Spectral contents of the signal at different stages of the transmitter. The baseband signal (a) has a total bandwidth W centered around DC. After the multiplication with the complex exponential (b), the spectrum is centered around f_c . Finally, taking the real part of the signal (c) creates a replica of the spectrum in the negative frequencies.

in **Fig. 2.2**, where it appears more clearly that the real (resp. imaginary) part of the complex baseband signal is the in-phase (resp. quadrature) component of the transmitted signal $x_m(t)$.

Receiver Structure The task of the receiver is to allow the recovery of the complex baseband signal $x(t)$ from the received signal $y_m(t)$. As the air interface is generally shared by multiple applications and users operating at different carrier frequencies, the signal that is present at the output of the receive antenna, $y_m(t)$, generally contains multiple spectral components. The signal of interest consists of two lobes centered around f_c and $-f_c$, whilst the rest of the spectrum consists of out-of-band noise and interference. The lobe around f_c is the signal of interest as it is an image of the baseband signal's spectrum. Several receiver architectures are possible in order to recover this signal of interest. We present one of the possible architectures in **Fig. 2.5**. The spectrum of the different signals present in this receiver architecture is shown in **Fig. 2.6**, where the part of the received signal originating from the transmitter is in bold on the figure. Multiplying the received signal with a complex exponential of frequency $-f_c$ shifts the whole spectrum to the left. The spectrum of the baseband signal is located in a bandwidth W centered around DC. Feeding the signal to an ideal low-pass filter allows to remove the undesired higher frequency components from the signal. It is then possible to recover the complex baseband signal $x(t)$ from the filtered signal $y(t)$, which is formally obtained as:

$$y(t) = y_m(t)e^{-j2\pi f_c t} \otimes f(t). \quad (2.2)$$

The output of the low-pass filter, which is called the *received passband signal* is then sampled at an appropriate rate. The resulting sequence is then processed by the decoder that estimates the transmitted data. We show below how the signals present at the different stages of such a receiver are formally described.

2.1.2 Channel Model

During the propagation, the transmitted waves are reflected or scattered by the objects that are placed in their path. Several reflected and scattered versions of the transmitted waveform arrive at the receive antenna. The received waveform is thus the superposition of multiple replicas of the emitted waveform and hence, the received signal is generally a distorted version of the transmitted one. The effects of the transmission channel are described by the channel's impulse response¹ $b(t)$.

¹We describe here the situation where the transmission channel has a stationary impulse response. The situation where the channel varies rapidly is studied in *Ch.6* of this thesis where we describe an alternative channel model that is suited to the time-varying channel situation.

Passband Model The received signal $y_m(t)$ is the linear convolution of $x_m(t)$ with the impulse response $b(t)$ of the physical channel, plus some background noise $\eta_m(t)$:

$$y_m(t) = x_m(t) \otimes b(t) + \eta_m(t). \quad (2.3)$$

Baseband Model Defining the equivalent baseband impulse response of the physical channel as $b_E(t) = e^{-j2\pi f_c t} b(t)$ and $\psi(t)$ as the compound impulse response of the transmit and receive filters ($\psi(t) = g(t) \otimes f(t)$), the equivalent baseband channel model (also called equivalent baseband pulse) is defined as:

$$h(t) = b_E(t) \otimes \psi(t). \quad (2.4)$$

The received passband signal can then be expressed as:

$$y(t) = \sum_{k=-\infty}^{+\infty} x[k] h(t - kT_s) + \eta(t), \quad (2.5)$$

where $\eta(t) = \eta_m(t)e^{-j2\pi f_c t}$. Note that the equivalent baseband impulse response $b_E(t)$ of the physical channel is the response of the channel to a signal with frequency components close to the carrier frequency f_c . It is generally a complex-valued function consisting of several discrete spikes that have their own delay, phase and amplitude response. Each of these spikes typically corresponds to a single reflection or scattering in the transmission channel.

The *Maximal Delay Spread* τ_{max} of the physical channel is the maximal delay between the first and the last received replica of the transmitted signal. It is generally a fixed quantity for a given environment. Typical values for the maximal delay spread of a channel range from 50 ns (office environment) to 250 ns (large open space) [15]. The value of the maximal delay spread has a significant impact on the shape of the equivalent baseband pulse $h(t)$.

When $\tau_{max} \ll T_s$, the equivalent impulse response of the physical channel is practically equivalent to an impulse and the equivalent baseband pulse is approximately equal to a scaled version of the transmit pulse. In this case, the transmission channel is called a *flat fading channel* since the Fourier transform of its impulse response is essentially flat within the channel bandwidth W .

When $\tau_{max} \approx T_s$ or $\tau_{max} > T_s$, the equivalent baseband pulse is a significantly distorted version of the transmit pulse. The transmission channel is then called a *frequency selective channel* since its frequency response varies significantly within the channel bandwidth W .

For a given physical channel, the type of equivalent baseband channel model will thus depend on the symbol rate $1/T_s$. This is illustrated in **Fig.2.7** where the equivalent baseband channel model corresponding to the same physical channel is shown when different symbol rates (and hence different front-end filters $f(t)$)

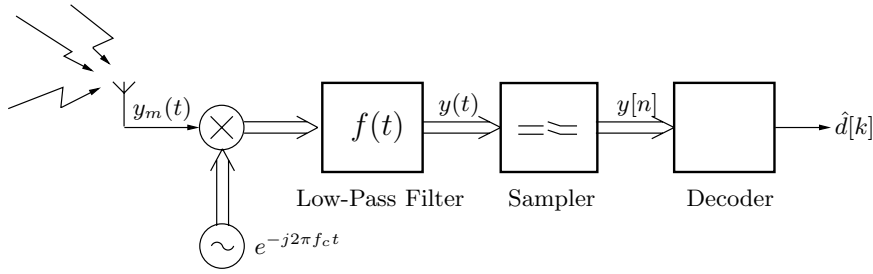


Figure 2.5: Example of a wireless receiver architecture. Single arrows indicate scalars, double arrows indicate complex numbers.

and $g(t)$) are considered. The equivalent baseband impulse response of the considered physical channel $b_E(t)$ consists in three fixed discrete pulses with a given maximal delay spread τ_{max} . These figures show that when T_s is relatively large, the resulting channel model is a flat-fading one and, as the datarate is increased and T_s gets smaller, the channel becomes frequency-selective.

Discrete Baseband Model Sampling the baseband signal $y(t)$ at the symbol rate T_s , we obtain the discrete sequence of complex numbers $y[n]$ defined as $y[n] = y(nT_s)$ that we call the sequence of *channel output samples*². They are related to the channel input sequence $x[n]$ by the following relationship:

$$y[n] = \sum_{k=-\infty}^{+\infty} x[n-k]h(kT_s) + \eta[n], \quad (2.6)$$

where the discrete noise sequence $\eta[n]$ is defined as $\eta[n] = \eta(nT_s)$. Note that the full expression of the equivalent baseband channel model $h(t)$ is not needed to derive this expression; its value at the sampling instants is sufficient. We therefore define the *discrete channel model* as $\mathbf{h} : h[k] = h(kT_s)$. Taking into account the limited time-span of the transmit pulse and receive filter, the finite delay spread of the physical channel and the causality of the transmission process, the discrete channel model has a limited number of non-zero elements: $\mathbf{h} = [h[0], h[1], \dots, h[L]]$, where L is called the *order* of the transmission channel.

²Note that when the pulse-shaping transmit filter has a non-zero excess bandwidth α , $y(t)$ should be sampled at the rate $\frac{(1+\alpha)}{T_s}$ in order to fulfill the Nyquist criterion. Sampling at the rate $1/T_s$ results in a sub-sampled system and so, part of the information contained in $y(t)$ is lost in the sampling process. In order to exploit all the received information, some systems use fractional sampling, i.e, they sample $y(t)$ at the rate P/T_s , where P is an integer. This process results in a SIMO model and the correlation between the different channels depends on the excess bandwidth (e.g. a system where $P = 2$ has uncorrelated subchannels if $\alpha = 1$, while the subchannels are correlated if $\alpha \leq 1$)

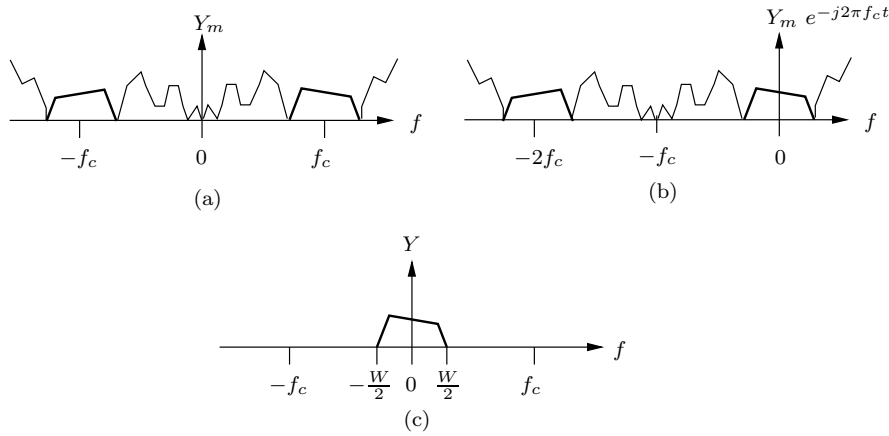


Figure 2.6: Spectral contents of the signal at different stages of the presented receiver. The spectrum of the received signal (a) is shifted to the left (b) after the multiplication with the complex exponential. The only signal left after low-pass filtering (c) is the signal of interest.

The discrete channel models for flat-fading and frequency-selective channels are presented in **Fig.2.7**.

When flat-fading channels are considered and the sampling delay is chosen properly, the discrete channel model is essentially a pulse and the channel order is $L = 0$. In this case, $y[n]$ is simply a scaled and noisy version of the transmitted data symbol $x[n]$. When frequency-selective channels are considered, the situation changes and $L \neq 0$. In this case, each channel output sample $y[n]$ is a function of the last $L + 1$ transmitted data symbols. The transmitted data symbols are thus interfering with each other; this phenomenon is called *Inter-Symbol Interference (ISI)*. Flat fading channels are also called *ISI-free channels*.

In the work we present here, we primarily focus on communication systems that aim at delivering high data rates, i.e. rates that are at least equal to those offered by the current standards for Wireless Local Area Networks (WLANs) such as IEEE802.11a or Hiperlan2. The symbol period for these two standards is set to $T_s = 50$ ns and we will thus focus on the situation where $T_s \leq 50$ ns. Considering that realistic values for the maximal delay spread of the physical channel range from 50 ns to 250 ns, we have to rely on the data models presented for frequency-selective channels where ISI is present in the transmission scheme. Hence, the channel will be modeled as a FIR filter of order L :

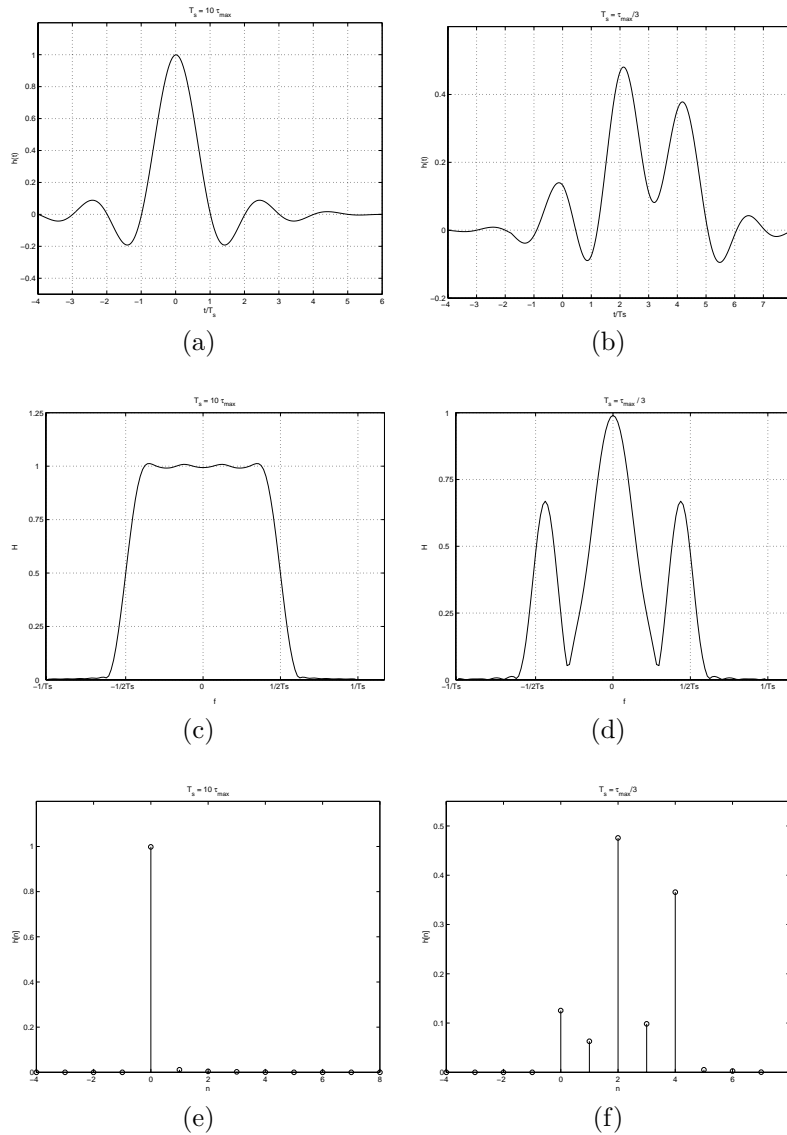


Figure 2.7: Equivalent baseband channel model for the same physical channel (fixed τ_{max}) but with different symbol rates. A low symbol rate ($T_s = 10\tau_{max}$) results in a flat-fading channel (a, c, e). A higher symbol rate ($T_s = \tau_{max}/3$) results in a frequency-selective channel (b, d, f). The resulting time-domain equivalent baseband pulses are shown in (a) and (b), their spectrum in (c) and (d) and their corresponding discrete channel model in (e) and (f).

$$y[n] = \sum_{l=0}^L x[n-l]h[l] + \eta[n]. \quad (2.7)$$

Different statistical models are presented in the literature to characterize the distribution of the channel coefficients. When there is a direct line of sight between the transmitter and the receiver, the first channel tap $h[0]$ will be dominant and its p.d.f. will be characterized by a *Ricean distribution*. This Ricean distribution also characterizes channels where a reflected or scattered wave is clearly dominant. When there is no line of sight and no such dominant wave between the transmitter and the receiver, the received signal results from the random superposition of reflected and scattered waves that can be characterized by a *Rayleigh* distribution, a situation which is referred to as the *Rayleigh-faded channel*. In this case, the channel tap is described with a zero-mean Gaussian p.d.f. with a fixed variance. In this thesis, we consider the situation where all the channel taps are Rayleigh faded. Moreover we consider that the different taps of the channel are uncorrelated, and that they all have the same variance, i.e. the different channel taps are i.i.d. Gaussian variables with a flat power profile:

$$\mathbb{E} \{ \mathbf{h} \mathbf{h}^H \} = \gamma^2 \mathbf{I}, \quad (2.8)$$

where γ is a constant representing the variance of the Gaussian process.

The noise term has two different origins. It is partly caused by the RF stage of the receiver where the thermal noise of the power amplifiers or the quantization noise of the ADC imposes a noise on the received signal. This first term is generally described as a zero-mean white Gaussian process. Moreover, when the air interface is shared by different users, the noise term accounts as well for the interference of the interfering users. This term is generally considered to be Gaussian as well, but, depending on the type of interfering signal, the resulting noise can either be white or colored. We consider here the situation where the noise is white, the extension to the colored noise situation being quite straightforward in most of the presented applications.

$$\mathbb{E} \{ \eta[n] \eta[k]^* \} = \delta_{k,n} \sigma^2 \mathbf{I}, \quad (2.9)$$

where σ^2 is the noise variance.

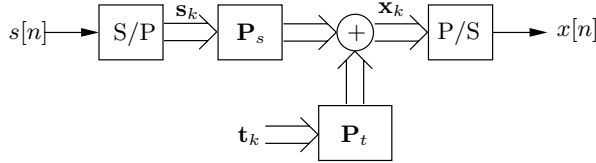


Figure 2.8: Organization of the transmitter in a general block transmission scheme

2.2 Block Transmission Techniques: Data Model

The channel model presented above describes the relationship that binds the channel output samples to the channel input sequence. However, we did not detail how the transmitter processes the binary data sequence $d[n]$ in order to obtain the sequence $x[n]$. The way this channel input sequence is obtained depends on the adopted transmission scheme, which translates in a specific structure for the encoder (see **Fig 2.1**). We describe below how the emitter processes a packet of data symbols $d[n]$ before it is transmitted over the channel. The packet of data is first mapped onto a sequence of complex-valued data symbols $s[n]$ by means of a constellation mapper. These data symbols can take their values within a finite set of discrete values that depend on the chosen constellation (BPSK, QPSK, 16-QAM, 8-PSK, PAM, ...). They thus have finite alphabet properties, each element of the alphabet corresponding to a logical sequence of data bits. Most classical transmission schemes directly transmit the sequence of data symbols over the channel, i.e. $x[n] = s[n]$, the channel input sequence is the sequence of data symbols. In block transmission, the transmitted data symbols are organized in blocks and fed to a precoding matrix before transmission. Training symbols can be mixed with data symbols, which is equivalent to affine precoding [14].

A block $\mathbf{s}_k(N_s \times 1)$ of data symbols is defined as a column vector of size N_s , $k = 1 \cdots K$ being the block index (K is the total number of data symbol blocks present in a packet of data):

$$\mathbf{s}_k = [s_k[1], \dots, s_k[N_s]]^T. \quad (2.10)$$

A block $\mathbf{t}_k(N_t \times 1)$ of training symbols is defined similarly:

$$\mathbf{t}_k = [t_k[1], \dots, t_k[N_t]]^T. \quad (2.11)$$

The k^{th} block of channel input symbols, $\mathbf{x}_k(N_x \times 1)$ is then computed as follows:

$$\mathbf{x}_k = \mathbf{P}_s \mathbf{s}_k + \mathbf{P}_t \mathbf{t}_k, \quad (2.12)$$

where $\mathbf{P}_s(N_x \times N_s)$ is the precoding matrix for data symbols and $\mathbf{P}_t(N_x \times N_t)$ is the precoding matrix for training symbols. The transmitted channel input sequence is the concatenation of all these blocks: $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_K^T]^T$. The transmitted packet thus contains $K N_s$ data symbols that are mapped onto a channel input sequence of length $K N_x$. The resulting transmitter scheme is presented in **Fig. 2.8**. We assume that the total transmission time of a packet of data is significantly smaller than the coherence time of the transmission channel, i.e. $KN_x T_s \ll \tau_{coh}$. Under this hypothesis, we can use the discrete baseband model proposed in *Ch. 1*.

We thus consider a stationary FIR convolutive channel of order L :

$$\mathbf{h} = [h[0], h[1], \dots, h[L]]^T. \quad (2.13)$$

The received sequence, $y[n]$ is the linear convolution of the channel input sequence with the channel impulse response:

$$y[n] = \sum_{l=0}^L h[l]x[n-l] + \eta[n], \quad (2.14)$$

where $\eta[n]$ is the Additive White Gaussian Noise (AWGN) at the receiver. This sequence of channel output samples can be organized in blocks $\mathbf{y}_k(N_x \times 1)$ of received symbols corresponding to the blocks of transmitted symbols:

$$\mathbf{y}_k = [y[(k-1)N_x + 1], \dots, y[kN_x]]^T. \quad (2.15)$$

Using these definitions, the transmission scheme can be expressed on a block level: the k^{th} received block contains contributions from the k^{th} transmitted block and noise plus some interference from the $(k-1)^{st}$ transmitted block (assuming $N_x > L$)³:

$$\mathbf{y}_k = \mathbf{H}_0 \mathbf{x}_k + \mathbf{H}_1 \mathbf{x}_{k-1} + \boldsymbol{\eta}_k, \quad (2.16)$$

where $\boldsymbol{\eta}_k$ is the AWGN vector with noise power σ^2 ($\mathbb{E}\{\boldsymbol{\eta}_k \boldsymbol{\eta}_k^H\} = \sigma^2 \mathbf{I}$) and the channel matrices $\mathbf{H}_0(N_x \times N_x)$ and $\mathbf{H}_1(N_x \times N_x)$ are the following Toeplitz matrices:

³We assume that the number of data symbols is chosen to be larger than the channel order. If this was not the case, the $(k-2)^{nd}$ block (and eventually other blocks) would cause interfere into \mathbf{y}_k as well

$$\mathbf{H}_0 = \begin{bmatrix} h[0] & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \\ h[L] & & \ddots & \ddots & \vdots \\ 0 & \ddots & & \ddots & \ddots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & h[L] & \cdots & h[0] \end{bmatrix}, \quad (2.17)$$

$$\mathbf{H}_1 = \begin{bmatrix} 0 & \cdots & 0 & h[L] & \cdots & h[1] \\ & & & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & h[L] \\ \vdots & & & & & 0 \\ & & & & & \vdots \\ 0 & \cdots & \cdots & & & 0 \end{bmatrix}. \quad (2.18)$$

Equalization or symbol detection will usually be done after a preprocessing step on the received block, which yields the vector $\mathbf{z}_k(N_z \times 1)$:

$$\mathbf{z}_k = \mathbf{R}\mathbf{y}_k, \quad (2.19)$$

where $\mathbf{R}(N_z \times N_x)$ is the preprocessing matrix at the receiver. The general scheme for block transmission can then be described by the following equation:

$$\boxed{\mathbf{z}_k = \mathbf{R}\mathbf{H}_0\mathbf{P}_s\mathbf{s}_k + \mathbf{R}\mathbf{H}_1\mathbf{P}_s\mathbf{s}_{k-1} + \mathbf{R}\mathbf{H}_0\mathbf{P}_t\mathbf{t}_k + \mathbf{R}\mathbf{H}_1\mathbf{P}_t\mathbf{t}_{k-1} + \mathbf{R}\boldsymbol{\eta}_k.} \quad (2.20)$$

Traditional block transmission techniques aim at avoiding Inter Block Interference (IBI, interference between two successive blocks of data). Formally, this is obtained if the precoding and preprocessing matrices \mathbf{P}_s and \mathbf{R} are designed s.t. $\mathbf{R}\mathbf{H}_1\mathbf{P}_s = \mathbf{0}$. This is usually done by inserting a guard band (a short sequence of symbols) between two consecutive blocks of data symbols. If the guard band duration is larger than the channel order, it collects all the interference from the preceding block of data symbols and so IBI is avoided. We investigate here two families of block transmission techniques that are commonly used in practical implementations: Cyclic Prefix (CP) transmission and Known Symbol Padding (KSP) transmission, which differ in the composition of the guard band.

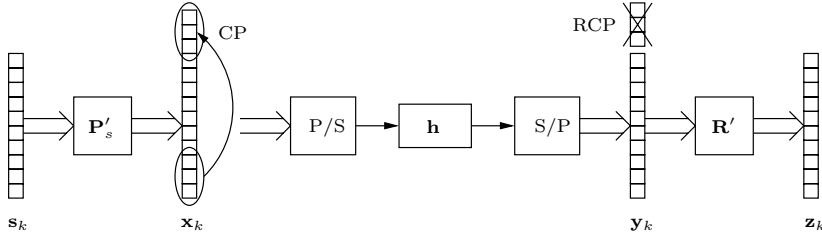


Figure 2.9: Organization of the transmitter in CP block transmission.

2.2.1 Cyclic Prefix Transmission

Cyclic Prefix (CP) techniques avoid IBI by appending a cyclic prefix of length μ to the beginning of each block of data symbols. As illustrated in **Fig. 2.9** where the general organization of a CP transceiver is depicted, this cyclic prefix is a copy of the last symbols of the block. In order to avoid the IBI, the length of the CP must be larger than the channel order ($\mu \geq L$). In that case, the CP collects all the interference from the data symbols of the previous block and is discarded at the receiver. In the presented framework, appending the CP translates into precoding matrices of the form:

$$\mathbf{P}_s = \mathbf{P}_\mu \mathbf{P}'_s, \quad \mathbf{P}_\mu = \left[\begin{array}{c|c} \mathbf{0}_{\mu \times (N_x - 2\mu)} & \mathbf{I}_\mu \\ \hline & \mathbf{I}_{N_x - \mu} \end{array} \right], \quad (2.21)$$

where \mathbf{P}'_s is a general precoding matrix of size $(N_x - \mu) \times N_s$. Note that no training symbols are mixed with the data symbols in the considered CP schemes ($\mathbf{t}_k = \mathbf{0}$). Discarding the CP at the receiver translates into preprocessing matrices of the form:

$$\mathbf{R} = \mathbf{R}' \mathbf{R}_\mu, \quad \mathbf{R}_\mu = \left[\begin{array}{c|c} \mathbf{0}_{(N_x - \mu) \times \mu} & \mathbf{I}_{N_x - \mu} \end{array} \right]. \quad (2.22)$$

where \mathbf{R}' is a general preprocessing matrix of size $N_z \times (N_x - \mu)$.

Using these precoding and preprocessing matrices, it is easy to check that, when the condition $\mu \geq L$ is respected, CP transmission yields two useful properties:

- The channel matrix effectively becomes a circulant matrix of size $(N_x - \mu) \times (N_x - \mu)$:

$$\mathbf{R}_\mu \mathbf{H}_0 \mathbf{P}_\mu = \mathbf{H}_{circ} = \begin{bmatrix} h[0] & 0 & \cdots & 0 & h[L] & \cdots & h[1] \\ \vdots & \ddots & & & & \ddots & \vdots \\ h[L] & & \ddots & & & & h[L] \\ 0 & \ddots & & \ddots & & & 0 \\ \vdots & & \ddots & & \ddots & & \vdots \\ \vdots & & & \ddots & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & h[L] & \cdots & h[0] \end{bmatrix}, \quad (2.23)$$

- The IBI is cancelled:

$$\mathbf{R}_\mu \mathbf{H}_1 \mathbf{P}_\mu = \mathbf{0}_{(N_x - \mu) \times (N_x - \mu)}. \quad (2.24)$$

Keeping in mind that $\mathbf{t}_k = \mathbf{0}$, the transmission scheme (2.20) can then be rewritten as:

$$\mathbf{z}_k = \mathbf{R}' \mathbf{H}_{circ} \mathbf{P}_s' \mathbf{s}_k + \mathbf{R}' \mathbf{n}_k, \quad (2.25)$$

where $\mathbf{n}_k = \boldsymbol{\eta}_k(\mu + 1 : N_z)$.

A nice property of circulant matrices is that they are diagonalized by the use of DFT and IDFT operations:

$$\mathbf{H}_{circ} = \mathcal{I}_P \mathbf{H}_{diag} \mathcal{F}_P, \quad (2.26)$$

where \mathcal{I}_P is the normalized IDFT matrix of size $P = N_x - \mu$, \mathcal{F}_P is the normalized DFT matrix of the same size and \mathbf{H}_{diag} is a diagonal matrix with the P-point DFT of the channel impulse response \mathbf{h} on the main diagonal:

$$\mathbf{H}_{diag} = \text{diag}(\sqrt{P} \mathcal{F}_P [\mathbf{h} \ 0 \ \cdots \ 0]^T). \quad (2.27)$$

This property can be used for implementing efficient and cheap equalization schemes that will be discussed in further details in the next chapter.

Two popular block transmission systems fall under this category of CP systems: **OFDM** systems and **SC-CP** systems. They differ in the way that the data are precoded before the CP is added (i.e., \mathbf{P}_s') and preprocessed before equalization (i.e., \mathbf{R}').

Orthogonal Frequency Division Multiplexing Systems (OFDM):

Also known as **MC-CP** (Multi-Carrier with Cyclic Prefix), these systems were the first block transmission techniques used in practical systems. They directly exploit the circulant structure of the channel matrix of CP systems. In OFDM, an IDFT is performed on data blocks of size N_s before the CP is added (i.e., $\mathbf{P}'_s = \mathcal{I}_P$ with $P = N_s$). The receiver throws away the CP and performs a DFT on the received block (i.e., $\mathbf{R}' = \mathcal{F}_P$ with $P = N_s$). OFDM systems can thus be described as follows:

$$\mathbf{z}_k = \mathcal{F}_P \mathbf{H}_{circ} \mathcal{I}_P \mathbf{s}_k + \mathcal{F}_P \mathbf{n}_k, \quad (2.28)$$

Recalling (2.26) and exploiting the fact that \mathcal{I}_P is the inverse of \mathcal{F}_P , this system can be further simplified into:

$$\boxed{\mathbf{z}_k = \mathbf{H}_{diag} \mathbf{s}_k + \mathcal{F}_P \mathbf{n}_k.} \quad (2.29)$$

This transmission scheme allows to transform the initial convolutive channel into a set of parallel flat-fading channels, and hence to get rid of the ISI using only simple DFT and IDFT operations. The effect of the IDFT performed at the transmitter is to encode the data symbols on independent orthogonal tones. The CP allows to maintain the orthogonality between these tones during the transmission over the convolutive channel. The data symbol received on a given tone is simply a noisy version of the transmitted data symbol multiplied by a complex scalar corresponding to the channel frequency response on the considered tone.

Single Carrier Cyclic Prefix Systems (SC-CP):

SC-CP directly adds a CP of length μ to the transmitted data symbol block without any extra encoding ($\mathbf{P}'_s = \mathbf{I}_{N_s}$). The CP is discarded at the receiver without preprocessing ($\mathbf{R}' = \mathbf{I}_{N_s}$), which results in the following transmission scheme:

$$\boxed{\mathbf{z}_k = \mathbf{H}_{circ} \mathbf{s}_k + \mathbf{n}_k.} \quad (2.30)$$

Note that SC-CP transmission is equivalent to an OFDM transmission scheme where the data are encoded across the tones by means of a DFT.

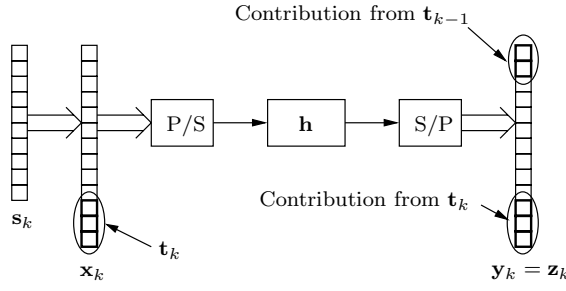


Figure 2.10: Organization of the transmitter in KSP block transmission.

2.2.2 KSP Transmission

In KSP transmission, IBI is avoided by padding blocks of known symbols between data blocks. The k^{th} transmitted block is made of N_s data symbols followed by N_t known symbols: $\mathbf{x}_k = [\mathbf{s}_k^T, \mathbf{t}_k^T]^T$. This translates into precoding matrices of the form:

$$\mathbf{P}_{s(N_s+N_t) \times N_s} = \begin{bmatrix} \mathbf{I}_{N_s} \\ \mathbf{0}_{N_t \times N_s} \end{bmatrix}, \quad \mathbf{P}_{t(N_s+N_t) \times N_t} = \begin{bmatrix} \mathbf{0}_{N_s \times N_t} \\ \mathbf{I}_{N_t} \end{bmatrix}. \quad (2.31)$$

No preprocessing is done at the receiver, which translates into an identity preprocessing matrix: $\mathbf{R} = \mathbf{I}_{N_s+N_t}$. In order to avoid interference into the following block of data symbols, the length of the padded sequence must be chosen such that $N_t \geq L$. In that case, it is easy to check that $\mathbf{H}_1 \mathbf{P}_s = \mathbf{0}_{(N_s+N_t) \times N_s}$, which means that there is no IBI. Using these precoding and preprocessing matrices, the total transmission scheme (2.20) translates into:

$$\mathbf{z}_k = \mathbf{H}_{KSP} \mathbf{s}_k + \mathbf{H}_{t0} \mathbf{t}_k + \mathbf{H}_{t1} \mathbf{t}_{k-1} + \boldsymbol{\eta}_k, \quad (2.32)$$

where

$$\mathbf{H}_{KSP(N_s+N_t) \times N_s} = \begin{bmatrix} h[0] & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ h[L] & & \ddots & 0 \\ 0 & \ddots & \ddots & h[0] \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h[L] \\ \hline & & & \mathbf{O}_{(N_t-L) \times N_s} \end{bmatrix}, \quad (2.33)$$

\mathbf{H}_{t1} is an $(N_s + N_t) \times N_t$ upper triangular Toeplitz matrix with $[0, \dots, 0, h[L], \dots, h[1]]$ on the first row and $\mathbf{H}_{t0} = [\mathbf{H}_{t0,L} | \mathbf{H}_{t0,R}]$ is the concatenation of two matrices: $\mathbf{H}_{t0,R}$ is an $(N_s + N_t) \times L$ lower triangular Toeplitz matrix with $[0, \dots, 0, h[0], \dots, h[L-1]]^T$ on the first column and $\mathbf{H}_{t0,L}$ is an $(N_s + N_t) \times (N_t - L)$ Toeplitz matrix with $[0, \dots, 0, \mathbf{h}^T]^T$ on the last column and zeros on the first row. Note that \mathbf{H}_{t0} reduces to $\mathbf{H}_{t0,R}$ when $N_t = L$.

Two KSP methods are used in practice: The first proposed KSP scheme is called **Zero Padding (ZP)**. In this case, a sequence of zeros is appended at the end of each block ($\mathbf{t}_k = \mathbf{0}$, $\forall k$). The alternative to ZP is known as **Non-Zero Padding (NZP)**. NZP corresponds to the general KSP situation described in this section where known symbols are padded at the end of each block of data symbols ($\mathbf{t}_k \neq \mathbf{0}$).

An important observation is that, when the same training sequence is used for all data blocks, i.e., $\mathbf{t}_k = \mathbf{t}$, $\forall k$, the system can again be described with a circulant channel matrix:

$$\mathbf{z}_k = \left[\mathbf{H}_{KSP} \mid \mathbf{H}_{t1} + \mathbf{H}_{t0} \right] \begin{bmatrix} \mathbf{s}_k \\ \mathbf{t} \end{bmatrix} + \boldsymbol{\eta}_k = \mathbf{H}_{circ} \mathbf{x}_k + \boldsymbol{\eta}_k, \quad (2.34)$$

where the matrix \mathbf{H}_{circ} has the same structure as in (2.23), now of size N_x instead of the original $N_x - \mu$. As will be seen later in the text, the possibility of describing the system with a circulant channel matrix is highly desirable as it allows the use of low-complexity frequency-domain equalizers relying on the diagonalization properties of circulant matrices.

Chapter 3

Channel Equalization

In *Ch. 2*, we reviewed different block transmission techniques and derived the data models that describe their input-output relationship. In this chapter, we focus on the role of the decoders (see **Fig. 2.5**) that operate at the receiver side. Their role is to estimate the transmitted data from the available channel output samples. Several approaches are possible for the design of the decoder, including maximum likelihood decoding, linear equalization or decision-feedback equalization. For each of these approaches, the final receiver design also depends on the adopted transmission scheme.

We analyze here how the transmitted sequence can be estimated when the different block transmission schemes are used. We mainly focus on the specific (linear or decision-feedback) equalization schemes of the presented block transmission techniques and discuss their compared advantages and drawbacks in the context of WLAN applications. We compare the different block transmission techniques and their related equalizer structures in terms of achievable bit error rate performance as well as in terms of computational complexity or other implementational aspects.

We derive the block linear equalizers corresponding to the different block transmission schemes in *sec. 3.1*. In *sec. 3.2*, we introduce recently developed Block Decision Feedback Equalizers (BDFE) suited to KSP transmission that allow significant performance gains at the cost of an increased computational complexity. We finally analyze the expected BER performance of these different block transmission schemes and perform an experimental comparison in the light of the IEEE 802.11a and Hiperlan2 standards for WLAN in *sec. 3.3*. We discuss the computational complexity of the presented equalizers in *sec. 3.4.1* and discuss some important implementation aspects in *sec. 3.4.2*.

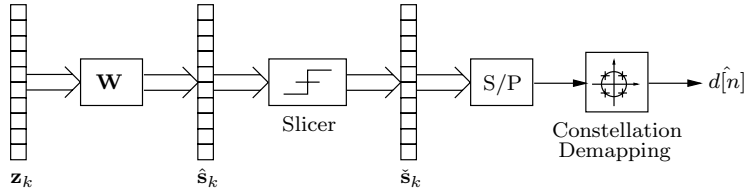


Figure 3.1: General organization of the receiver when a block linear equalizer is used to detect the transmitted data symbols.

3.1 Block Linear Equalizers

When data symbols are transmitted over a wireless channel, the received sequence is corrupted by ISI and AWGN. Linear equalizers aim at reconstructing the transmitted symbols using linear combinations of the channel output samples.

Classical transmission schemes compensate the effects of FIR channels with *serial linear equalizers (SLE)* that are usually implemented as an FIR filter. However, as the inverse of an FIR channel is a (possibly unstable) IIR filter, the SLE will at best be able to approximate the true channel inverse within the allowed computational complexity and decoding delay. Hence the performance of classical SLEs is largely sub-optimal. The design rules of such equalizers are briefly introduced in *sec. 3.1.1*.

In block transmission, specific equalizer designs that process the channel output samples on a block-per-block basis can be implemented, avoiding the problems that are encountered by classical SLEs. We focus here on *Block Linear Equalizers (BLE)*. When there is no IBI (which is the case when the CP or the padded sequences are longer than the channel order), the transmitted block of data symbols \mathbf{s}_k can be estimated relying on \mathbf{z}_k only. The BLE linearly combines the elements of \mathbf{z}_k in order to produce an estimate $\hat{\mathbf{s}}_k$ of \mathbf{s}_k . Ignoring for the time being the known symbol terms \mathbf{t}_k , the output of the equalizer \mathbf{W} can be expressed as:

$$\hat{\mathbf{s}}_k = \mathbf{W}\mathbf{z}_k = \mathbf{W}\mathbf{R}\mathbf{H}_0\mathbf{P}_s\mathbf{s}_k + \mathbf{W}\mathbf{R}\boldsymbol{\eta}_k. \quad (3.1)$$

The resulting data symbol estimates $\hat{\mathbf{s}}_k$ do not have the finite alphabet properties of the transmitted data symbols as the effect of the noise and/or ISI cannot be totally cancelled by the equalizers. Hence, they are fed to a decision device or *slicer*, which projects them onto the constellation of the data symbols and whose output is denoted $\check{\mathbf{s}}_k$. Usually, $\check{\mathbf{s}}_k$ is chosen according to the minimum distance criterion, i.e. $\check{s}_k[i]$ is chosen in the data symbol constellation such

that $|\hat{s}_k[i] - s_k[i]|^2$ is minimized. Finally, the resulting complex symbols are demapped into logical data bits. The resulting transmission scheme is depicted in **Fig. 3.1**. *Symbol errors* occur when a data symbol is wrongly decoded ($\hat{s}_k[i] \neq s_k[i]$). The goal of the equalizers is to minimize the probability of such error events. Two different criteria for the design of the equalizers are commonly used towards this goal:

The **Zero Forcing (ZF)** criterion aims at designing a linear equalizer that cancels the ISI effects of the transmission channel on the considered block of data symbols without taking the noise effects into account. In the general context of block transmission, the ZF criterion is expressed as:

$$\mathbf{W}\mathbf{R}\mathbf{H}_0\mathbf{P}_s = \mathbf{I}. \quad (3.2)$$

Inverting the channel matrix without taking the noise into account may result in noise enhancement that will cause a degradation of the system performance. The **Minimum Mean Square Error (MMSE)** criterion takes both effects (noise and ISI) into account. It aims at finding the equalizer that statistically minimizes the distance between the estimated data symbols and the actually transmitted symbols:

$$\min_{\mathbf{W}} \mathbb{E} \{ \|\hat{\mathbf{s}}_k - \mathbf{s}_k\|^2 \}. \quad (3.3)$$

or equivalently:

$$\min_{\mathbf{W}} \mathbb{E} \{ \|\mathbf{W}\mathbf{z}_k - \mathbf{s}_k\|^2 \}. \quad (3.4)$$

The solution to this problem can be expressed as

$$\mathbf{W} = \mathbb{E} \{ \mathbf{s}_k \mathbf{z}_k^H \} \mathbb{E} \{ \mathbf{z}_k \mathbf{z}_k^H \}^{-1}. \quad (3.5)$$

Depending on the transmission scheme that is used, the design rules for the equalizer will differ. *In the next paragraphs we investigate the design of equalizers for different transmission schemes under the hypothesis that the receiver has perfect channel knowledge.*

3.1.1 Serial Linear Equalizers

Serial Linear Equalizers (see e.g. [16] for more details on how these can be designed) are used in classical transmission schemes where $x[n] = s[n]$. The

linear equalizer \mathbf{w} of fixed length N_w operates directly on the sequence $y[n]$ of channel output samples. Using (2.14), the n^{th} output of the equalizer can be expressed as:

$$\hat{s}[n] = \mathbf{w}^H \mathbf{y}^n = \mathbf{w}^H \mathbf{H} \mathbf{s}^n + \mathbf{w}^H \boldsymbol{\eta}^n, \quad (3.6)$$

where $\mathbf{y}^n (N_w \times 1) = [y[n-\delta], \dots, y[n], \dots, y[n-\delta+N_w-1]]^T$, $\mathbf{s}^n (N_w+L \times 1) = [s[n-\delta-L], \dots, s[n], \dots, s[n-\delta+N_w-1]]^T$; δ being the decoding delay and $\boldsymbol{\eta}^n$ is the AWGN vector. The channel matrix \mathbf{H} of size $N_w \times (N_w+L)$ is described as:

$$\mathbf{H} = \begin{bmatrix} h[L] & \cdots & h[0] & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & h[L] & \cdots & h[0] \end{bmatrix}. \quad (3.7)$$

ZF Equalization: In this framework, the ZF criterion is expressed as:

$$\mathbf{w}^H \mathbf{H} = [0, \dots, 0, 1, 0, \dots, 0], \quad (3.8)$$

where the number of zeros preceding the 1 is equal to the decoding delay δ . Since \mathbf{H} has more columns than rows the system is over-determined and no exact solution can be found. We can only find a LS solution to this problem whose accuracy will grow with the length of the equalizer N_w :

$$\mathbf{w}^H = [0, \dots, 0, 1, 0, \dots, 0] \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1}. \quad (3.9)$$

MMSE Equalization: The MMSE criterion (3.3) turns into:

$$\min_{\mathbf{w}} E \{ \|\hat{s}[n] - s[n]\|^2 \}, \quad (3.10)$$

or equivalently:

$$\min_{\mathbf{w}} E \{ \|\mathbf{w}^H \mathbf{y}^n - s[n]\|^2 \}. \quad (3.11)$$

The solution to this problem is:

$$\mathbf{w}^H = [0, \dots, 0, 1, 0, \dots, 0] \mathbf{H}^H (\mathbf{H} \mathbf{H}^H + \sigma^2 \mathbf{I})^{-1}, \quad (3.12)$$

where $\sigma^2 = E \{ \eta[n] \eta[n]^* \}$ is the noise variance. Here again, the accuracy of the solution will increase with the complexity N_w that is allowed for the equalizer.

3.1.2 Block Linear Equalizers for CP Transmission

In the context of CP transmission, the output of the BLE can be described using (2.25):

$$\hat{\mathbf{s}}_k = \mathbf{W} \mathbf{R}' \mathbf{H}_{circ} \mathbf{P}'_s \mathbf{s}_k + \mathbf{W} \mathbf{R}' \mathbf{n}_k. \quad (3.13)$$

ZF Equalization: In this context, the ZF criterion turns into:

$$\mathbf{W} \mathbf{R}' \mathbf{H}_{circ} \mathbf{P}'_s = \mathbf{I}_{N_s}, \quad (3.14)$$

whose solution depends on the precoding and preprocessing matrices \mathbf{R}' and \mathbf{P}'_s .

- In **OFDM** systems, this criterion simplifies into (see (2.29)):

$$\mathbf{W} \mathbf{H}_{diag} = \mathbf{I}_{N_s}, \quad (3.15)$$

whose solution is:

$$\mathbf{W} = \mathbf{H}_{diag}^{-1}. \quad (3.16)$$

Note that \mathbf{W} is very simple to compute: it is a diagonal matrix with the inverse of the coefficients of \mathbf{H}_{diag} on the main diagonal. Note as well that this solution only exists when there is no zero on the DFT grid of the channel impulse response \mathbf{h} .

- In **SC-CP** systems, this criterion turns into (see (2.30)):

$$\mathbf{W} = \mathbf{H}_{circ}^{-1}. \quad (3.17)$$

The solution only exists when the circulant channel matrix \mathbf{H}_{circ} is full rank. Note that one can use the diagonalization properties of circulant matrices $\mathbf{H}_{circ} = \mathcal{I}_P \mathbf{H}_{diag} \mathcal{F}_P$ to find an equivalent implementation of

this equalizer which is computationally less demanding, leading to the following *Frequency Domain ZF Equalizer*:

$$\mathbf{W} = (\mathcal{I}_P \mathbf{H}_{diag} \mathcal{F}_P)^{-1} = \mathcal{I}_P \mathbf{H}_{diag}^{-1} \mathcal{F}_P. \quad (3.18)$$

This equalizer first performs a DFT on the received blocks, followed by a multiplication with the inverse of the diagonal channel matrix and finally performs an IDFT on the processed block. It is strictly equivalent to the previous one but, relying on the fact that the DFT and IDFT operations can be implemented respectively as FFT and IFFT operations, this FD equalizer only requires $\mathcal{O}(\log P)$ operations per estimated symbol rather than $\mathcal{O}(P)$ for the previous solution. Again, this equalizer only exists if there is no zero on the DFT grid of the channel. This condition of existence is equivalent to the full rank condition on the circulant channel matrix: when the circulant channel matrix is rank deficient, this translates into one or more zeros on the DFT grid of the channel and vice-versa.

MMSE Equalization: In CP transmission, under the hypothesis that the data and the noise are both white and uncorrelated, the MMSE criterion translates into equalizers of the form:

$$\mathbf{W} = (\mathbf{R}' \mathbf{H}_{circ} \mathbf{P}'_s)^H \left((\mathbf{R}' \mathbf{H}_{circ} \mathbf{P}'_s) (\mathbf{R}' \mathbf{H}_{circ} \mathbf{P}'_s)^H + \sigma^2 \mathbf{R}' \mathbf{R}'^H \right)^{-1}. \quad (3.19)$$

- In **OFDM** transmission, exploiting the properties of DFT and IDFT matrices ($\mathcal{F}_P^{-1} = \mathcal{F}_P^H = \mathcal{I}_P$ and $\mathcal{I}_P^{-1} = \mathcal{I}_P^H = \mathcal{F}_P$), the MMSE Equalizer simplifies into:

$$\mathbf{W} = \mathbf{H}_{diag}^H (\mathbf{H}_{diag} \mathbf{H}_{diag}^H + \sigma^2 \mathbf{I}_P)^{-1}. \quad (3.20)$$

Note that this solution always exists (even when there are zeros on the DFT grid of the channel) as soon as there is noise in the transmission scheme ($\sigma^2 \neq 0$).

- In **SC-CP** transmission, the MMSE equalizer becomes:

$$\mathbf{W} = \mathbf{H}_{circ}^H (\mathbf{H}_{circ} \mathbf{H}_{circ}^H + \sigma^2 \mathbf{I}_P)^{-1}. \quad (3.21)$$

Here again, one can exploit the diagonalization property of circulant matrices and the properties of DFT and IDFT matrices in order to simplify this expression. It is easy to show that this equalizer is strictly equivalent

to the following *Frequency Domain MMSE Equalizer*, which has again the advantage of a reduced computational complexity:

$$\mathbf{W} = \mathcal{I}_P \mathbf{H}_{diag}^H (\mathbf{H}_{diag} \mathbf{H}_{diag}^H + \sigma^2 \mathbf{I}_P)^{-1} \mathcal{F}_P. \quad (3.22)$$

It is the same as the OFDM equalizer preceded by a DFT and followed by an IDFT. Here again, the solution exists as soon as there is noise in the transmission scheme.

3.1.3 Block Linear Equalizers for KSP Transmission

When both the channel and the padded symbols are perfectly known at the receiver, the contribution of the padded known symbols can be subtracted from the received sequence, making the NZP-KSP situation equivalent to ZP-KSP:

$$\mathbf{z}'_k = \mathbf{z}_k - \mathbf{H}_{t0} \mathbf{t}_k - \mathbf{H}_{t1} \mathbf{t}_{k-1} = \mathbf{H}_{KSP} \mathbf{s}_k + \boldsymbol{\eta}_k. \quad (3.23)$$

Equalization can then be done on \mathbf{z}'_k :

$$\hat{\mathbf{s}}_k = \mathbf{W} \mathbf{z}'_k = \mathbf{W} \mathbf{H}_{KSP} \mathbf{s}_k + \mathbf{W} \boldsymbol{\eta}_k. \quad (3.24)$$

Optimal Linear ZF Equalization: The ZF criterion translates into

$$\mathbf{W} \mathbf{H}_{KSP} = \mathbf{I}_{N_s}. \quad (3.25)$$

Since \mathbf{H}_{KSP} is a tall matrix, this is an under-determined system that will have several solutions. Since \mathbf{H}_{KSP} always has full column rank by construction, its left pseudo-inverse will always exist and hence the minimum norm equalizer can be expressed as:

$$\mathbf{W} = \mathbf{H}_{KSP}^\dagger = (\mathbf{H}_{KSP}^H \mathbf{H}_{KSP})^{-1} \mathbf{H}_{KSP}^H. \quad (3.26)$$

As opposed to CP transmission, this ZF equalizer always exists, even when the channel has zeros on the DFT grid (except in the unpractical situation where all the taps of the channel are equal to zero).

Optimal Linear MMSE Equalization: Using the hypothesis that noise and data are white and zero-mean, the MMSE equalizer (3.5) can be expressed as

$$\mathbf{W} = \mathbf{H}_{KSP}^H (\mathbf{H}_{KSP} \mathbf{H}_{KSP}^H + \sigma^2 \mathbf{I}_{N_s})^{-1}. \quad (3.27)$$

Using the matrix inversion lemma¹, this equalizer can also be expressed as:

$$\mathbf{W} = (\mathbf{H}_{KSP}^H \mathbf{H}_{KSP} + \sigma^2 \mathbf{I}_{N_s})^{-1} \mathbf{H}_{KSP}^H, \quad (3.28)$$

which always exists and converges to the ZF equalizer in the noiseless case.

Frequency Domain Equalization: As we have seen previously, when the padded sequence does not change from block to block, the transmission scheme can be expressed with a circulant channel matrix (2.34). This situation is equivalent to SC-CP transmission, the padded sequence at the tail of the previous block acts as a cyclic prefix and the frequency-domain equalization schemes of CP transmission can be used, offering the advantage of a lower complexity. Since we do not need to estimate the padded symbols which are known to the receiver, the equalizers are slightly modified to focus only on the sought data symbols: only the N_s first lines of the IDFT are kept in the equalizer (the DFT size is here $P = N_s + N_t$). This yields for the **Frequency - Domain ZF equalizer**:

$$\mathbf{W} = \mathcal{I}_P(1 : N_s, :) \mathbf{H}_{diag}^{-1} \mathcal{F}_P. \quad (3.29)$$

This alternative equalization scheme differs from the optimal linear ZF equalizer (3.26). The FD ZF equalizer will only exist if there is no zero on the DFT grid of the channel, whereas the existence of the optimal linear ZF equalizer is guaranteed irrespective of the channel. The reason for this is that, the equalization problem being under-determined, there is a whole space of acceptable ZF solutions. The optimal linear ZF equalizer (3.26) is the minimal norm solution, i.e., the solution that cancels the impact of the received signal components that lie in the null space of the channel matrix (i.e., part of the noise) whereas other solutions (such as the FD ZF) tend to boost these components. In extreme cases (a zero on the DFT grid of the channel for instance), this results in unbounded noise enhancement.

The Frequency-Domain MMSE is defined equivalently as:

$$\mathbf{W} = \mathcal{I}_P(1 : N_s, :) \mathbf{H}_{diag}^H (\mathbf{H}_{diag} \mathbf{H}_{diag}^H + \sigma^2 \mathbf{I}_P)^{-1} \mathcal{F}_P. \quad (3.30)$$

¹For \mathbf{A} , \mathbf{B} , \mathbf{C} and \mathbf{D} of compatible dimensions, the matrix inversion lemma states $(\mathbf{A} - \mathbf{C}\mathbf{B}^{-1}\mathbf{D})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{C}(\mathbf{B} - \mathbf{D}\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{D}\mathbf{A}^{-1}$

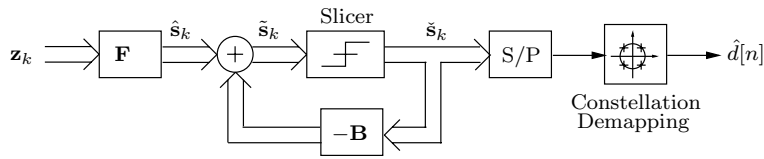


Figure 3.2: Block diagram describing the functional organization of a Block Decision Feedback equalizer. \mathbf{F} is the feedforward filter and \mathbf{B} is the feedback filter

This solution also differs from the optimal linear MMSE equalizer (3.28). However, thanks to the noise term, the existence of this sub-optimal MMSE equalizer is always guaranteed irrespective of the channel.

These FD equalizers are thus sub-optimal and are generally outperformed by the optimal linear KSP equalizers (3.26) (3.28). Their computational complexity and BER performance are exactly the same as those of their SC-CP equivalents. Their main advantage is that they are computationally much less demanding than the optimal linear KSP equalizers for several reasons. The number of operations needed per equalized symbol is $\mathcal{O}(\log(\mathcal{P}))$ compared to $\mathcal{O}(\mathcal{P})$ for the optimal linear equalizers. Moreover, when the channel matrix is made circulant by the padded sequences, equalization can be performed directly on \mathbf{z}_k and the contribution of the padded sequences \mathbf{t}_k must not be subtracted to derive \mathbf{z}'_k , which is required to perform optimal linear KSP equalization. Finally, it is less complex to derive the expression of the FD equalizer (inverse of a diagonal matrix) than the optimal linear KSP equalizer (left pseudo-inverse of a tall matrix).

3.2 Block Decision Feedback Equalizers for KSP Transmission

The block linear equalizers that were presented above directly estimate the transmitted data symbols using linear combinations of the channel output samples. The *Block Decision Feedback Equalizer (BDFE)* adopts a different approach: it introduces a non-linearity in the equalizer structure; namely, the decision device, which is only used after the equalization step in the block linear equalizer. In this section, we detail how such BDFEs are obtained in the context of KSP transmission.

The principle of the BDFE, which is depicted in **Fig. 3.2**, can be summarized as follows. The block of channel output samples \mathbf{z}_k is first fed to a feedforward filter \mathbf{F} whose role is approximately equivalent to a classical block linear

equalizer: it outputs a soft estimate $\hat{\mathbf{s}}_k$ of the transmitted block \mathbf{s}_k . An iterative procedure is then used to make the decisions on $\hat{\mathbf{s}}_k$ and to produce the final hard estimate $\check{\mathbf{s}}_k$. In this iterative procedure, which is represented by the feedback loop in **Fig. 3.2**, a decision $\check{\mathbf{s}}_k[i]$ is taken on a single data symbol $\tilde{\mathbf{s}}_k[i]$. Assuming that the decision is correct, the contribution of $\check{\mathbf{s}}_k[i]$ into $\hat{\mathbf{s}}_k$ is estimated by means of the feedback filter \mathbf{B} and is subtracted from $\hat{\mathbf{s}}_k$ to obtain $\tilde{\mathbf{s}}_k$. The procedure can be restarted and a decision is taken for another data symbol $\tilde{\mathbf{s}}_k[j]$. The procedure is repeated until all the data symbols of the block have been estimated. In order to make causal decisions, the feedback matrix is constrained to an *upper triangular* structure with zero elements on the main diagonal². Under this constraint, a decision $\check{\mathbf{s}}_k[N_s]$ can be taken directly on $\hat{\mathbf{s}}_k[N_s]$ since, by the structure of \mathbf{B} , there is no feedback on this last symbol. In the second iteration, $\check{\mathbf{s}}_k[N_s - 1]$ is estimated. The structure of \mathbf{B} requires that the contribution of $\check{\mathbf{s}}_k[N_s]$ is subtracted from $\hat{\mathbf{s}}_k[N_s - 1]$ before a decision can be taken, which can effectively be done since the decision $\check{\mathbf{s}}_k[N_s]$ is already available. With the upper triangular structure of \mathbf{B} , the iterative procedure can proceed from $\mathbf{s}_k[N_s]$ to $\mathbf{s}_k[1]$ and only past decisions are needed at each step of the procedure.

When the feedforward and feedback filters \mathbf{F} and \mathbf{B} are designed properly and if the hard estimates of $\check{\mathbf{s}}_k$ that are fed back were equal to the transmitted data symbols, the BDFE equalizer would significantly outperform the BLE. The reason for this improved performance is that the ISI is more efficiently cancelled when the interfering data symbols are known than when the ISI is directly estimated from the channel output samples (which corresponds to the BLE approach). Hence, when the past decisions are correct, the probability of making a wrong decision is significantly reduced. However, when a data symbol is erroneously detected, its contribution in the soft estimate is erroneously estimated by the feedback filter and then the probability of error of the surrounding data symbols is significantly increased, a phenomenon which is known as *error propagation*. Therefore, the symbol errors produced by a BDFE are usually not randomly distributed in the stream of data symbols, they are often grouped in a sequence of consecutive symbol errors that result from this error propagation mechanism.

We outline below how the expression of the feedforward and feedback filters of the BDFE are derived. More details on the presented derivations can be found in [17], [18] and [19] (see also [20] or [21] for a similar approach in a CDMA context). First note that the equalization is done on \mathbf{z}'_k (see eq. 3.23) rather than on \mathbf{z}_k (i.e. the contribution of the padded sequence is subtracted from the received blocks, making the NZP-KSP situation equivalent to ZP-KSP).

Assuming that past decisions are correct, i.e $\check{\mathbf{s}}_k = \mathbf{s}_k$, the input $\tilde{\mathbf{s}}_k$ of the slicer

²The feedback matrix can alternatively be constrained to a lower triangular structure, in which case the iterative procedure starts at the beginning of the block.

is expressed as

$$\tilde{\mathbf{s}}_k = \mathbf{F}\mathbf{H}_{KSP}\mathbf{s}_k - \mathbf{B}\mathbf{s}_k + \mathbf{F}\boldsymbol{\eta}_k. \quad (3.31)$$

Focusing first on the **MMSE BDFE**, we can express the MSE $\mathcal{J} = \mathbb{E}\{\|\mathbf{s}_k - \tilde{\mathbf{s}}_k\|^2\}$ at the input of the slicer as

$$\mathcal{J} = \text{tr} \left\{ \mathbf{F}(\mathbf{H}_{KSP}\mathbf{H}_{KSP}^H + \sigma^2\mathbf{I})\mathbf{F}^H - 2\Re \left\{ (\mathbf{B} + \mathbf{I})\mathbf{H}_{KSP}^H\mathbf{F}^H \right\} + (\mathbf{B} + \mathbf{I})(\mathbf{B} + \mathbf{I}_{N_s})^H \right\}. \quad (3.32)$$

Solving $\partial\mathcal{J}/\partial\mathbf{F} = 0$ w.r.t. \mathbf{F} yields the MMSE feedforward filter, which can be expressed as

$$\mathbf{F} = (\mathbf{B} + \mathbf{I})(\mathbf{H}_{KSP}^H\mathbf{H}_{KSP} + \sigma^2\mathbf{I})^{-1}\mathbf{H}_{KSP}^H. \quad (3.33)$$

Using this \mathbf{F} in (3.32) yields the following expression for \mathcal{J} :

$$\mathcal{J} = \text{tr} \left\{ (\mathbf{B} + \mathbf{I})\boldsymbol{\Omega}(\mathbf{B} + \mathbf{I})^H \right\}, \quad (3.34)$$

with $\boldsymbol{\Omega} = \sigma^2(\mathbf{H}_{KSP}^H\mathbf{H}_{KSP} + \sigma^2\mathbf{I})^{-1}$. Solving $\partial\mathcal{J}/\partial\mathbf{B} = 0$ with the new expression under the constraint that \mathbf{B} is a zero diagonal upper triangular matrix yields the expression of the MMSE feedback filter.

Noting that $\boldsymbol{\Omega}$ is positive definite by definition (and hence also its inverse $\boldsymbol{\Omega}^{-1}$), it is possible to perform the Cholesky decomposition of $\boldsymbol{\Omega}^{-1}$ and find a diagonal matrix $\mathbf{\Lambda}$ and a lower triangular matrix \mathbf{L} with ones on the main diagonal such that

$$\text{Chol} \{ \boldsymbol{\Omega}^{-1} \} = \mathbf{L}\mathbf{\Lambda}\mathbf{L}^H. \quad (3.35)$$

The feedback matrix that minimizes \mathcal{J} can then be expressed as:

$$\mathbf{B} = \mathbf{L}^H - \mathbf{I}. \quad (3.36)$$

It is interesting to note that, in the absence of feedback, the feedforward filter is strictly equivalent to the optimal linear KSP equalizer that we derived earlier. The **ZF-BDFE** is readily obtained using the above expressions with $\sigma^2 = 0$.

Finally note that the BDFE that is presented here in the KSP context can formally be adapted to CP transmission schemes. Its expression is obtained by replacing the channel matrix \mathbf{H}_{KSP} by the corresponding channel matrices,

i.e., \mathbf{H}_{circ} for SC-CP transmission and \mathbf{H}_{diag} for OFDM transmission. The resulting BDFE has some merits in SC-CP transmission but we will not detail it here. In OFDM transmission however, as there is no ISI in the transmission scheme, the feedback structure vanishes. It is straightforward to check that using the channel matrix \mathbf{H}_{diag} in the expression of the BDFE indeed yields an all-zero feedback filter and the classical BLE as feedforward filter.

3.3 Performance Analysis and Simulation Results

In this section, we compare the performance of the different equalization schemes presented above under the hypothesis that both the channel and the noise statistics are perfectly known at the receiver. Perfect timing and frequency synchronization are assumed as well. We first discuss the expected behavior of the proposed equalization schemes and then experimentally analyze their compared performance in an introductory experiment. We then analyze in more detail the performance of the proposed block equalization schemes in a realistic Hiperlan2 / IEEE 802.11a context. All the presented experimental results are performed on a set of 400 random channel realizations and approximately 200.000 data symbols are transmitted over each of these channels. As mentioned in *Chapter 2*, we consider Rayleigh fading channels with a flat power profile. We consider randomly generated additive white Gaussian noise of variance σ^2 : $E\{\eta[i]\eta[j]^*\} = \delta_{ij}\sigma^2$. Different constellations are considered for the data symbols, i.e. BPSK, QPSK, 16-QAM and 64-QAM. These constellations are always scaled such that the variance of the data symbols is 1, i.e., $E\{s_k[i]s_k[i]^*\} = 1$. The signal to noise ratio at the receiver (SNR) is defined as $SNR = E\{\|\mathbf{h}\|^2\}/\sigma^2$.

3.3.1 BER Analysis

Let us first analyze the expected performance of the different block transmission schemes before proceeding with the experimental comparison.

When OFDM transmission is used, the convolutive time-domain channel is effectively transformed in a set of parallel frequency domain flat fading channels also called tones or sub-channels. The symbol received on a given tone is the transmitted data symbol multiplied by the complex response of the channel at the frequency of the considered tone plus some noise. Given the frequency-selective nature of the transmission channel, the multiplicative coefficient changes from tone to tone and so the SNR varies from tone to tone.

When the channel is known to the transmitter, the power of the transmitted

signal and the data symbol constellation can be finely tuned for each tone using power- and bit-loading techniques (see e. g [4], [1] or [22]). When the number of tones is sufficiently large, the use of these techniques allows to approach the channel capacity. These loading techniques cannot be used when other block transmission schemes (SC-CP and KSP) are used, which can be expected to harm their performance. In [23], a theoretical analysis of the achievable bit error rates of different block transmission schemes indeed shows that OFDM systematically outperforms SC-CP transmission when the channel is known to the transmitter and these loading techniques are used. The same result holds when the FD equalizers for KSP transmission are used as they have the same performance as SC-CP equalizers. However, this study also shows that KSP transmission with block decision feedback equalizers yields the same performance as OFDM with bit loading, even though the PSD and the constellations of the transmitted signal is not specifically adapted to the channel characteristics. This remarkable result originates in the high accuracy of the BDFE, be it that the performance is achieved with an increased computational complexity. Hence, when the channel is stationary and known to the transmitter (e.g. DSL applications), OFDM appears to be more interesting than the other block transmission schemes.

In the context of wireless transmission however, the coherence time of the channel is significantly reduced and it is often problematic to have a precise and up to date knowledge of the communication channel at the transmitter. Hence, in many wireless applications where OFDM transmission is used (including IEEE 802.11a and Hiperlan2), *the transmitter does not have channel knowledge* and hence cannot use these loading techniques. The transmitted signal thus has a flat PSD and the same constellation is used for all the tones. The same problem occurs in *Digital Audio Broadcasting (DAB)* and *Digital Video Broadcasting (DVB)* applications where it is not possible to jointly adapt the transmitted signal to all the receiver channels.

In this scenario, **OFDM** transmission becomes very sensitive to channel fades in the frequency domain and the total BER performance of the system is dominated by the weakest tones. Define the noise to signal ratio $\beta(k)$ on the k^{th} tone as $\beta(k) = \sigma^2/\mathbf{H}_{diag}(k, k)$. When QPSK constellations are used, the probability of error of the OFDM transmission scheme can be expressed as (see e.g. [24] where a BER analysis of different block transmission schemes is proposed) as³:

$$\mathcal{P}_{OFDM} = \frac{1}{N_s} \sum_{i=1}^{N_s} Q\left(\frac{1}{\sqrt{\beta(i)}}\right), \quad (3.37)$$

³Note that this expression of the error probability holds for both ZF and MMSE equalizers. Indeed, the output of the MMSE equalizer is simply a scaled version of the ZF equalizer's output. In BPSK and QPSK transmission, this has no impact on the final decision. When higher order constellations are used, correcting the bias introduced by the MMSE equalizer yields the same result as the ZF equalizer.

where $Q(x)$ is the complementary distribution function of a standard Gaussian distribution (see e.g. [16]):

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\alpha^2/2} d\alpha. \quad (3.38)$$

This error probability is simply the sum (divided by N_s) of the different sub-channels' error probabilities. When there are one or more zero or near-zero channel coefficients, the probability of error for the data symbols transmitted on these tones approaches 0.5. The total performance of the OFDM system is then dominated by these weak tones that will yield detection errors even for relatively large values of the global SNR. A commonly used solution to avoid this problem consists in encoding the data across the tones. In the Hiperlan2 and IEEE 802.11a standards for instance, the binary data are encoded by means of a convolutional code whose output is interleaved across the tones. Doing so, a single bit of information is spread over several tones and only the part of this information that is transmitted over the weakest tones has a significant chance of being lost in the transmission process. Hence, the chance of losing all the information related to one bit of information is significantly reduced and many symbol errors can be corrected at the receiver. An alternative way of protecting oneself against channel fades in the frequency domain consists in spreading the data symbols $\mathbf{s}_k[i]$ across the tones by means of a linear precoder, i.e. the block of data symbols \mathbf{s}_k is multiplied by a coding matrix before transmission (see e.g. [9], [25], [26] or [27] where the design of such precoding matrices is discussed).

The **SC-CP** transmission scheme can be viewed as such a linearly precoded OFDM scheme where the precoding matrix is the DFT matrix of size N_s . Indeed, if the data symbols are encoded with a matrix \mathcal{F} in a linearly precoded OFDM scheme, the IDFT that is performed on the encoded symbols cancels the effect of \mathcal{F} and the transmitter becomes equivalent to a SC-CP transmitter. Hence, one can expect a reduced performance sensitivity of the SC-CP scheme to the frequency-selectivity of the transmission channel. The analysis proposed in [24] shows that, when *ZF equalizers* are used with QPSK data symbols, the BER performance of SC-CP systems can be computed as:

$$\mathcal{P}_{SC-CP,ZF} = Q \left(1 / \sqrt{\frac{1}{N_s} \sum_{i=1}^{N_s} \beta(i)} \right). \quad (3.39)$$

As opposed to OFDM transmissions, all the data symbols of $\mathbf{s}_k[i]$ have the same error probability. This results from the fact that the SC-CP equalizer is a circulant matrix. Hence, all the estimated data symbols are obtained with the same linear equalizer whose inputs are circularly shifted and they all have

the same SNR at the input of the slicer. Note that the effective SNR is not the mean of all the SNRs on each tone, it is the inverse of the mean of the noise to signal ratio over the tones. If there is a zero on the DFT grid of the channel, the noise to signal ratio on this tone is infinite and the global SNR at the input of the slicer is zero, making the probability of error equal to 0.5. The sensitivity to frequency-domain channel fades thus remains important when ZF equalizers are used. This sensitivity is significantly reduced when MMSE equalizers are used as the noise enhancement is then limited and the unbounded noise enhancement that happens on the null tones in the ZF scheme is not possible anymore.

In [24], considering that the residual ISI of the *MMSE equalizer* has a Gaussian distribution and defining the averaged noise to signal ratio $\bar{\beta}$ as:

$$\bar{\beta} = \frac{1}{N_s} \sum_{k=1}^{N_s} \frac{\beta(k)}{1 + \beta(k)}, \quad (3.40)$$

the authors show that, when QPSK constellations are used, the probability of error of the MMSE SC-CP scheme can be expressed as:

$$\mathcal{P}_{SC-CP,MMSE} = Q\left(\sqrt{\bar{\beta}^{-1} - 1}\right). \quad (3.41)$$

Furthermore, the authors of [24] prove that this probability of error is always smaller than that of the OFDM scheme:

$$\mathcal{P}_{SC-CP,MMSE} \leq \mathcal{P}_{OFDM}, \quad (3.42)$$

with equality only when the channel is flat-fading and the SNR is the same on all the tones.

In **KSP** transmission, the error probability is the same as $\mathcal{P}_{SC-CP,MMSE}$ when the FD equalizers are used. When the *optimal linear KSP equalizers* are used, the probability of error cannot be expressed as nicely as in the CP case and it is hard to analytically derive a direct relationship between the KSP and CP error probabilities (see e.g. [28] for a full expression of the error probability of the optimal linear ZF KSP equalizer). However, as the optimal linear ZF KSP equalizer is the minimum norm equalizer, the noise enhancement is less important than when the FD equalizer is used. Moreover, as the channel matrix has full column rank, the optimal linear ZF KSP equalizer can accommodate an arbitrary number of channel zeros in the frequency domain (channel-irrespective symbol recovery). It is thus clear that the optimal linear ZF KSP equalizer outperforms its FD counterpart and hence the ZF SC-CP equalizer. Similarly, the optimal linear MMSE KSP equalizer truly minimizes the mean squared

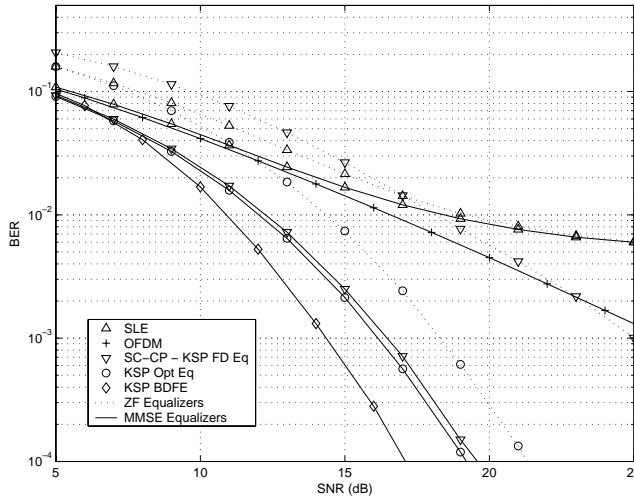


Figure 3.3: BER vs. SNR for different MMSE and ZF equalization schemes with different transmission schemes (serial, OFDM, SC-CP and KSP)

error at the input of the slicer, i.e. the mean squared error of the MMSE FD equalizer is larger. Hence, the optimal linear MMSE KSP equalizer has a lower probability of error than its FD counterpart or the MMSE SC-CP equalizer.

It thus appears that KSP outperforms other transmission schemes when the optimal linear equalizers are used. Moreover, besides the interesting channel-irrespective symbol recovery property (see [2] where this was first introduced), the use of the optimal linear KSP equalizers allows to fully exploit the *delay diversity* offered by the multipath channel, which is equal to $L + 1$ (see [9] where this was first introduced and [25] where it was further developed). The diversity exhibited by a transmission scheme is the slope of its averaged (with respect to the random channel) BER as a function of the SNR in the high SNR region, i.e. it is a measure of how fast the number of symbol errors drops as the SNR increases. In contrast to KSP, CP systems do not fully exploit the diversity of the channel and OFDM only exhibits diversity one. Hence, the advantage of KSP will grow as the SNR increases.

Finally, the performance of the *block decision feedback equalizer* for KSP is very sensitive to error propagations. Its theoretical performance can be derived only assuming that past decisions are correct. Under this hypothesis, a theoretical value of the BDFE error probability can be derived (see e.g. [18] for the full expression of this error probability). In [28], the authors prove that this theoretical error probability is lower than that of the optimal linear KSP equalizers. However, the true system performance can differ substantially from

these theoretical values as the error propagation mechanisms can significantly harm the overall system performance. Hence, the performance of the BDFE will be evaluated experimentally. Note that an important advantage of the BDFE over the classical serial DFE is that the error propagation mechanism is stopped at the end of each block which limits its impact on the overall system performance.

3.3.2 Introductory Experiment

To obtain a first insight on how the proposed equalization schemes compare, we analyze their BER performance when QPSK data symbols are transmitted over a Rayleigh-Fading channel of order $L = 5$. We compare the Bit Error Rate (BER) performance of a traditional SLE of length 15 with OFDM, SC-CP and KSP transmission.

For KSP, we consider FD Equalizers as well as the optimal linear equalizers and the BDFE. Note that, in this context of known channels, FD KSP equalizers have the same performance as the SC-CP equalizers. In the simulation results presented here and in the next section, amongst these two systems, we only mention the performance of FD KSP equalizers, but the conclusions that are drawn naturally hold for SC-CP transmission. We consider both ZF and MMSE equalizers. We chose a typical length of $N_s = 64$ for the blocks of data symbols and we use a guard band of length $N_t = \mu = 5$ to avoid IBI.

The results are presented in **Fig. 3.3**. As expected, we see that the classical SLEs are outperformed by the block equalizers, especially in the high SNR region. These results show as well that OFDM transmission is outperformed by most other block transmission schemes. When the noise statistics are unknown and ZF equalizers have to be used, ZF FD KSP equalizers only outperform OFDM for high SNR values because the noise enhancement that equally affects all the equalized symbols is more important in the low SNR region. As expected, we see that optimal linear ZF KSP equalizers significantly outperform their FD counterpart.

When the noise statistics are known and MMSE equalizers can be used, the performance is improved for both optimal and FD linear KSP equalizers. The performance improvement is more spectacular for the low-complexity FD equalization scheme whose performance becomes comparable to the more complex optimal linear equalization scheme. Finally, the BDFE KSP clearly outperforms all others, allowing to gain several dBs at the cost of an increased complexity.

3.3.3 Equalizers Performance in the Context of the Hiperlan2 and IEEE 802.11a Standards

In order to perform a more complete comparison between the different proposed block transmission techniques and their related block equalizers, we analyze their performance in a context similar to the Physical Layer that is shared by two well known standards for WLAN: Hiperlan2 and IEEE 802.11a. Both of these standards rely on OFDM. We analyze here how the overall system performance is changed when other block transmission techniques are used whilst the other parameters proposed in the standard (constellation mapping, error coding, data interleaving, block length, guard band duration, etc...) are not changed.

Note that in order to maintain the complexity of the performed simulations at a reasonable level, we had to perform the error correction step at the receiver with a hard Viterbi decoder rather than the soft Viterbi decoder that is generally used in real implementations. This choice results in a 3dB performance degradation for all the coded schemes that are considered in the rest of the text.

Standard description: Both standards consider an OFDM transmission scheme with 64 subcarriers and a guard band duration of 16 samples (or 8 in an optional mode). In the adopted notation, this translates into $N_s = 64$, $\mu = N_t = 16$ (or 8) and $N_x = 80$ (or 72). Note that this results in different FFT sizes for CP and KSP schemes: $P = 64$ for OFDM and SC-CP and $P = 80$ (or 72) in KSP⁴. The data are transmitted in long packets (typically 50 to 500 OFDM symbols depending on the constellation and coding rate). Before transmission, the packet of binary data is encoded with a binary convolutional code. The coded data are organized in blocks, each block corresponding to a single OFDM symbol. A constellation-dependent interleaver operating on a block-per-block basis is used in order to improve the efficiency of the convolutional code. Interleaved binary data are then mapped onto data symbols before OFDM transmission (IDFT + CP insertion). A preamble (i.e. a long sequence of known symbols) used by the receiver for channel estimation, carrier and timing synchronization is appended at the beginning of each transmitted packet. At the receiver, the data symbols are estimated with a classical OFDM equalizer, a demapper translates them into binary data that are then de-interleaved before going through a Viterbi decoder that estimates the initial flow of uncoded data. The data rate of the system can be tuned (in order to optimally exploit the quality of the radio link) by the choice of the data symbol con-

⁴If one wants to have the same FFT size in KSPas in OFDM, the number of data symbols per block must be reduced s.t. $N_s + N_t$ is equal to the desired P . In the context of the considered standard, this would result in $N_s = 48$ and $N_t = 16$ (or $N_s = 56$ and $N_t = 8$ in the optional mode). Such a choice reduces the effective throughput of the system by 6.25% (1.56% in the optional mode).

Data Rate (Mbits/s)	Modulation	Coding Rate
6	BPSK	1/2
9	BPSK	3/4
12	QPSK	1/2
18	QPSK	3/4
24	16 QAM	1/2
36	16 QAM	3/4
48	64 QAM	2/3
54	64 QAM	3/4

Table 3.1: Data rate obtained with different combinations of the constellations and the coding rates (IEEE 802.11a)

stellations and the rate of the FEC-encoder. The available constellations for the mapping of binary data into data symbols are Gray-coded BPSK, QPSK, 16-QAM and 64-QAM (respectively 1, 2, 4 and 6 bits per data symbol). Depending on the chosen constellation, the standard allows to pick rate 1/2, 3/4 or 2/3 codes (rate $r = 2/3$ and $r = 3/4$ codes result from the puncturing of a unique rate $r = 1/2$ mother code). The allowed combinations of constellations and coding rates are detailed in **Table 3.1.** together with their resulting bit rates.

In **Fig. 3.4.**, we compare the *uncoded BER performance* (i.e., the system performance without the convolutional encoder and Viterbi decoder) of the different block transmission schemes (OFDM, FD, optimal linear and BDFE KSP equalizers) when the different symbol constellations of the standard are used. The results are averaged over 400 random realizations of a fifth-order ($L = 5$) Rayleigh-fading channel using the optional short guard band of the standard (i.e., $\mu = 8$ for OFDM and SC-CP, $N_t = 8$ for KSP). A total of approximately 750.000 data bits is transmitted for each of these channel realizations. The presented results are obtained with MMSE equalizers assuming perfect synchronization and channel knowledge. As expected, OFDM transmission is largely outperformed by the different KSP schemes. Furthermore, this experiment confirms that OFDM only achieves a diversity advantage of 1 whereas KSP exhibits a higher diversity. In fact, the BER curves of the optimal linear KSP equalizers in the low BER region corresponds to a diversity advantage of $L + 1$ for all constellations, which confirms that KSP fully exploits the delay diversity of the channel as foreseen by the theory. In this experiment where the diversity advantage is equal to 6, the BER is divided by $10^{0.6} = 3.98$ every time the SNR is increased by 1dB, whereas it is only divided by $10^{0.1} = 1.26$ when OFDM is used. Note as well that FD and optimal linear KSP equalizers exhibit similar performance for low-order constellations (BPSK and QPSK) and hence, the MMSE FD equalizers appear to exploit the delay diversity for these constellations. However, they are outperformed by the optimal linear equalizers

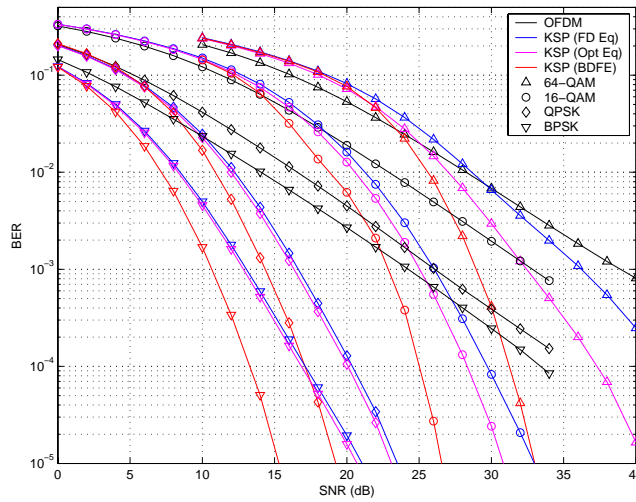
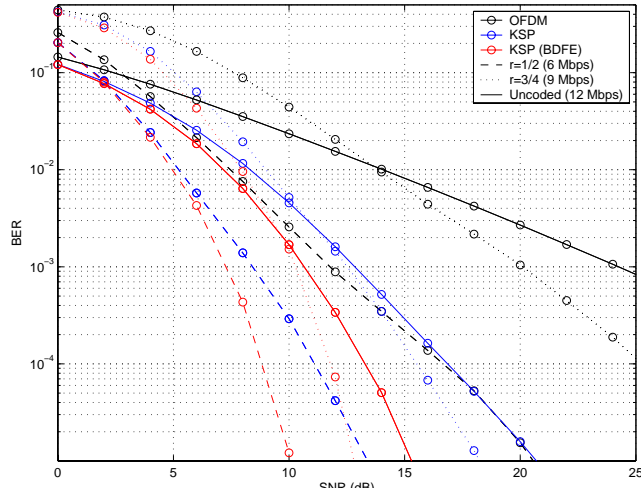


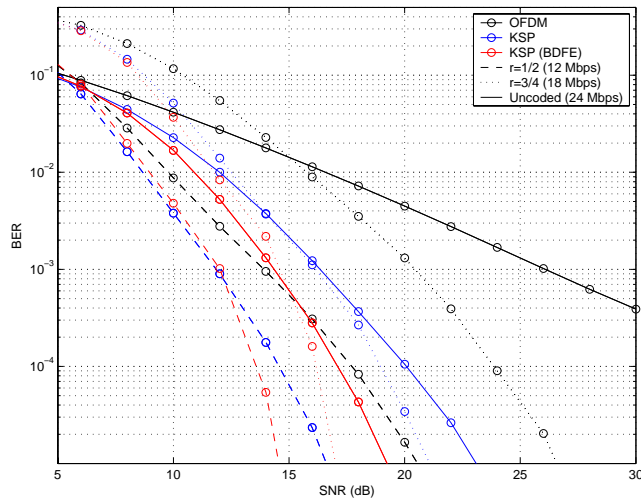
Figure 3.4: Uncoded BER performance vs. SNR for OFDM transmission and KSP transmission with resp. FD, optimal linear equalization and BDFE when the MMSE Block Equalization Schemes are used with different constellations. The equalizers are computed with the exact channel model.

when higher order constellations (16-QAM and 64-QAM) are used and thus do not fully exploit the diversity for these constellations. Finally, the non-linear BDFE clearly outperforms all the block linear equalizers and appears to be the most efficient of the equalization schemes we are reviewing.

The convolutional code of the standards protects the OFDM transmission against the frequency domain channel fades that harmed its performance in the previous experiment with uncoded transmission. Indeed, when combined with the interleaver, the encoder effectively spreads each bit of information over several non-adjacent carriers and thereby reduces the sensitivity to the frequency-selectivity of the channel. Besides their intrinsic diversity and coding gains (the coding gain corresponds to a horizontal shift of the BER curves in the presented figures and depends on both the depth and the rate of the encoder, whereas the diversity gain describes the asymptotic slope of the curve and depends on the coding rate only), one can expect the convolutional code to offer an additional diversity advantage resulting from the exploitation of the delay diversity of the channel. Indeed, if the interleaver is well designed and the encoding depth is sufficient, the data symbols onto which one bit of information is spread will be transmitted on tones that are sufficiently distant to have a limited correlation, thereby allowing to receive up to $L+1$ independent replicas of the same bit of information and effectively exploiting the channel's delay diversity. For KSP transmission, as the channel diversity is intrinsically

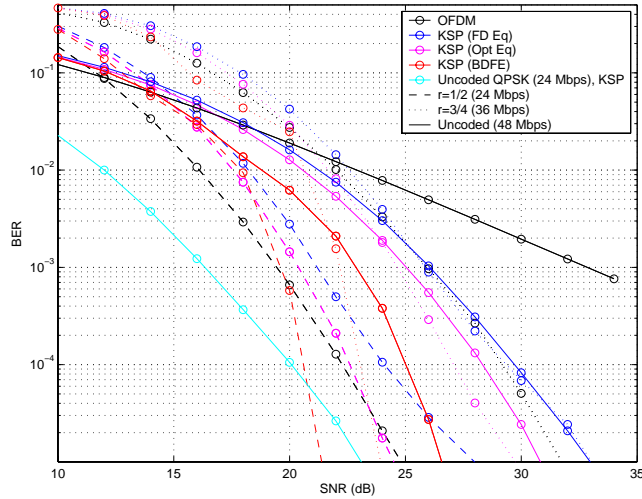


(a)

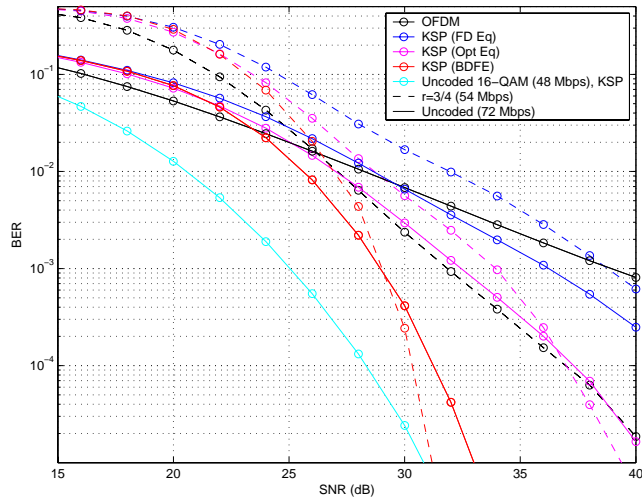


(b)

Figure 3.5: BER performance vs. SNR for different Coding rates (1/2, 3/4, Uncoded) in BPSK (a) and QPSK (b) modulation with exact channel knowledge when different block transmission techniques (OFDM and KSP) are used with their respective equalization schemes.



(a)



(b)

Figure 3.6: BER performance vs. SNR for different Coding rates (1/2, 3/4, Uncoded) in 16-QAM (a) and 64-QAM (b) modulation with exact channel knowledge when different block transmission techniques (OFDM and KSP) are used with their respective equalization schemes.

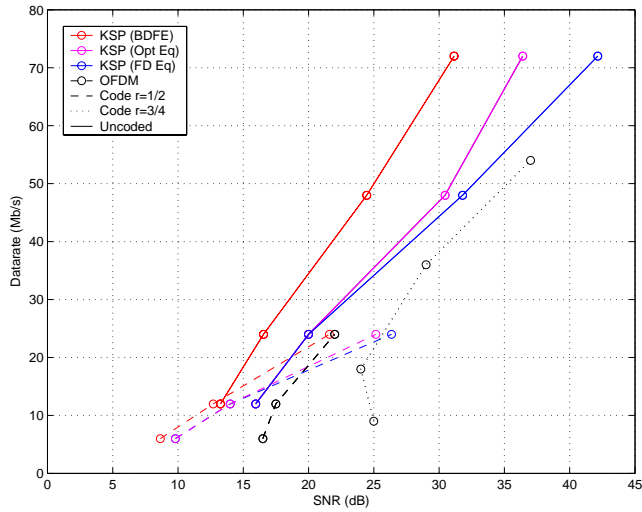
exploited by the transmission scheme, the advantage yielded by the use of the codes does not exceed their intrinsic coding and diversity gains. In **Fig. 3.5** and **Fig. 3.6**, we investigate the impact of the *different coding schemes* of the standards on the system performance when different block transmission techniques are used. Each figure presents the performance for one constellation under the different possible coding schemes. Uncoded transmission resulting in higher bit rates (not foreseen in the standards) is added for comparison on the figures. Since FD and optimal linear KSP equalizers have similar performance for small constellations, we group both transmission schemes under the label KSP when BPSK and QPSK constellations are used (the performance of the optimal linear KSP equalizer is actually shown). The difference is made for 16- and 64-QAM constellations as optimal linear equalization outperforms its FD counterpart. Rate 2/3 codes foreseen in 64-QAM are not investigated.

The results for BPSK and QPSK constellations are presented in **Fig. 3.5** (resp. **a** and **b**). Uncoded OFDM is the worst performer as expected but its performance is significantly improved when $r = 3/4$ codes are used. The use of the $r = 1/2$ codes in OFDM allows to equal or slightly outperform linearly equalized uncoded KSP scheme. KSP transmission appears to perform much better than OFDM in these simulations and the BDFE equalizer outperforms its linear counterpart. Note that the performance advantage yielded by the use of the coded schemes is much less important in KSP than in OFDM.

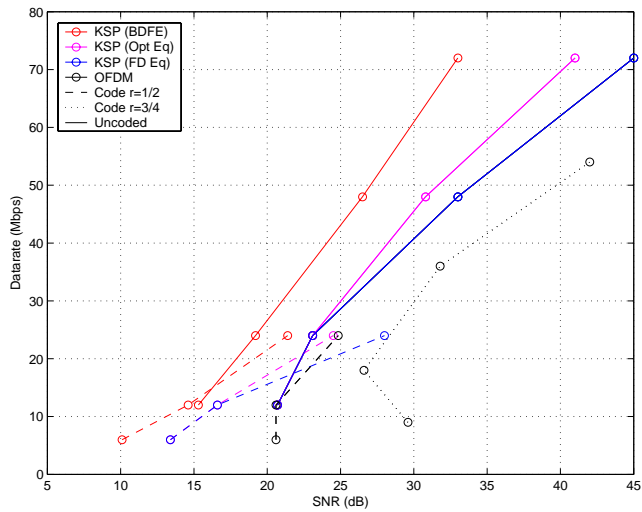
The results presented in **Fig. 3.6** show that the performance gap between OFDM and KSP is significantly reduced or even closed when 16-QAM and 64-QAM constellations are used. Uncoded OFDM is still outperformed by uncoded KSP, but the difference is relatively small when FD KSP equalizers are considered (especially in 64-QAM). The performance advantage of uncoded KSP is restored when the optimal linear equalizer is considered and further increased with the BDFE. When the coded schemes are considered, OFDM outperforms FD KSP and approaches the performance of the optimal linear KSP equalizer very closely. Only the BDFE still maintains a significant advantage over OFDM.

These simulation results confirm that the performance improvement brought by the use of the codes is considerably larger for OFDM transmission than for KSP transmission. Coded OFDM always performs better than uncoded OFDM and has a significantly increased diversity. In contrast, when the linear equalizers are used, the diversity advantage brought by the use of the codes is marginal in KSP, whereas the coding gain is significant for $r = 1/2$ codes only and barely noticeable for $r = 3/4$ codes (or even negative in 64 QAM with FD equalization).

The reason for this poor performance of the coded KSP scheme probably finds its origin in the specificity of the codes adopted in the standards. These have indeed been designed for the OFDM scheme that is characterized by a non-



(a)



(b)

Figure 3.7: Achievable Data rates (Mbits/s) in function of the SNR for target BER of 10^{-4} (a) and 10^{-5} (b) for different transmission schemes and their related equalizers. Each curve is obtained with a fixed equalization scheme and a fixed coding rate but with different constellations. The equalizers are computed with the exact channel model.

uniform distribution of the error probability across the tones. In contrast, this distribution is uniform in KSP and the coded performance would be significantly improved if a specific coding scheme was used.

It is worth noting that uncoded KSP offers interesting performance and always compares with its $r = 3/4$ coded counterpart. The absence of coding not only saves the computationally demanding steps of encoding and decoding, it also avoids the redundant symbols introduced by the codes and hence increases the throughput. Furthermore, for a given constellation, the uncoded schemes offer similar bit rates than the coded schemes ($r = 1/2$) of the higher order constellation. In order to compare these uncoded schemes with the coded schemes offering similar bit rates, they are introduced in the figure of the higher order constellation (e.g., uncoded KSP transmission with QPSK constellation is introduced in **Fig. 3.6(a)**, where 16-QAM schemes are shown). This addition to the plots shows that the uncoded schemes outperform the $r = 1/2$ coded schemes offering similar bit rates, especially for high order constellations. Finally note that KSP BDFE significantly outperforms other schemes for all constellations and all coding schemes. Note that the interleaver that we used for the BDFE differs from the standard interleaver. The data are not only interleaved inside a single block, they are also interleaved from block to block, thereby reducing the impact of the error propagation mechanism, which explains the increased advantage of the coded BDFE schemes (note that, in these simulations, the transmission channel remains constant during the transmission of the full packet of data, hence, this interleaving scheme does not bring any diversity advantage).

In order to further assess the different proposed transmission schemes, we compare the achievable transmission speed for a target BER as a function of the SNR in **Fig. 3.7**. These plots are directly derived from the results of the experiments presented above. For the sake of clarity, only the best performing schemes are included in the figures. Uncoded OFDM is not shown, neither is $r = 3/4$ coding for SC-CP and KSP transmission. The results are shown in **Fig. 3.7(a)** for a target BER of 10^{-4} and in **Fig. 3.7(b)** for a target BER of 10^{-5} . Uncoded KSP dominates the other techniques in terms of achievable data rate for a given SNR in the high SNR region, rate 1/2 coded KSP dominates in the low SNR region. The FD and optimal linear KSP equalizers offer similar performance for most SNRs and the optimal linear KSP equalizers only for the highest SNRs, but always outperforms OFDM. The BDFE significantly outperforms all other schemes in all circumstances.

Mflops/sec	OFDM	SC-CP or FD - KSP	Optimal KSP	BDFE KSP
Rx	200	380	2540	3800
Tx	180	0	0	0
Total	380	380	2540	3800

Table 3.2: Processing power requirements for the different block transmission schemes and their related equalizers in an Hiperlan2 / IEEE 802.11a setup

3.4 Complexity Analysis and Implementation Issues

Besides the achievable BER performance that has been discussed in previous section, a number of other considerations play an important role when the implementability of a given transmission method needs to be assessed. It is beyond the scope of this research to assess in details all the aspects that come into consideration before a given technique can be fully implemented. In this section, we aim at gaining a first insight into the implementability of the block transmission techniques presented above. Towards that goal, we review some important issues that arise in the considered block transmission techniques.

3.4.1 Computational Complexity

An important aspect for the implementability of the presented block transmission schemes is their computational complexity. We present here an overview of the computational complexity that is required for the implementation of the equalization schemes discussed previously.

One of the main advantages of the first proposed block transmission scheme, OFDM, is its low computational complexity. The main reason for this low computational complexity is the possibility of implementing the DFT and IDFT operations respectively as FFT and IFFT operations. Indeed, if P is the size of the DFT (or IDFT), a total of $2P^2 - P$ operations (P^2 additions and $P^2 - P$ multiplications) are needed to perform the DFT, but the computational complexity is reduced to $\mathcal{O}(P \log_2[P])$ when the FFT is used. The exact number of operations required to perform the FFT depends on the chosen implementation. In the rest of the text, we consider the *decimation-in-time radix-2 FFT* [29] that requires a total of $1.5P \log_2[P]$ Flops (floating point operations) to compute a P -point FFT. Taking into account the one tap frequency-domain equalization step, the total complexity of an OFDM receiver is $1.5 \log_2[P] + 1$ Flops per data symbol ($1.5 \log_2[P]$ Flops for an OFDM transmitter).

When the FD equalizers are used, the total computational complexity (transmit

+ receive) is the same as in OFDM for both SC-CP and KSP schemes and so the average processing power required for their implementation is the same as in the OFDM case. However, as all the processing is done by the receiver, the peak processing load that such a terminal has to handle has to be evaluated when it operates in receive mode and is equal to $3\log_2[P] + 1$ Flops per symbol.

In KSP transmission, when the optimal linear equalizers are used, the complexity of the receiver is significantly increased as a full matrix multiplication is required for each estimated block of data. In this case, the receiver complexity is $2P - 1$ Flops per symbol (P multiplications and $P - 1$ additions).

When the BDFE equalizer is used, the feedforward filter \mathbf{F} yields the same complexity as the optimal linear KSP equalizer. Given the structure of the feedback matrix \mathbf{B} , the matrix multiplication involves $\sum_{i=1}^P i - 1 = (P^2 - P)/2$ scalar multiplications and $(P^2 - 3P)/2$ additions. Subtracting the output of \mathbf{B} from $\hat{\mathbf{s}}_k$ requires another P additions. Hence, the total computational complexity of the BDFE is $3P - 2$ Flops per symbol.

In order to give an indication of how these requirements translate in terms of processing speed requirements, we give in **Table 3.2** an overview of the number of Flops that are required by the different schemes in a Hiperlan2 / IEEE 802.11a environment where the block size is set to $P = 64$ and the symbol rate is $20 \cdot 10^6$ symbols/second.

The outcome of this analysis is that, when the FD equalizers are used, all the proposed transmission schemes have a comparable complexity, with a small advantage for OFDM, which spreads more evenly the processing power between the transmitter and the receiver. When the optimal linear KSP equalizers or the BDFE equalizers are used, the computational complexity is increased approximately by a factor 10. Even though the computational complexity of these higher performance equalization schemes is significantly higher than for FD equalization schemes, the difference is not dramatic. Current high-end DSP chipsets are able to compute up to $2.5 \cdot 10^9$ multiplications per second, which approximately corresponds to the requirements for the optimal linear KSP equalizers in the considered setup. The foreseeable progression of the DSP chipsets should thus enable the use of the higher complexity equalizers in a near future.

Besides the channel equalization task, when the data are encoded by means of error-correcting codes (as in Hiperlan2), the receiver has to run a decoding algorithm, which is often computationally demanding. As discussed earlier, such codes are not mandatory to reach acceptable BER performance in KSP transmission, whereas their use is mandatory to achieve acceptable performance in OFDM. Numerous algorithms of various performance and complexity have been developed for the decoding of such codes and hence we have not included the decoding task in our complexity discussion. However, the possibility of

avoiding this decoding task in KSP transmission reduces the complexity gap between the BDFE or optimal linear KSP equalizers and the FD equalizers.

Note that the presented analysis does not take into account the computation of the equalizer coefficients, the estimation of the channel and other tasks such as timing and frequency offset estimation and compensation. These also have an important impact on the total required processing power and need to be analyzed as well in order to get a complete view of the computational complexity.

Note as well that the higher requirements of the KSP scheme in terms of processing power also has an impact on the power consumption of the wireless terminal. The high-end DSPs that are currently able to cope with the complexity of the equalization are quite greedy in terms of power consumption. Hence, the possible power savings resulting from the lower SNR requirements of KSP transmission scheme will be at least partly counter-balanced by these computational aspects. However, it is foreseeable that this problem will progressively disappear with the natural evolution of the DSP chipsets technology.

3.4.2 Implementational Issues

Besides the computational complexity where OFDM has a clear advantage over other block transmission techniques, a number of issues are known to cause problems when implementing MC-CP transmission techniques. We review here the most important of these issues and see how they impact SC-CP and KSP transmission.

a) Peak To Average Power Ratio

The time-domain OFDM signal consists of the superposition of multiple complex exponentials that are randomly modulated in phase and amplitude by the data symbols. When these different sine waves interfere in a constructive way, large peaks appear in the time-domain signal, whereas the signal can be near-zero when the interference is destructive. Hence, the resulting time-domain signal has a large dynamic range, or *Peak to Average Power Ratio (PAPR)*, which is an important drawback of OFDM systems for several reasons. Classical transmission schemes transmit finite alphabet signals that have a limited dynamic range. Hence, low-quality power amplifiers can be used for the transmitted signal as a small linear region is sufficient to transmit such signals. An OFDM signal transmitted with such amplifiers will undergo non-linear distortions as it uses the amplifier significantly beyond its linearity zone. These distortions destroy the orthogonality between the carriers, which dramatically harms the system performance as the resulting ICI is generally not compensated

Constellation	PAPR in OFDM	PAPR in KSP	Ratio
BPSK	18.06 dB	0.97 dB	17.09 dB
QPSK	18.06 dB	0.97 dB	17.09 dB
16 QAM	20.61 dB	3.52 dB	17.09 dB
64 QAM	21.74 dB	4.65 dB	17.09 dB
256 QAM	22.29 dB	5.20 dB	17.09 dB

Table 3.3: Peak-to-average power ratio of the time-domain signal resulting from the use of Nyquist pulses for different block transmission schemes for a system with 64 sub-carriers [28]

at the receiver [30]. Moreover, these distortions are the source of out-of-band radiations that can affect neighboring systems. The large PAPR also requires the use of ADC and DAC converters that have a large dynamic range. With this increased dynamic range, one has to choose between an increased quantization noise or a larger number of bits for a finer digital representation of the analog signals. The first solution reduces the effective SNR whilst the second results in an increased complexity of the DSP stage.

In contrast, in SC-CP and KSP, the transmitted signals have finite alphabet properties. Even though the transmitted signals are not constant modulus when higher-order constellations are used, the dynamic range of the transmitted signal is significantly smaller than in OFDM. A detailed study in a Hiperlan2 context, which was presented in [28] and whose results are detailed in **Table 3.3**, shows that the PAPR difference between SC-CP or KSP and OFDM is equal to 17.09 dB independently of the considered constellation.

Several methods have been proposed recently to address this PAPR problem (see e. g [31], [32], [33], [34], [35] or [36]). They either increase the overhead, or the computational complexity, or introduce noise in the transmission scheme and often fail to meet the power masks requirements. Moreover, the resulting PAPR remains several dBs larger than in SC-CP or KSP systems. These techniques are generally not foreseen in current OFDM standards.

A family of techniques that are commonly used to avoid the non-linear zone of the power amplifiers is termed *clipping* (see e.g. [37], [33], [38], [39] or [40] and references therein). These techniques avoid the saturation of the power amplifiers by limiting the amplitude of the transmitted signals in the digital domain and often consist in a non-linear saturation stage where the value of the largest samples are limited to an acceptable threshold.

These clipping techniques allow to reduce the amount of out-of band radiations but do not impact the ICI problem and its associated performance degradation [41]. In order to avoid this clipping as much as possible or to at least reduce the resulting ICI to an acceptable level, the transmitted signals are of-

ten downscaled before transmission [42], which reduces the operating SNR of the system.

b) Carrier Frequency Offset

In order to modulate or demodulate the RF signal on the carrier frequency f_c , the terminals rely on a local oscillator whose operating frequency is close to the desired frequency f_c with some error margin. Hence, when a signal is transmitted it is actually modulated on a carrier frequency $f_{c,t}$ and demodulated with an oscillator of frequency $f_{c,r}$. Both are close to the desired f_c but actually differ by the so-called *Carrier Frequency Offset (CFO)*, $f_{off} = |f_{c,t} - f_{c,r}|$.

This frequency offset (see e.g. [28] where its effects are discussed) acts as a multiplicative noise that reduces the useful signal amplitude in single-carrier systems such as KSP. In OFDM systems, the CFO destroys the orthogonality between the subcarriers, which again harms the performance as the resulting ICI is not compensated. Note that the CFO affects all the FD equalization schemes presented here (i.e. not only OFDM but also SC-CP and FD-KSP equalizers).

When the CFO is small compared to the carrier spacing, it has been proved in [43] that the effects of the CFO can be modeled as an SNR loss γ_{snr} which is approximated for the different systems as

$$\gamma_{snr,ofdm} \approx \frac{10}{3\ln 10} (\pi f_0 N_s)^2 \frac{\lambda^2}{\sigma^2} \text{dB}, \quad (3.43)$$

$$\gamma_{snr,ksp} \approx \frac{10}{3\ln 10} (\pi f_0)^2 \text{dB}, \quad (3.44)$$

where f_0 is the CFO normalized with respect to the sampling frequency, i.e. $f_0 = f_{off} T_s$, and λ^2/σ^2 is the SNR of the subchannel. The SNR degradation is proportional to the squared normalized CFO in both cases, but is $N_s^2 \lambda^2/\sigma^2$ times larger in OFDM. Hence, the performance degradation caused by the CFO is significantly more important in OFDM, SC-CP or FD KSP than when the BDFE or optimal linear KSP equalizers are used.

In order to avoid this performance degradation, several techniques have been developed to compensate for the CFO in OFDM systems, see e.g. [44] and references therein. An accurate estimate of the CFO is needed in order to apply this compensation. Several techniques have been developed to estimate the CFO based on the CP-induced cyclostationarity of OFDM signals (see e.g. [45], [46], [47] or [48]). Pilot tones are inserted in many OFDM standards, including Hiperlan2 in order to allow more simple CFO estimation schemes, and several approaches are possible to use these pilots (see e.g. [49] or [50]

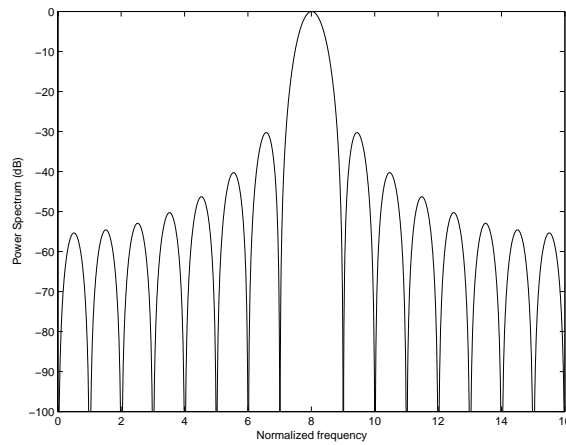


Figure 3.8: Power Spectral Density of the DFT/IDFT filter in a system with 16 subcarriers

and references therein). On top of these pilots, a part of the preamble that is appended at the beginning of each packet is specifically foreseen for CFO estimation, see e.g. [51] or [52] and references therein.

Amongst these techniques, the CP-based or preamble-based techniques are readily adapted to SC-CP and FD KSP schemes (cyclostationarity is also present in KSP schemes where FD equalization is possible), while the pilot-based techniques cannot be used.

In KSP transmission, time and frequency offsets can be estimated accurately relying on the knowledge of the padded sequences (see e.g. [8], [10], [11]), thereby allowing to avoid the insertion of bandwidth-consuming pilot tones or reducing the importance of the preamble. Hence, for FD-KSP schemes, one can expect a more accurate estimation and compensation of the CFO than in OFDM.

For BDFE and optimal linear equalizers, the compensation of the CFO is not really crucial, but can be realized accurately using the KSP-specific techniques mentioned above.

c) Narrowband Interference and Out of Band Noise

Another important OFDM implementational issue directly results from the spectral shape of the DFT filters that are used at the receiver. This spectrum is illustrated in **Fig 3.8**, where the power spectral density of the filter that estimates the 8th tone of a 16-carrier system is shown. This figure shows the

presence of important side-lobes that are characterized by a quite slow decay. When a narrowband interference is present in the spectrum of the received signal it will harm all the subcarriers rather than disturbing only those whose frequencies are close to that of the interfering signal as one would have expected. In SC-CP or KSP transmission, the presence of a narrowband interferer simply results in colored noise and its impact can be mitigated by adapting the MMSE equalizers to the actual noise correlation. This narrowband interference problem is rather important for the WLAN applications that often operate in unlicensed bands where they are confronted to other applications that pollute the spectrum. The 2.4 GHz band that is used by WLAN standards is also employed by Bluetooth devices, which is an importance source of interference.

Another consequence of these side-lobes arises only in OFDM transmission and results from the use of the IDFT filter at the transmitter side. The spectral shape of the IDFT filters is similar to that of the DFT and it thus exhibits similar side-lobes. In wireless applications, the transmitter is only allowed to transmit within a limited portion of the spectrum with strict spectral mask requirements. If data are modulated on the carriers that are located at the edge of the available spectrum, the side-lobes of the IDFT filters will cause *out-of band noise*, which cannot be tolerated. Hence, it is often needed to switch off the outer tones of OFDM systems in order to meet the spectral mask requirements. In Hiperlan2 for instance, 12 tones from a total of 64 remain unused because of this problem.

This problem is not present in SC-CP and KSP transmission and none of the transmitted symbols must remain unused to avoid radiating power outside the dedicated spectrum. Replacing OFDM by either of these techniques in a Hiperlan2-like setup would then increase the effective data throughput by 21.43%. However, these SC-CP and KSP systems require the use of pulse shaping filters (see *Ch. 2*), which always have a non-zero excess bandwidth in practical systems. Hence, the possible throughput gain resulting from the absence of IFFT filters at the transmitter is compensated at least in part by the excess bandwidth of these transmit filters.

In order to mitigate these effects in OFDM applications, several techniques have been developed that reduce the side-lobes of the FFT/IFFT filters. The most popular solution is called *windowing* [53], [54], [55] and consists of multiplying the time-domain signal with a time-domain window with relatively smooth edges.

3.5 Conclusions

In this chapter, we have presented channel equalization schemes that are suited to the different block transmission techniques. Their performance has been an-

alyzed in the light of the WLAN standards IEEE802.11a and Hiperlan2. From the presented perspective, KSP appears to be an interesting alternative to the traditional OFDM scheme as adopted in the standards. Uncoded KSP appears to be especially interesting in good channel conditions where it outperforms coded OFDM in terms of achievable data rate under similar channel conditions. Coded KSP appears to be most interesting in poor channel conditions as the SNR requirements are significantly less stringent than in OFDM. Overall, KSP seems to have several advantages over OFDM.

The *first advantage* results from the implicit encoding of the transmitted information over the tones: as a result one can avoid the use of error correcting codes, thereby reducing the complexity of the transmission scheme (no Viterbi decoding is needed at the receiver) and increasing the achieved data rates (the redundancy of the error correcting codes is reduced) whilst operating at lower SNR than required by the OFDM scheme.

The *second advantage* resides in the flexibility offered to the receiver for the choice of the equalization scheme. When the padded sequences are designed appropriately, it is possible to use the computationally efficient frequency-domain equalizer whose complexity is comparable to that of the classical OFDM equalizers. The low-complexity FD equalizers offer acceptable performance, but if the receiver has enough processing power, the system performance can be significantly improved by the use of either the BDFE or the optimal linear KSP equalizer.

Other advantages of KSP include the reduced impact of the implementational problems discussed in this chapter (PAPR, CFO, out of band noise) on the system performance or the possible exploitation of the padded sequences towards other goals than the prevention of IBI (e.g. synchronization, carrier frequency offset estimation, channel estimation).

Chapter 4

Channel Estimation

In order to enable the design of the previously discussed block equalizers, the receiver needs to rely on an estimate of the actual transmission channel. The final BER performance of the transmission system largely depends on the accuracy of the available channel estimate. In this chapter, we analyze the specific channel estimation possibilities that are offered by the different block transmission techniques and verify the BER performance that are obtained *when realistic channel estimates are considered*.

Before presenting the strategies that are traditionally used with the different block transmission techniques in order to obtain the *Channel State Information (CSI)* and perform an experimental performance analysis, we introduce a new method that is well suited for the KSP transmission scheme.

In KSP, the short sequences of known symbols, which primarily act as guard bands between blocks of data symbols, can indeed also be viewed as short training sequences. The method we develop here allows to optimally exploit the knowledge of these repeatedly inserted training sequences towards the identification of the channel. However, as the presented method is applicable in other contexts than KSP, we present it in a slightly more general framework than the one we have considered until now; i.e. we also consider the situation where the training sequences are shorter than the channel order. This situation arises in KSP transmission when the channel is longer than the padded sequences, which results in IBI. It also arises in *Pilot Symbol Assisted Modulation (PSAM)* [56], a transmission scheme where blocks of data symbols are separated by a unique pilot symbol.

The discussion on the proposed channel estimation method in the KSP context indicates that changing the composition of the padded sequence after each block of data allows to increase significantly the accuracy of the channel estimate.

Nevertheless, this requirement on the composition of the padded sequences contradicts the requirement of constant padded sequences that is needed to enable the use of the low-complexity FD KSP equalizers. In this chapter, we propose a novel KSP scheme, namely *Shifted Known Symbol Padding (S-KSP)* that enables the use of the FD KSP equalizers whilst offering the possibility of estimating the channel accurately.

The structure of the chapter is as follows:

In a first part, we discuss the new channel estimation technique. In *sec. 4.1*, we introduce the channel estimation problem in the KSP context in more details. In *sec. 4.2*, we present the data model that will be used to derive the new method. In *sec. 4.3*, we derive an expression for the Gaussian likelihood function of a channel estimate and introduce some approximations on which we have to rely in order to derive low-complexity ML channel estimates. In *sec. 4.4*, we show that the proposed approximation has a negligible impact on the achievable performance of ML channel estimation through a CRB analysis. We then propose an iterative algorithm that converges to the ML channel estimate (*sec. 4.5*). We derive an approximate closed-form expression of the ML channel estimate, both for a constant (*sec. 4.6.1*) and for a changing training sequence (*sec. 4.6.2*). We then discuss the identifiability conditions of the proposed methods (*sec. 4.6.5*). We experimentally test the proposed methods and compare them with existing ML methods in *sec. 4.8*.

In *sec. 4.9*, we discuss the channel estimation possibilities offered by the different block transmission techniques and present the techniques that are traditionally used for channel estimation. We introduce a new block transmission technique (S-KSP) that is well suited for accurate channel estimation and low-complexity FD equalization in *sec. 4.10* and we experimentally analyze the performance of the different block transmission schemes when realistic channel estimates are considered in *sec. 4.11*. Finally, we draw some conclusions in *sec. 4.12*.

4.1 Introduction

Existing channel identification algorithms can be divided into three families which are termed “blind”, “semi-blind” and “training-based”.

Training-based techniques assume that known symbols (training or pilot symbols) are inserted in the transmitted signals. It is then possible to identify the channel at the receiver exploiting the knowledge of these symbols (see e.g. [57], [58] or [59]). Most existing methods require the training sequences to be significantly longer than the channel impulse response and only exploit the channel output samples that are not corrupted by unknown data symbols. In

Additive White Gaussian Noise (AWGN) conditions, the problem of performing *Maximum Likelihood (ML)* channel identification when only these data-free received symbols are used is then a simple least squares problem.

Blind algorithms on the other hand estimate the channel based on properties of the transmitted signals (finite alphabet properties, higher order statistics, cyclo-stationarity, ... see e.g. [60] or [61], and references therein) and do not exploit the knowledge of pilots or training sequences inserted in the stream of transmitted symbols.

Semi-blind techniques that have been proposed recently show improved performance as they exploit the knowledge of the pilots or training symbols as well as intrinsic properties of the transmitted signals towards the determination of the channel model. These techniques are most useful when the transmitted signals contain a mix of training symbols and unknown data symbols, in which case neither blind nor training-based techniques are optimal. These semi-blind methods appear to be well suited to the context of KSP where relatively short sequences of known symbols are repeatedly inserted in the stream of data symbols. A method that explicitly combines a blind and a training-based cost function was proposed in [62] and seems to be the first sub-optimal semi-blind method proposed in the literature. Later, this idea has also been adopted in [63], [64] and [65]. A deterministic method exploiting all the energy that is received from the training sequences was presented in [7] and is specifically suited to the KSP context. Several semi-blind ML channel estimation techniques have been proposed as well. Depending on the hypothesis upon which the expression of the likelihood function is built, one can distinguish between three families of ML methods: *Deterministic ML* for which the data symbols are considered as deterministic disturbances, *Gaussian ML* for which the data symbols are assumed to be Gaussian distributed and *Stochastic ML* where the true distribution of the data symbols is exploited. Some deterministic ML methods are presented in [66] for instance. In [67], a theoretical comparison of the Cramer-Rao Bounds (CRB) indicates that Gaussian ML methods outperform deterministic ML methods. In [68], [69] and [70] Gaussian ML methods are proposed. The method presented in [70] is based on the Expectation Maximization (EM) algorithm and will be used as a benchmark for our method. Stochastic ML methods theoretically perform better as they assume a more realistic distribution for the data symbols, but they suffer from a higher complexity. A method based on hidden Markov models and the EM algorithm was presented in [71], but performs significantly worse than the achievable CRB.

The method proposed in this chapter considers a SISO transmission scheme where short training sequences (possibly shorter than the channel impulse response length) are repeatedly inserted between blocks of data symbols, which corresponds to KSP or PSAM [56] transmission. As has been done until now, we consider multipath channels whose coherence time is significantly larger than the duration of a packet of data symbols. Hence, the channel model stays

constant during the transmission of the whole packet. We investigate appropriate semi-blind identification strategies for this specific context and propose a new Gaussian ML method for semi-blind channel identification. The proposed method is able to cope with arbitrarily short training sequences (possibly shorter than the channel impulse response length) and performs channel estimation jointly exploiting all the received symbols of that contain contributions from the training sequences.

Note that the method that is described in this chapter does not consider separately the problems of channel acquisition and tracking which is generally the case of practical channel estimation techniques. We consider the somewhat theoretical setup where all the data symbols of a packet are received and stored in a buffer. Channel estimation and equalization happens only after the full reception of the packet, which would yield important problems of latency and stack size in practical implementations. Hence, the results presented here must be considered as an upper bound on the achievable performance of practical methods that may be derived from the presented method. Note that the RLS method presented in *CH. 6* in the framework of doubly selective channels is a possible practical implementation of the method presented here.

The proposed method asymptotically achieves the CRB, has a small computational complexity and seems to outperform existing semi-blind methods in a similar context. For the sake of simplicity, we consider all the training sequences to have the same length, but it is straightforward to adapt the method to the more general case of training sequences of variable length. As the channel stays constant during the transmission of several blocks of data, it is possible to use several training sequences to construct the channel estimate. We investigate both the situation where the same training sequence is repeated after each block of data and the situation where the training sequence is changed after each block of data.

4.2 Data Model

While the transmission setup is similar to what has been used until now, we remind the data model for the sake of clarity. We consider a stationary channel $\mathbf{h} = [h[0], h[1], \dots, h[L]]^T$ of order L and a sequence $x[n]$ of symbols transmitted over the channel. The received sequence $y[n]$ is obtained using (2.7):

$$y[n] = \sum_{i=0}^L h[i]x[n-i] + \eta[n], \quad (4.1)$$

where $\eta[n]$ is the AWGN at the receiver.

As mentioned in the introduction, we consider a transmission scheme where

constant length training sequences are inserted between blocks of data symbols. There are two different possibilities: either the same training sequence is repeated after each block of data, or the training sequence is changed after each block. We refer to these two alternative schemes as the constant training sequence case and the changing training sequence case. As we will see later, these two alternative schemes yield different channel identification procedures. We describe the more general situation of a changing training sequence as often as possible, and only analyze the case of a constant training sequence when explicitly needed.

A total number of K training sequences is inserted in the stream of unknown data symbols. The k^{th} training sequence, $\mathbf{t}_k = [t_k[1], \dots, t_k[N_t]]^T$, starts at position n_k : $[x[n_k], \dots, x[n_k + N_t - 1]]^T = \mathbf{t}_k$. Define the vector \mathbf{u}_k of received symbols that contain a contribution from the k^{th} transmitted training sequence: $\mathbf{u}_k = [y[n_k], \dots, y[n_k + N_t + L - 1]]^T$. The vector \mathbf{u}_k contains a contribution from the training sequence \mathbf{t}_k plus an additional term that collects the contributions from both the unknown surrounding data symbols and the noise. We can thus describe \mathbf{u}_k as the sum of a deterministic and a stochastic term:

$$\mathbf{u}_k = \mathbf{T}_k \mathbf{h} + \boldsymbol{\epsilon}_k, \quad (4.2)$$

where \mathbf{T}_k is an $(N_t + L) \times (L + 1)$ tall Toeplitz matrix with $[\mathbf{t}_k^T, 0, \dots, 0]^T$ as its first column and $[t_k[1], 0, \dots, 0]$ as its first row. $\mathbf{T}_k \mathbf{h}$ is the deterministic term; the stochastic term, $\boldsymbol{\epsilon}_k$, is described as follows:

$$\boldsymbol{\epsilon}_k = \underbrace{\begin{bmatrix} h_L & \cdots & h_1 & & & & & & \mathbf{0} \\ & & & \ddots & & & & & \\ & & & & \vdots & & & & \\ & & & & & h_0 & & & \\ & & & & & \vdots & \ddots & & \\ & & & & & & & \ddots & \\ \mathbf{0} & & & & h_L & & & & \\ & & & & & h_{L-1} & \cdots & & h_0 \end{bmatrix}}_{\mathbf{H}_s(N_t + L) \times (2L)} \check{\mathbf{s}}_k + \boldsymbol{\eta}_k, \quad (4.3)$$

where $\check{\mathbf{s}}_k = [s_k[1], \dots, s_k[2L]] = [x[n_k - L], \dots, x[n_k - 1], x[n_k + N_t], \dots, x[n_k + N_t + L - 1]]^T$ is the vector of surrounding data symbols, and $\boldsymbol{\eta}_k = [\eta[n_k], \dots, \eta[n_k + N_t + L - 1]]$ is the AWGN term. Assuming that both the noise and the data are white and zero-mean ($E\{s_k[i]s_k[j]^*\} = E\{\eta[i]\eta[j]^*\} = 0, \forall i, j, k : i \neq j$, and $E\{s_k[i]\} = E\{\eta[k]\} = 0$), we can say that $\boldsymbol{\epsilon}_k$ is zero-mean. Defining n_s as the length of the shortest sequence of data symbols ($n_s = \min_k\{n_{k+1} - (n_k + N_t - 1)\}$), we assume $n_s \geq 2L$. This ensures that the $\check{\mathbf{s}}_k$'s are uncorrelated, i.e. $E\{\check{\mathbf{s}}_k \check{\mathbf{s}}_l^H\} = \mathbf{0} \quad \forall k, l : k \neq l$. Defining the signal and noise variances as $\lambda^2 = E\{s_k[i]s_k[i]^*\}$ and $\sigma^2 = E\{\eta[k]\eta[k]^*\}$ respectively, we can derive the covariance matrix of $\boldsymbol{\epsilon}_k$ from (4.3) as $E\{\boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^H\} \triangleq \mathbf{Q} = \lambda^2 \mathbf{H}_s \mathbf{H}_s^H + \sigma^2 \mathbf{I}$. The first and second order statistics of the stochastic term are thus as follows:

$$E\{\boldsymbol{\epsilon}_k\} = \mathbf{0}_{(N_t + L) \times 1},$$

$$\begin{aligned} \mathbb{E} \{ \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^H \} \triangleq \mathbf{Q} &= \lambda^2 \mathbf{H}_s \mathbf{H}_s^H + \sigma^2 \mathbf{I}, \\ \mathbb{E} \{ \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_l^H \} &= \mathbf{0}_{(N_t+L) \times (N_t+L)} \quad \forall k, l : k \neq l. \end{aligned} \quad (4.4)$$

4.3 Maximum Likelihood Approach for Channel Identification

As discussed in the introduction, there are two alternative ways of expressing the Likelihood function of our system. The expression of the deterministic likelihood function is obtained by considering the unknown data symbols as deterministic disturbances. The Gaussian likelihood function is established by making the hypothesis that these symbols are Gaussian variables, hence $\mathbf{u}_k = \mathcal{N}(\mathbf{T}_k \mathbf{h}, \mathbf{Q})$. It has been shown in [67], that the Gaussian hypothesis yields more accurate channel estimates than its deterministic counterpart. Adopting this hypothesis, we can express (up to a constant term) the negative Gaussian log-likelihood function of the system as:

$$-\mathcal{L} = K \ln |\mathbf{Q}| + \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}_k \mathbf{h})^H \mathbf{Q}^{-1} (\mathbf{u}_k - \mathbf{T}_k \mathbf{h}). \quad (4.5)$$

Relying on the definition of \mathbf{Q} , the log-likelihood can be expressed as a direct function of the unknown parameters \mathbf{h} and σ^2 . The corresponding ML channel estimate minimizes this expression w.r.t. \mathbf{h} and σ^2 . This minimization problem boils down to a computationally demanding $(L+2)$ -dimensional nonlinear search.

To overcome this complexity problem, we propose to disregard the structure of \mathbf{Q} , and ignore the relation that binds it to the parameters \mathbf{h} and σ^2 . We thus assume that the covariance matrix \mathbf{Q} of the stochastic term $\boldsymbol{\epsilon}_k$ can be any symmetric positive definite matrix, regardless of \mathbf{h} and σ^2 . This hypothesis turns the initial ML problem into a new one. We call the initial problem the *parametric ML problem*; the problem resulting from the proposed approximations will be called the *non-parametric ML problem*. The non-parametric ML channel estimate thus maximizes the likelihood function w.r.t \mathbf{h} and \mathbf{Q} (instead of \mathbf{h} and σ^2).

The approximations that are needed to formulate the non-parametric ML problem are summarized as:

- $\boldsymbol{\epsilon}_k$ is Gaussian distributed
- $\mathbf{Q} = \mathbf{Q}^H > 0$ does not depend on the channel model

These assumptions transform the parametric ML problem in \mathbf{h} and σ^2 into a new optimization problem which is separable in its two variables \mathbf{h} and \mathbf{Q} . We exploit this separability property in the next sections in order to solve the optimization problem in a less complex way than the $(L + 2)$ -dimensional nonlinear search of the parametric ML problem.

The solution of the non-parametric ML problem differs from the solution of the parametric ML. Hence, it is worthwhile to first check the impact of the proposed hypothesis on the accuracy of the resulting ML channel estimates. This is what we do in the next section through an analysis of the respective Cramer-Rao Bounds.

4.4 Cramer Rao Bounds

We show later in the text (see *sec. 4.7*) that the channel estimates derived from the non-parametric ML problem are consistent and thus asymptotically unbiased. The Cramer-Rao Bound (CRB) is a theoretical lower bound on the covariance matrix of such an unbiased estimate and can thus be used as a benchmark for the presented method. In this section, we analyze the impact of the non-parametric hypothesis on the accuracy of the derived channel estimate through this theoretical bound. It can be shown (see e.g. [72, p. 562]) that for any unbiased estimate $\hat{\mathbf{h}}$ of the parameter vector \mathbf{h} , the following inequality holds:

$$\text{cov}(\hat{\mathbf{h}}) \geq \mathcal{J}(\mathbf{h})^{-1},$$

where the covariance matrix of the channel estimate is defined as $\text{cov}(\hat{\mathbf{h}}) = E \left\{ (\hat{\mathbf{h}} - \mathbf{h})(\hat{\mathbf{h}} - \mathbf{h})^H \right\}$, and the Fisher Information Matrix (FIM) is defined as $\mathcal{J}(\mathbf{h}) = E \left\{ \left(\frac{\partial \mathcal{L}}{\partial \mathbf{h}} \right)^T \left(\frac{\partial \mathcal{L}}{\partial \mathbf{h}} \right) \right\}$, \mathcal{L} being (up to a constant term) the Gaussian log-likelihood function of the system.

Adopting the results presented in [67], the FIM of the *parametric ML problem* can be formulated as:

$$\mathcal{J}(\mathbf{h}) = 2 \begin{bmatrix} \text{Re}(\mathbf{J}_1) & -\text{Im}(\mathbf{J}_1) \\ \text{Im}(\mathbf{J}_1) & \text{Re}(\mathbf{J}_1) \end{bmatrix} + 2 \begin{bmatrix} \text{Re}(\mathbf{J}_2) & -\text{Im}(\mathbf{J}_2) \\ \text{Im}(\mathbf{J}_2) & \text{Re}(\mathbf{J}_2) \end{bmatrix}, \quad (4.6)$$

where

$$\mathbf{J}_1(i, j) = \left(\sum_{k=1}^K \mathbf{T}_k^H \mathbf{Q}^{-1} \mathbf{T}_k \right) (i, j) + \text{tr} \left\{ \mathbf{Q}^{-1} \frac{\partial \mathbf{Q}}{\partial h[i-1]^*} \mathbf{Q}^{-1} \frac{\partial \mathbf{Q}}{\partial h[j-1]^*} \right\}, \quad (4.7)$$

$$\mathbf{J}_2(i, j) = \text{tr} \left\{ \mathbf{Q}^{-1} \frac{\partial \mathbf{Q}}{\partial h[i-1]^*} \mathbf{Q}^{-1} \frac{\partial \mathbf{Q}}{\partial h[j-1]^*} \right\}, \quad (4.8)$$

and

$$\frac{\partial \mathbf{Q}}{\partial h[i]^*} = \lambda^2 \mathbf{H}_s \left(\frac{\partial \mathbf{H}_s}{\partial h[i]^*} \right). \quad (4.9)$$

The approximation inserted in the *non-parametric ML problem* simplifies the expression of the FIM since it makes the $\frac{\partial \mathbf{Q}}{\partial h[i]^*}$ terms equal to zero. The expression of the FIM then becomes:

$$\mathcal{J}(\mathbf{h}) = \left(\sum_{k=1}^K \mathbf{T}_k^H \mathbf{Q}^{-1} \mathbf{T}_k \right). \quad (4.10)$$

Since the traces in (4.7) and (4.8) are always positive, the *CRB of the parametric ML problem will always be tighter than the CRB of its non-parametric counterpart*. However, numerical evaluations of the CRB in realistic situations show that the impact of these trace terms is negligible: the relative difference between the two CRBs is less than 10^{-4} for experimental setups similar to the ones that are used in *sec. 4.8*. This shows that the impact of the hypothesis of an unstructured \mathbf{Q} is very limited. The difference between the solutions of the parametric and non-parametric ML problems will hardly be noticeable.

We can thus safely work under the proposed hypothesis and consider the solutions we will obtain in this framework as the true ML channel estimates.

Further in the text, the CRB will be used as a benchmark for the proposed methods. The CRB can be evaluated in the non-parametric framework as:

$$\boxed{\mathcal{J}(\mathbf{h})^{-1} = \left(\sum_{k=1}^K \mathbf{T}_k^H \mathbf{Q}^{-1} \mathbf{T}_k \right)^{-1}} \quad (4.11)$$

Note that this bound depends both on the channel realization (through the covariance matrix \mathbf{Q}) and on the chosen training sequences (through the training sequence matrices \mathbf{T}_k).

4.5 Iterative Procedure

When a minimization problem is separable in its variables, a common approach to find the solution is an iterative one. One iteration consists in analytically

minimizing the cost function with respect to one variable whilst keeping the other(s) fixed. The variable with respect to which the cost function is minimized is changed in each iteration (see e.g. [73] where this approach is used to jointly estimate the transmitted data symbols and the channel). This procedure converges to a minimum of the cost function. If the starting point is accurate enough, the point of convergence is the global minimum of the cost function. In the sequel, we apply this approach to the likelihood function of the system, which leads to the ML estimate of \mathbf{Q} and \mathbf{h} .

Assume that at the i^{th} iteration an estimate $\hat{\mathbf{Q}}^i$ of the covariance matrix \mathbf{Q} is available. We first seek the channel estimate $\hat{\mathbf{h}}^i$ that minimizes the cost function (4.5) with respect to \mathbf{h} for a fixed $\mathbf{Q} = \hat{\mathbf{Q}}^i$, i.e. we compute $\hat{\mathbf{h}}^i = \mathbf{h}_{ML}(\hat{\mathbf{Q}}^i)$, where $\mathbf{h}_{ML}(\mathbf{Q}) = \arg \min_{\mathbf{h}} -\mathcal{L}$. The solution to this optimization problem can be computed as:

$$\mathbf{h}_{ML}(\mathbf{Q}) = \left(\sum_{k=1}^K \mathbf{T}_k^H \mathbf{Q}^{-1} \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \mathbf{Q}^{-1} \mathbf{u}_k. \quad (4.12)$$

We then seek the covariance matrix $\hat{\mathbf{Q}}^{i+1}$ that minimizes (4.5) with respect to \mathbf{Q} for a fixed $\mathbf{h} = \hat{\mathbf{h}}^i$: $\hat{\mathbf{Q}}^{i+1} = \mathbf{Q}_{ML}(\hat{\mathbf{h}}^i)$, where $\mathbf{Q}_{ML}(\mathbf{h}) = \arg \min_{\mathbf{Q}} -\mathcal{L}$, the solution to this optimization problem can be computed as¹:

$$\mathbf{Q}_{ML}(\mathbf{h}) = K^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}_k \mathbf{h}) (\mathbf{u}_k - \mathbf{T}_k \mathbf{h})^H. \quad (4.13)$$

$\hat{\mathbf{Q}}^{i+1}$ is then used as a starting point for the next iteration. The procedure is stopped when there is no significant difference between the estimates produced by two consecutive iterations. As will be shown later in the text (see *sec. 4.6.4*), this iterative method yields very good convergence properties when it is initialized with a least squares channel estimate. It is straightforward to show that applying the proposed method with an identity matrix as initial covariance matrix yields exactly this LS channel estimate as initial channel estimate. We thus propose to initialize the iterative ML method with $\hat{\mathbf{Q}}^0 = \mathbf{I}$.

The iterative procedure is thus summarized as follows:

¹The derivation of this result is detailed in Appendix A

Initialization:	$\hat{\mathbf{Q}}^0 = \mathbf{I}$
Iterations:	$\hat{\mathbf{h}}^i = \left(\sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}^{i-1} \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}^{i-1} \mathbf{u}_k$ $\hat{\mathbf{Q}}^{i+1} = K^{-1} \sum_{k=1}^K \left(\mathbf{u}_k - \mathbf{T}_k \hat{\mathbf{h}}^i \right) \left(\mathbf{u}_k - \mathbf{T}_k \hat{\mathbf{h}}^i \right)^H$

4.6 Closed Form Solution

An alternative strategy to the iterative procedure described above consists of directly finding an analytical expression for the global minimum of the likelihood function (4.5). The separability property of the cost function can be exploited again in order to find this global minimum. The idea is to analytically minimize the cost function with respect to one variable. This minimum is a function of the other variable. The first variable can then be eliminated in the original cost function, which then becomes a single variable expression. When the problem is separable in its two variables, minimizing this new expression of the cost function with respect to the only variable left yields the global minimum (see e.g. [74]). This approach is often used in ML problems and is known as the process of *compressing* or *concentrating* the Likelihood function onto one variable (see e.g. [75] or [76]). In this section, we concentrate the likelihood function on the variable \mathbf{h} in order to derive the closed-form channel estimate.

4.6.1 Constant Training Sequence

In order to indicate that the training sequence after each block is the same, we simply omit the block index (subscript k) for the vector \mathbf{t} and the matrix \mathbf{T} . We first minimize the cost function with respect to \mathbf{Q} , the solution of which is given by (4.13) with a constant \mathbf{T} matrix:

$$\mathbf{Q}_{ML}(\mathbf{h}) = K^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}\mathbf{h}) (\mathbf{u}_k - \mathbf{T}\mathbf{h})^H. \quad (4.14)$$

Replacing \mathbf{Q} by $\mathbf{Q}_{ML}(\mathbf{h})$ in the likelihood function as expressed in (4.5) yields the following expression of the concentrated likelihood function:

$$-\mathcal{L} = K \text{tr}(\mathbf{I}) + K \ln \left| K^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}\mathbf{h}) (\mathbf{u}_k - \mathbf{T}\mathbf{h})^H \right| + \text{cst.} \quad (4.15)$$

The ML channel estimate \mathbf{h}_{ML} minimizes (4.15) with respect to \mathbf{h} :

$$\mathbf{h}_{ML} = \arg \min_{\mathbf{h}} \left| K^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}\mathbf{h})(\mathbf{u}_k - \mathbf{T}\mathbf{h})^H \right|. \quad (4.16)$$

Define

$$\begin{aligned} \hat{\mathbf{R}} &\triangleq K^{-1} \sum_{k=1}^K \mathbf{u}_k \mathbf{u}_k^H, \\ \bar{\mathbf{u}} &\triangleq K^{-1} \sum_{k=1}^K \mathbf{u}_k, \\ \hat{\mathbf{Q}} &\triangleq \hat{\mathbf{R}} - \bar{\mathbf{u}} \bar{\mathbf{u}}^H, \end{aligned} \quad (4.17)$$

where $\hat{\mathbf{Q}}$ is assumed to be positive definite². Using these definitions, the matrix in the minimization problem (4.16) can be re-expressed as:

$$\begin{aligned} & K^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}\mathbf{h})(\mathbf{u}_k - \mathbf{T}\mathbf{h})^H \\ &= K^{-1} \left(\sum_{k=1}^K \mathbf{u}_k \mathbf{u}_k^H - \sum_{k=1}^K \mathbf{u}_k \mathbf{h}^H \mathbf{T}^H - \sum_{k=1}^K \mathbf{T}\mathbf{h} \mathbf{u}_k^H + \sum_{k=1}^K \mathbf{T}\mathbf{h} \mathbf{h}^H \mathbf{T}^H \right) \\ &= \hat{\mathbf{R}} - \bar{\mathbf{u}} \mathbf{h}^H \mathbf{T}^H - \mathbf{T}\bar{\mathbf{u}} \mathbf{h}^H + \mathbf{T}\mathbf{h} \mathbf{h}^H \mathbf{T}^H \\ &= \hat{\mathbf{R}} + (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H - \bar{\mathbf{u}} \bar{\mathbf{u}}^H \\ &= \hat{\mathbf{Q}} + \hat{\mathbf{Q}} \hat{\mathbf{Q}}^{-1} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H \\ &= \hat{\mathbf{Q}} \left(\mathbf{I} + \hat{\mathbf{Q}}^{-1} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H \right), \end{aligned}$$

Keeping in mind that $\hat{\mathbf{Q}}$ is positive definite, our minimization problem (4.16) is thus equivalent to:

$$\mathbf{h}_{ML} = \arg \min_{\mathbf{h}} \left| \mathbf{I} + \hat{\mathbf{Q}}^{-1} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H \right|. \quad (4.18)$$

It can be shown³ that

$$\left| \mathbf{I} + \hat{\mathbf{Q}}^{-1} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H \right| = 1 + (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H \hat{\mathbf{Q}}^{-1} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}}).$$

²We see from (4.17) that a necessary condition for this to hold is $K \geq N_t + L$. When this condition is fulfilled, the randomness of the noise and the data ensures that $\hat{\mathbf{Q}}$ is a positive definite matrix with probability 1.

³We detail the proof in Appendix B.

Hence, the minimization problem (4.18) is equivalent to:

$$\mathbf{h}_{ML} = \arg \min_{\mathbf{h}} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H \hat{\mathbf{Q}}^{-1} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}}). \quad (4.19)$$

The solution is obtained by nulling the partial derivative of this expression with respect to \mathbf{h} , which yields:

$$\boxed{\mathbf{h}_{ML} = \left(\mathbf{T}^H \hat{\mathbf{Q}}^{-1} \mathbf{T}\right)^{-1} \left(\mathbf{T}^H \hat{\mathbf{Q}}^{-1} \bar{\mathbf{u}}\right)}. \quad (4.20)$$

This ML channel estimate is easy to compute and also intuitively quite appealing for it shows that the ML channel estimate is simply a fit of $\mathbf{T}\mathbf{h}$ to $\bar{\mathbf{u}}$ in a weighted least squares sense.

Once the ML channel estimate is obtained, we can derive the corresponding ML covariance matrix estimate. First observe that, using the notations introduced in (4.17), the expression (4.14) for the ML estimate of \mathbf{Q} as a function of \mathbf{h} can be rewritten as:

$$\begin{aligned} \mathbf{Q}_{ML}(\mathbf{h}) &= \hat{\mathbf{R}} - \bar{\mathbf{u}}\mathbf{h}^H \mathbf{T}^H - \mathbf{T}\mathbf{h}\bar{\mathbf{u}}^H + \mathbf{T}\mathbf{h}\mathbf{h}^H \mathbf{T}^H \\ &= \hat{\mathbf{Q}} + (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H. \end{aligned} \quad (4.21)$$

\mathbf{Q}_{ML} is then derived by inserting \mathbf{h}_{ML} into this expression:

$$\boxed{\mathbf{Q}_{ML} = \mathbf{Q}_{ML}(\mathbf{h}_{ML}) = \hat{\mathbf{Q}} + (\mathbf{T}\mathbf{h}_{ML} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h}_{ML} - \bar{\mathbf{u}})^H}, \quad (4.22)$$

which differs from $\hat{\mathbf{Q}}$ by a rank-one term.

4.6.2 Changing Training Sequence

We proceed in the same way as for the constant training sequence case. First observe that the likelihood function (4.5) can be expressed as:

$$-\mathcal{L} = K \ln |\mathbf{Q}| + \text{tr} \left(\mathbf{Q}^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}_k \mathbf{h})(\mathbf{u}_k - \mathbf{T}_k \mathbf{h})^H \right). \quad (4.23)$$

We first minimize this cost function with respect to \mathbf{Q} leading to $\mathbf{Q}_{ML}(\mathbf{h})$ as given by (4.13). Replacing \mathbf{Q} by $\mathbf{Q}_{ML}(\mathbf{h})$ in (4.23) leaves us with an expression of the concentrated likelihood function that only depends on \mathbf{h} :

$$-\mathcal{L} = K \ln \left| K^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}_k \mathbf{h})(\mathbf{u}_k - \mathbf{T}_k \mathbf{h})^H \right| + K \text{tr}(\mathbf{I}). \quad (4.24)$$

The ML channel estimate is thus computed as:

$$\mathbf{h}_{ML} = \arg \min_{\mathbf{h}} \left| \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}_k \mathbf{h}) (\mathbf{u}_k - \mathbf{T}_k \mathbf{h})^H \right|. \quad (4.25)$$

Although this problem seems similar to (4.16), the varying \mathbf{T}_k forces us to adopt a different approach which will only lead us to an approximate solution. Let us first introduce the following notations:

$$\begin{aligned} \mathbf{h}_{LS} &\triangleq \left(\sum_{k=1}^K \mathbf{T}_k^H \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \mathbf{u}_k, \\ \mathbf{g} &\triangleq \mathbf{h} - \mathbf{h}_{LS}, \\ \mathbf{g}_{ML} &\triangleq \mathbf{h}_{ML} - \mathbf{h}_{LS}, \\ \mathbf{e}_k &\triangleq \mathbf{u}_k - \mathbf{T}_k \mathbf{h}_{LS}, \\ \hat{\mathbf{Q}}' &\triangleq K^{-1} \sum_{k=1}^K \mathbf{e}_k \mathbf{e}_k^H. \end{aligned} \quad (4.26)$$

where $\hat{\mathbf{Q}}'$ is assumed to be positive definite⁴. Using these notations, the minimization problem (4.25) can be rephrased as:

$$\mathbf{g}_{ML} = \arg \min_{\mathbf{g}} \left| \sum_{k=1}^K (\mathbf{e}_k - \mathbf{T}_k \mathbf{g}) (\mathbf{e}_k - \mathbf{T}_k \mathbf{g})^H \right|. \quad (4.27)$$

The determinant in this last expression can be expressed as (up to a positive factor):

$$\left| \mathbf{I} + \hat{\mathbf{Q}}'^{-1} K^{-1} \sum_{k=1}^K [\mathbf{T}_k \mathbf{g} \mathbf{g}^H \mathbf{T}_k^H - \mathbf{T}_k \mathbf{g} \mathbf{e}_k^H - \mathbf{e}_k \mathbf{g}^H \mathbf{T}_k^H] \right|. \quad (4.28)$$

When K is large, both \mathbf{h}_{LS} and \mathbf{h}_{ML} are close to the true \mathbf{h} (this hypothesis is confirmed by the experimental results presented in *sec. 4.8*). We can thus assume that \mathbf{g} , and consequently the second term in (4.28), is small in the vicinity of the solution. It is well known that, for $\|\Delta\| \ll 1$, $|\mathbf{I} + \Delta| \approx 1 + \text{tr}(\Delta)$. Hence, for $K \gg 1$, the minimization problem in (4.27) can be approximated by:

$$\hat{\mathbf{g}}_{ML} = \arg \min_{\mathbf{g}} \text{tr} \left(\hat{\mathbf{Q}}'^{-1} \sum_{k=1}^K [\mathbf{T}_k \mathbf{g} \mathbf{g}^H \mathbf{T}_k^H - \mathbf{T}_k \mathbf{g} \mathbf{e}_k^H - \mathbf{e}_k \mathbf{g}^H \mathbf{T}_k^H] \right),$$

⁴We see from (4.26) that a necessary condition for this to hold is $K \geq N_t + L$. When this condition is fulfilled, the randomness of the noise and the data ensures that $\hat{\mathbf{Q}}'$ is a positive definite matrix with probability 1.

where $\hat{\mathbf{g}}_{ML}$ is an approximation of \mathbf{g}_{ML} . Exploiting the permutation property of the trace of a product, this problem can be rephrased as:

$$\hat{\mathbf{g}}_{ML} = \arg \min_{\mathbf{g}} \sum_{k=1}^K \left[\mathbf{g}^H \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k \mathbf{g} - \mathbf{e}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k \mathbf{g} - \mathbf{g}^H \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{e}_k \right]. \quad (4.29)$$

The solution to this minimization problem is obtained by nulling the partial derivative of this expression with respect to \mathbf{g} , and is given as:

$$\hat{\mathbf{g}}_{ML} = \left(\sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{e}_k$$

We know from (4.26) that $\mathbf{h} = \mathbf{g} + \mathbf{h}_{LS}$. If we additionally replace \mathbf{e}_k by $\mathbf{u}_k - \mathbf{T}_k \mathbf{h}_{LS}$, we obtain the following approximation $\hat{\mathbf{h}}_{ML}$ of the channel estimate \mathbf{h}_{ML} :

$$\begin{aligned} \hat{\mathbf{h}}_{ML} = \mathbf{h}_{LS} &+ \left(\sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{u}_k \\ &- \left(\sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k \mathbf{h}_{LS}, \end{aligned}$$

which effectively simplifies into:

$$\boxed{\hat{\mathbf{h}}_{ML} = \left(\sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{u}_k.} \quad (4.30)$$

Based on this approximation of the ML channel estimate, it is possible to derive an approximation $\hat{\mathbf{Q}}_{ML}$ of the ML covariance matrix \mathbf{Q}_{ML} by inserting $\hat{\mathbf{h}}_{ML}$ into (4.13). First note that, using (4.26), the expression (4.13) of $\mathbf{Q}_{ML}(\mathbf{h})$ can be rewritten as:

$$\begin{aligned} \mathbf{Q}_{ML}(\mathbf{h}) = \hat{\mathbf{Q}}' &+ K^{-1} \sum_{k=1}^K (\mathbf{T}_k \mathbf{h} - \mathbf{u}_k) (\mathbf{T}_k \mathbf{h} - \mathbf{u}_k)^H \\ &- K^{-1} \sum_{k=1}^K (\mathbf{T}_k \hat{\mathbf{h}}_{LS} - \mathbf{u}_k) (\mathbf{T}_k \hat{\mathbf{h}}_{LS} - \mathbf{u}_k)^H. \end{aligned} \quad (4.31)$$

The approximate ML covariance matrix $\hat{\mathbf{Q}}_{ML}$ is obtained by replacing \mathbf{h} by $\hat{\mathbf{h}}_{ML}$ in this last expression:

$$\hat{\mathbf{Q}}_{ML} = \hat{\mathbf{Q}}' + K^{-1} \sum_{k=1}^K (\mathbf{T}_k \mathbf{h}_{ML} - \mathbf{u}_k) (\mathbf{T}_k \mathbf{h}_{ML} - \mathbf{u}_k)^H - K^{-1} \sum_{k=1}^K (\mathbf{T}_k \hat{\mathbf{h}}_{LS} - \mathbf{u}_k) (\mathbf{T}_k \hat{\mathbf{h}}_{LS} - \mathbf{u}_k)^H. \quad (4.32)$$

Note that a possible limitation of the proposed closed-form channel estimates comes from the fact that the channel is estimated only *after* the reception of a packet of data. Hence, all the channel output samples of the packet must first be stored in a possibly large buffer. The channel estimation procedure can start only when the whole packet or at least a significant number of data blocks has been received. The corresponding equalizers can then be computed and the data symbols estimated whilst a new packet is eventually being received. In *Ch. 6*, we propose a Recursive Least Squares (RLS) implementation of the closed form estimation described here. This RLS implementation, though not truly optimal can be adopted as a practically attractive alternative to the GML closed form estimate as it allows 'on the fly' channel estimation and equalization.

4.6.3 Noise Variance Estimate

Besides an accurate channel estimate, which is helpful in many applications, an estimate of the noise variance can be worthwhile obtaining as well. This is the case for instance when Minimum Mean Square Error (MMSE) equalizers are sought. When both the ML estimates \mathbf{h}_{ML} and \mathbf{Q}_{ML} (or approximations thereof) are available, it is possible to estimate the noise variance σ^2 as outlined below.

Using the available ML channel estimate, we rely on the definition (4.3) of the matrix \mathbf{H}_s in which we replace \mathbf{h} by \mathbf{h}_{ML} (or $\hat{\mathbf{h}}_{ML}$) to build $\hat{\mathbf{H}}_s$. If our estimates are accurate enough, subtracting $\lambda^2 \hat{\mathbf{H}}_s \hat{\mathbf{H}}_s^H$ from \mathbf{Q}_{ML} (or $\hat{\mathbf{Q}}_{ML}$) should leave us with a diagonal matrix with σ^2 repeated over the diagonal. We can thus use the following noise variance estimate:

$$\hat{\sigma}^2 = (N_t + L)^{-1} \text{tr} \left(\mathbf{Q} - \lambda^2 \hat{\mathbf{H}}_s \hat{\mathbf{H}}_s^H \right), \quad (4.33)$$

where \mathbf{Q} is replaced by \mathbf{Q}_{ML} (or $\hat{\mathbf{Q}}_{ML}$).

4.6.4 Comparison with the Iterative Method

In this section, we show that initializing the iterative method with $\hat{\mathbf{Q}}^0$ as recommended earlier indeed yields good convergence properties. Towards that goal, let us compare the estimates provided by the iterative method after one iteration with the closed form expressions derived above. We assume $K \gg 1$ throughout this discussion.

Let us first discuss the *constant training sequence* case. Initializing the algorithm with $\hat{\mathbf{Q}}^0 = \mathbf{I}$ yields $\hat{\mathbf{h}}^0 = (\mathbf{T}^H \mathbf{T})^{-1} \mathbf{T}^H \bar{\mathbf{u}}$. Next, exploiting the expression (4.21) of $\mathbf{Q}_{ML}(\mathbf{h})$, we can write $\hat{\mathbf{Q}}^1 = \hat{\mathbf{Q}} + (\mathbf{T}\hat{\mathbf{h}}^0 - \bar{\mathbf{u}}) (\mathbf{T}\hat{\mathbf{h}}^0 - \bar{\mathbf{u}})^H$.

Observing that $\hat{\mathbf{h}}^0$ is the least squares fit of $\mathbf{T}\mathbf{h}$ to $\bar{\mathbf{u}}$, we can say that $\mathbf{T}\hat{\mathbf{h}}^0 - \bar{\mathbf{u}}$ is small. The second term of the right-hand side of $\hat{\mathbf{Q}}^1$ is a second-order function of this small term. We can thus safely neglect it and make the following approximation: $\hat{\mathbf{Q}}^1 \approx \hat{\mathbf{Q}}$. Using (4.20), we then observe that $\hat{\mathbf{h}}^1$ is close to \mathbf{h}_{ML} :

$$\hat{\mathbf{h}}^1 = \mathbf{h}_{ML}(\hat{\mathbf{Q}}^1) \approx \mathbf{h}_{ML}(\hat{\mathbf{Q}}) = \mathbf{h}_{ML}.$$

When *changing training sequences* are used, it is straightforward to see that initializing the iterative method with $\hat{\mathbf{Q}}^0 = \mathbf{I}$ yields $\hat{\mathbf{h}}^0 = \hat{\mathbf{h}}_{LS}$. It then follows that $\hat{\mathbf{Q}}^1 = \hat{\mathbf{Q}}'$ and, observing the similarity between (4.12) and (4.30), $\hat{\mathbf{h}}^1 = \hat{\mathbf{h}}_{ML}$. Hence the iterative procedure yields the approximate closed form ML channel estimate after one iteration.

The first conclusion of this discussion is that in both constant and changing training sequences situations the iterative procedure has good convergence properties as the channel estimate obtained after one iteration, $\hat{\mathbf{h}}^1$, is a good approximation of the true ML channel estimate \mathbf{h}_{ML} in both cases. Experimental results indeed show that the iterative procedures almost converges in one single iteration in most cases and in 2 or 3 iterations in the worst case scenarios.

The second conclusion of this discussion is related to the practical interest of the iterative method. In this case the conclusions differs for constant and changing training sequences.

When *constant training sequences* are used, the closed-form channel estimate is the true ML channel estimate. The iterative method is then of little practical interest as it will at best reach the same accuracy as the closed-form estimate, but only after several iterations.

When *changing training sequences* are used, the situation is different. In this case, the closed-form solution proposed above is only an approximation of the true ML channel estimate. The iterative method reaches this closed-form esti-

mate after just one iteration, but keeps getting closer to the true ML channel estimate as the iterations proceed. Hence, when the closed-form expression is a sufficiently accurate approximation of the true channel estimate, the iterative method is not very useful. However, if the approximations that were needed to derive the closed-form estimate do not fully hold and the resulting channel estimate differs significantly from the true ML estimate, performing a few iterations will allow to improve the quality of the channel estimate. Simulation results that are detailed later in this chapter show that the approximate closed-form estimate is of sufficient accuracy in most practical situations but that there are still a few practical situations where the use of the iterative method significantly improves the accuracy of the channel estimate.

4.6.5 Identifiability Conditions

In theory (see e.g. [77] for a broader discussion on channel identifiability), the channel is identifiable in the considered Gaussian problem formulation when two conditions are fulfilled: The number of channel output samples considered for channel identification is larger than the channel order L and there is at least one non-zero training symbol not located at the edges of the considered burst.

The first condition is always fulfilled in the considered data model and the second is fulfilled as soon as there is a non-zero element in each training sequence.

However, as we need to invert the estimate of the matrix \mathbf{Q} in the proposed methods (both the iterative method and the closed-form solutions), this estimate has to have a full rank. Relying on the randomness of the noise and the unknown data symbols, this happens with probability 1 as soon as $K \geq N_t + L$. Hence, the identifiability conditions of the proposed method are summarized as follows:

$$\begin{array}{l} N_t \geq 1 \\ K \geq N_t + L \end{array}$$

4.6.6 Complexity Analysis

In this section, we briefly analyze the processing power that is required by the proposed closed-form channel estimation method. This analysis is useful for the assessment of the total computational complexity of the system as the channel estimation often requires a significant part of the available processing power. It is indeed common that the required complexity is comparable to that required by the equalizer.

First note that the total complexity of (4.30) is dominated by the operations that are repeated for each received block. The impact of the matrix inversions

present in (4.30) (the inversion of the sum and the inversion of $\hat{\mathbf{Q}}'$) on the total complexity is negligible.

Recalling that the total complexity of a matrix multiplication $\mathbf{A}_{(a \times b)}\mathbf{B}_{(b \times c)}$ is $a(2b - 1)c$ Flops, the computational complexity of the recurring operations can be computed as (Flops per block):

- Computation of \mathbf{h}_{LS} : $(L + 1)(2N_t + 2L - 1)$.
- Computation of \mathbf{e}_k : $(N_t + L)(2L + 1)$.
- Computation of $\hat{\mathbf{Q}}'$: $(N_t + L)^2$.
- Computation of $\mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k$: $(L + 1)(2N_t + 2L - 1)(N_t + 2L + 1)$.
- Computation of $\mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{u}_k$: $(L + 1)(2N_t + 2L - 1)(N_t + L + 1)$.

The total complexity is $\mathcal{O}(L^3)$. In order to have a better idea in terms of processing requirement, let us compute the total complexity of this method in terms of Flops/symbol for the experimental setup considered further in the text for the simulation results. In that setup, $L = 5$, $N_t = 8$ and the number of symbols per block is $N_s = 64$. The total complexity of the channel estimation step in this setup is approximately 5400 Flops per received block, or 85 Flops per received symbol. In comparison, the BDFE requires 190 Flops per symbol, while the optimal and FD linear equalizers for KSP require respectively 127 and 19 Flops per symbol.

4.7 Asymptotic Properties of the Closed Form Channel Estimates

In this section, we study the asymptotic properties of the proposed closed form ML channel estimates, that is their properties when the number of transmitted data blocks, K , is large. We find that the proposed closed form channel estimates are *consistent* and *asymptotically efficient* both for constant and changing training sequences.

4.7.1 Constant Training Sequence

Let us first note that $\hat{\mathbf{Q}}$ can be rewritten as:

$$\hat{\mathbf{Q}} = \hat{\mathbf{R}} - \bar{\mathbf{u}}\bar{\mathbf{u}}^H$$

$$= K^{-1} \sum_{k=1}^K \mathbf{u}_k \mathbf{u}_k^H - K^{-2} \sum_{k=1}^K \mathbf{u}_k \sum_{k=1}^K \mathbf{u}_k^H.$$

Keeping in mind that $\mathbf{u}_k = \mathbf{T}\mathbf{h} + \boldsymbol{\epsilon}_k$, we have:

$$\hat{\mathbf{Q}} = K^{-1} \sum_{k=1}^K \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^H - K^{-2} \sum_{i,j=1}^K \boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_j^H. \quad (4.34)$$

Using the central limit theorem, the above time averages can be replaced by their expected values when K tends to infinity:

$$\begin{aligned} \lim_{K \rightarrow \infty} K^{-1} \sum_{k=1}^K \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^H &= \mathbb{E} \{ \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^H \} = \mathbf{Q}, \\ \lim_{K \rightarrow \infty} K^{-2} \sum_{i,j=1}^K \boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_j^H &= \lim_{K \rightarrow \infty} K^{-1} \sum_{j=1}^K \mathbb{E} \{ \boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_j^H \} \\ &= \lim_{K \rightarrow \infty} K^{-1} \mathbf{Q} \\ &= 0. \end{aligned}$$

Therefore,

$$\lim_{K \rightarrow \infty} \hat{\mathbf{Q}} = \mathbf{Q}. \quad (4.35)$$

Replacing $\hat{\mathbf{Q}}$ by its limit \mathbf{Q} and defining the shorthand notations $\boldsymbol{\mathcal{Q}} = \mathbf{T}^H \mathbf{Q}^{-1} \mathbf{T}$ and $\hat{\boldsymbol{\mathcal{Q}}} = \mathbf{T}^H \hat{\mathbf{Q}}^{-1} \mathbf{T}$, it follows that the ML channel estimate (4.20) is *consistent*:

$$\begin{aligned} \lim_{K \rightarrow \infty} \mathbf{h}_{ML} &= \lim_{K \rightarrow \infty} K^{-1} \sum_{k=1}^K \hat{\boldsymbol{\mathcal{Q}}}^{-1} \mathbf{T}^H \hat{\mathbf{Q}}^{-1} \mathbf{u}_k \\ &= \mathbf{h} + \lim_{K \rightarrow \infty} K^{-1} \sum_{k=1}^K \boldsymbol{\mathcal{Q}}^{-1} \mathbf{T}^H \mathbf{Q}^{-1} \boldsymbol{\epsilon}_k \\ &= \mathbf{h} + \boldsymbol{\mathcal{Q}}^{-1} \mathbf{T}^H \mathbf{Q}^{-1} \mathbb{E} \{ \boldsymbol{\epsilon}_k \} \\ &= \mathbf{h}. \end{aligned} \quad (4.36)$$

Relying on these results, we can prove that the normalized sample covariance matrix of the ML channel estimate converges to the normalized CRB as K tends to infinity (i.e. *asymptotic efficiency*):

$$\lim_{K \rightarrow \infty} \mathbb{E} \{ K(\mathbf{h}_{ML} - \mathbf{h})(\mathbf{h}_{ML} - \mathbf{h})^H \}$$

$$\begin{aligned}
&= \mathbb{E} \left\{ \lim_{K \rightarrow \infty} \hat{\mathbf{Q}}^{-1} \mathbf{T}^H \hat{\mathbf{Q}}^{-1} K^{-1} \sum_{i=1}^K \boldsymbol{\epsilon}_i \sum_{j=1}^K \boldsymbol{\epsilon}_j^H \hat{\mathbf{Q}}^{-1} \mathbf{T} \hat{\mathbf{Q}}^{-1} \right\} \\
&= \mathbf{Q}^{-1} \mathbf{T}^H \mathbf{Q}^{-1} \lim_{K \rightarrow \infty} \mathbb{E} \left\{ K^{-1} \sum_{i,j=1}^K \boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_j^H \right\} \mathbf{Q}^{-1} \mathbf{T} \mathbf{Q}^{-1} \\
&= \mathbf{Q}^{-1} \mathbf{T}^H \mathbf{Q}^{-1} \sum_{j=1}^K \mathbb{E} \{ \boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_j^H \} \mathbf{Q}^{-1} \mathbf{T} \mathbf{Q}^{-1} \\
&= \mathbf{Q}^{-1} = (\mathbf{T}^H \mathbf{Q}^{-1} \mathbf{T})^{-1}. \tag{4.37}
\end{aligned}$$

This last expression is equal to the normalized CRB (see (4.11)).

4.7.2 Changing Training Sequence

Defining the shorthand notations $\mathbf{Q}_k = \mathbf{T}_k^H \mathbf{Q}^{-1} \mathbf{T}_k$, $\hat{\mathbf{Q}}_k = \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k$ and $\mathcal{P}_k = \mathbf{T}_k^H \mathbf{T}_k$, let us first note that \mathbf{e}_k can be rewritten as

$$\begin{aligned}
\mathbf{e}_k &= \mathbf{u}_k - \mathbf{T}_k \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \sum_{l=1}^K \mathbf{T}_l^H \mathbf{u}_l \\
&= \mathbf{T}_k \mathbf{h} + \boldsymbol{\epsilon}_k - \mathbf{T}_k \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \left(\sum_{l=1}^K \mathcal{P}_l \mathbf{h} + \boldsymbol{\epsilon}_l \right) \\
&= \boldsymbol{\epsilon}_k - \mathbf{T}_k \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \sum_{l=1}^K \mathbf{T}_l^H \boldsymbol{\epsilon}_l
\end{aligned}$$

Based on this observation, it is possible to check that the shorthand notation $\hat{\mathbf{Q}}'$ defined in (4.26) converges to the true \mathbf{Q} when K tends to infinity:

$$\begin{aligned}
\lim_{K \rightarrow \infty} \hat{\mathbf{Q}}' &= \lim_{K \rightarrow \infty} K^{-1} \sum_{k=1}^K \mathbf{e}_k \mathbf{e}_k^H \\
&= \lim_{K \rightarrow \infty} K^{-1} \sum_{k=1}^K \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^H - K^{-1} \sum_{k=1}^K \boldsymbol{\epsilon}_k \left(\sum_{l=1}^K \boldsymbol{\epsilon}_l^H \mathbf{T}_l \right) \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \mathbf{T}_k^H \\
&\quad - K^{-1} \sum_{k=1}^K \mathbf{T}_k \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \left(\sum_{l=1}^K \mathbf{T}_l^H \boldsymbol{\epsilon}_l \right) \boldsymbol{\epsilon}_k^H \\
&\quad + K^{-1} \sum_{k=1}^K \mathbf{T}_k \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \left(\sum_{l=1}^K \mathbf{T}_l^H \boldsymbol{\epsilon}_l \right) \left(\sum_{l=1}^K \boldsymbol{\epsilon}_l^H \mathbf{T}_l \right) \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \mathbf{T}_k^H.
\end{aligned}$$

Using the central limit theorem, we can replace the limit of $K^{-1} \sum_k$ by its expected value, which yields:

$$\begin{aligned} \lim_{K \rightarrow \infty} \hat{\mathbf{Q}}' &= \mathbf{Q} - \mathbf{Q} \mathbb{E} \left\{ \mathbf{T}_k \lim_{K \rightarrow \infty} \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \mathbf{T}_k^H \right\} \\ &\quad - \mathbb{E} \left\{ \mathbf{T}_k \lim_{K \rightarrow \infty} \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \mathbf{T}_k^H \right\} \mathbf{Q}^H \\ &\quad + \mathbb{E} \left\{ \mathbf{T}_k \lim_{K \rightarrow \infty} \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \left(\sum_{l=1}^K \mathcal{Q}_l \right)^{-1} \left(\sum_{l=1}^K \mathcal{P}_l \right)^{-1} \mathbf{T}_k^H \right\}. \end{aligned} \quad (4.38)$$

Since the training sequences have a non-zero energy, the limits present in this last expression are equal to zero. Hence, given the finite norm of \mathbf{Q} , all the terms containing these limits are equal to zero:

$$\lim_{K \rightarrow \infty} \hat{\mathbf{Q}}' = \mathbf{Q} \quad (4.39)$$

Replacing \mathbf{u}_k by its equivalent $\mathbf{T}_k \mathbf{h} + \boldsymbol{\epsilon}_k$ in (4.30) yields:

$$\hat{\mathbf{h}}_{ML} = \mathbf{h} + \left(\sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \boldsymbol{\epsilon}_k. \quad (4.40)$$

If the training sequences have a constant non-zero energy, it is clear that

$$\lim_{K \rightarrow \infty} K^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \mathbf{T}_k = \lim_{K \rightarrow \infty} K^{-1} \sum_{k=1}^K \mathbf{T}_k^H \mathbf{Q}^{-1} \mathbf{T}_k \neq \mathbf{0}. \quad (4.41)$$

Furthermore,

$$\lim_{K \rightarrow \infty} K^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}'^{-1} \boldsymbol{\epsilon}_k = \mathbb{E} \{ \mathbf{T}_k^H \mathbf{Q}^{-1} \boldsymbol{\epsilon}_k \}. \quad (4.42)$$

Since \mathbf{T}_k and $\boldsymbol{\epsilon}_k$ are independent, the expected value of the product is the product of the expected values. Since $\mathbb{E}\{\mathbf{T}_k\}$ is bounded and $\mathbb{E}\{\boldsymbol{\epsilon}_k\} = \mathbf{0}$, the limit is equal to zero. This shows that the approximate ML channel estimate is *consistent*:

$$\lim_{K \rightarrow \infty} \hat{\mathbf{h}}_{ML} = \mathbf{h}. \quad (4.43)$$

Here also, it is possible to derive an expression for the expected value of the normalized covariance matrix as K tends to infinity. Using the additional

shorthand notations $\mathcal{M}_k = \mathbf{T}_k^H \mathbf{Q}^{-1}$ and $\hat{\mathcal{M}}_k = \mathbf{T}_k^H \hat{\mathbf{Q}}^{-1}$, it is possible to show the *asymptotic efficiency* of the proposed channel estimate:

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \mathbb{E} \left\{ K \left(\hat{\mathbf{h}}_{ML} - \mathbf{h} \right) \left(\hat{\mathbf{h}}_{ML} - \mathbf{h} \right)^H \right\} \\
&= \mathbb{E} \left\{ \lim_{K \rightarrow \infty} \left(K^{-1} \sum_{k=1}^K \hat{\mathbf{Q}}_k \right)^{-1} K^{-1} \sum_{k=1}^K \mathcal{M}_k \boldsymbol{\epsilon}_k \sum_{k=1}^K \boldsymbol{\epsilon}_k^H \hat{\mathcal{M}}_k^H \left(K^{-1} \sum_{k=1}^K \hat{\mathbf{Q}}_k \right)^{-1} \right\} \\
&= \lim_{K \rightarrow \infty} \left(K^{-1} \sum_{k=1}^K \mathbf{Q}_k \right)^{-1} \mathbb{E} \left\{ K^{-1} \sum_{k=1}^K \mathcal{M}_k \boldsymbol{\epsilon}_k \sum_{k=1}^K \boldsymbol{\epsilon}_k^H \mathcal{M}_k^H \right\} \left(K^{-1} \sum_{k=1}^K \mathbf{Q}_k \right)^{-1} \\
&= \lim_{K \rightarrow \infty} \left(K^{-1} \sum_{k=1}^K \mathbf{Q}_k \right)^{-1} \mathbb{E} \{ \mathbf{Q}_k \} \left(K^{-1} \sum_{k=1}^K \mathbf{Q}_k \right)^{-1} \\
&= \lim_{K \rightarrow \infty} \left(K^{-1} \sum_{k=1}^K \mathbf{Q}_k \right)^{-1} K^{-1} \sum_{k=1}^K \mathbf{Q}_k \left(K^{-1} \sum_{k=1}^K \mathbf{Q}_k \right)^{-1} \\
&= \lim_{K \rightarrow \infty} \left(K^{-1} \sum_{k=1}^K \mathbf{T}_k^H \mathbf{Q}^{-1} \mathbf{T}_k \right)^{-1}
\end{aligned}$$

The above expression is equal to the normalized CRB (see (4.11)).

4.8 Simulation Results

The performance metric that is used throughout this section is the Normalized Mean Square Error (NMSE) of the proposed channel estimate:

$$\text{NMSE} = E \left\{ \frac{\|\hat{\mathbf{h}} - \mathbf{h}\|^2}{\|\mathbf{h}\|^2} \right\}. \quad (4.44)$$

The results that are presented are obtained with the closed form channel estimates (4.20) and (4.30). When the iterative method results are investigated, we explicitly state it in the text. We use the CRB as a benchmark in the experiments. The CRB curves displayed on the graphs represent the NMSE of an estimator that achieves the CRB, which is $E \left\{ \frac{\text{tr}(\mathcal{J}^{-1})}{\|\mathbf{h}\|^2} \right\}$. The experiments are performed on convolutive Rayleigh fading channels of varying order L , which correspond to the channel model that has been used in previous chapters. The training and data sequences are white QPSK sequences. The energy of the transmitted symbols (both data and training) is set to $\lambda^2 = 1$. The presented

results are obtained after averaging over a set of 100 channel realizations. For each of these channel realizations, the results are averaged over 100 different sets of training sequences in the changing training sequence case and over 100 different training sequences in the constant training sequence case. Note that this averaging is also done for the CRB results since the CRB depends both on the channel realization and the training sequences. The Signal to Noise Ratio (SNR) is defined as in previous chapters, i.e. $SNR = E\{\|\mathbf{h}\|^2\} \frac{\lambda^2}{\sigma^2}$.

In *sec. 4.8.1*, we analyze in details the channel estimation methods that have been discussed in this chapter. We simulate their performance in various channel conditions, for different compositions of the padded sequences and compare their performance to the CRB. In *sec. 4.8.2*, we compare the proposed method to recently proposed channel estimation techniques that can be used in the considered context.

4.8.1 Performance of the Proposed Method

In this section, we analyze and compare the algorithms proposed for the two situations that have been considered throughout the presented analysis: the constant and changing training sequence cases.

Comparison of the Cramer-Rao Bounds

To have a first insight in how the CRBs compare, we check the CRB performance for these two configurations. We consider a transmission scheme where the length of the training sequences is set to $N_t = 10$ and the number of observed training sequences is set to $K = 100$. The CRB for different channel orders in that context is presented as a function of the SNR in **Fig. 4.1**. We see that the use of changing training sequences systematically results in a reduced CRB for all channel orders. In both situations, the CRB decreases with a constant slope as the SNR increases when there is an exact solution to the channel identification problem in the noiseless case, i.e. when $N_t \geq 2L + 1$ for constant training sequences and when $N_t \geq L + 1$ for changing training sequences. The CRB saturates at high SNR, if there is no exact solution to the channel identification problem in the noiseless case, i.e. when $N_t < 2L + 1$ for constant training sequences and when $N_t < L + 1$ for changing training sequences. When N_t is fixed and the channel order is in the interval $\frac{N_t-1}{2} \leq L \leq N_t - 1$, using changing training sequences will yield a constant slope in the CRB for increasing SNR whereas a floor will appear at high SNR if a constant training sequence is used. For channel orders outside this interval, both methods show a similar behavior (constant slope for small channel orders and saturation for large channel orders), but there is still an advantage in using changing training sequences.

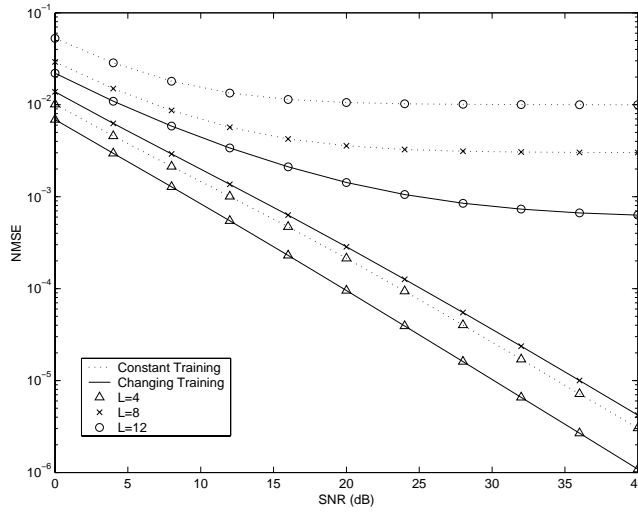


Figure 4.1: Comparison of the CRB of the constant and changing training sequence cases vs. the SNR for different channel orders when the length of the padded sequence is fixed to $N_t = 10$ and the number of data blocks considered for channel estimation is set to $K = 100$.

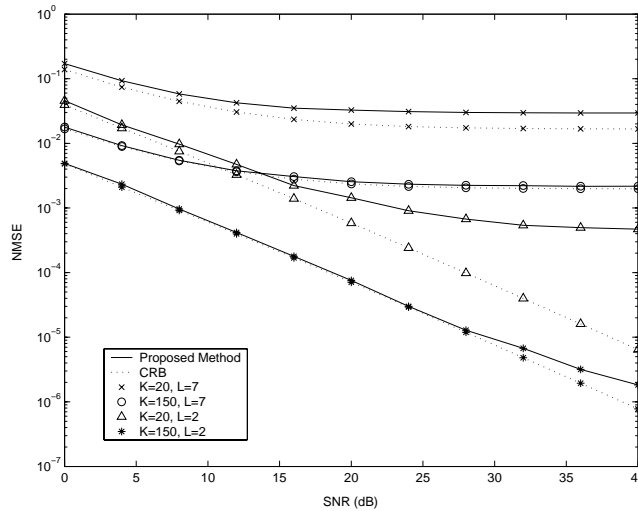


Figure 4.2: Comparison of the simulated NMSE and the CRB vs. SNR for different channel orders when $N_t = 5$ and changing training sequences are used. The results are plotted for two different values of K , namely 20 and 150

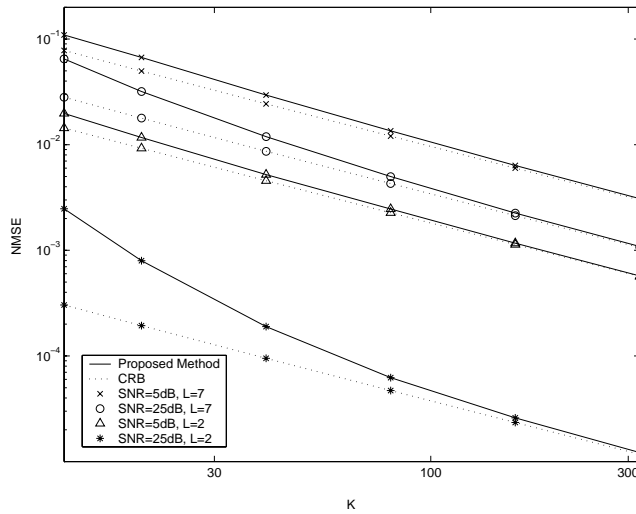


Figure 4.3: Comparison of the simulated NMSE and the CRB vs. K for different channel orders when $N_t = 5$ and changing training sequences are used. The results are plotted for two different values of the SNR, namely 5 and 25dB.

Changing Training Sequences

After this discussion on the CRB, we verify how the proposed closed form channel estimates match this theoretical bound. We first analyze the closed-form ML channel estimate proposed in the context of changing training sequences. In **Fig. 4.2**, we compare the simulated performance of our method with the corresponding CRB as a function of the SNR. We perform this comparison for two different channel orders: one for which the CRB has a constant slope, the other being large enough to have the CRB saturating at high SNR. We repeat these experiments for two different values of K : a relatively small one ($K = 20$) and a larger one ($K = 150$). When the channel order is large (saturation of the CRB at high SNRs), we observe a relatively good match between the CRB and the experimental curves. The match is tighter for a larger value of K . When the channel order is small (constant slope in the CRB curves), the theoretical and experimental curves match quite well when K is large, but we see the emergence of a floor on the experimental NMSE for higher SNR when K is small. In **Fig. 4.3**, we evaluate the impact of the number of data blocks K on the channel estimate NMSE. The simulations are done for two different channel orders and two different values of the SNR. The presented results confirm the asymptotical efficiency of the closed-form channel estimate as the experimental performance systematically achieves the CRB for large values of K . When K gets small, the match between the CRB and the experimental results remains

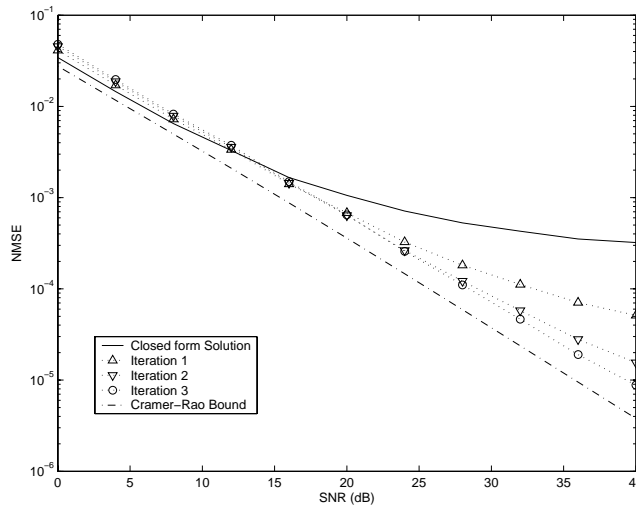


Figure 4.4: Comparison of the simulated NMSE and the CRB vs. SNR for $L = 2$, $N_t = 5$ and $K = 20$ using changing training sequences. When the iterative method is used, the NMSE converges to the CRB after a few iterations whereas the closed-form performs significantly worse than the CRB.

acceptable, except in the situation where the channel order is small and the SNR is high. In this case, the closed-form estimate is at a significant distance from the CRB. This difference between the CRB and the experimental results originates from the approximations that were needed to derive the approximate closed-form ML channel estimate. These approximations do not hold when the SNR is large, K is small and the channel order is small. However, when the iterative method is used, the channel estimate converges to \mathbf{h}_{ML} . The experiments presented in **Fig. 4.4** and **Fig. 4.5** show that the gap between the CRB and the closed-form estimate is closed after a few iterations. Hence, for large K and high SNRs, performing a few iterations allows us to effectively achieve the CRB.

Constant Training Sequences

In **Fig. 4.6** and **Fig. 4.7**, we perform a similar analysis for the closed form ML channel estimate in the context of a constant training sequence. The figures show us that there is no significant difference between the CRB and the experimental results, except for very small values of K . This improved match between the CRB and the experimental results originates in the fact that we did not have to make any approximation when deriving the expression of the closed-

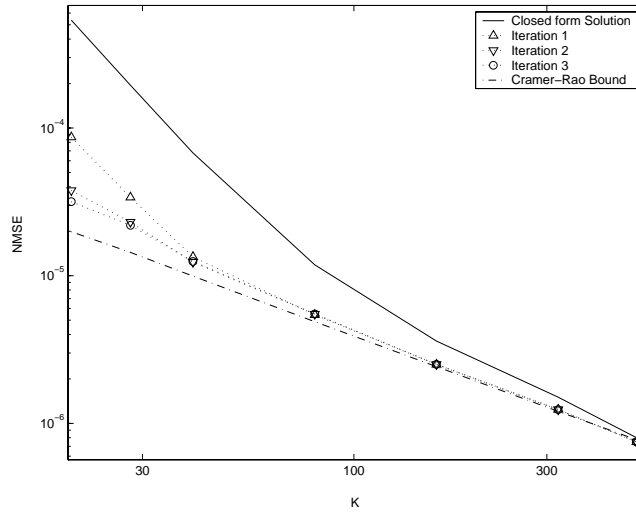


Figure 4.5: Comparison of the simulated NMSE and the CRB vs. K for $L = 2$, $N_t = 5$ and $\text{SNR}=35\text{dB}$ using changing training sequences. When the iterative method is used, the NMSE converges to the CRB after a few iterations whereas the closed-form performs significantly worse than the CRB.

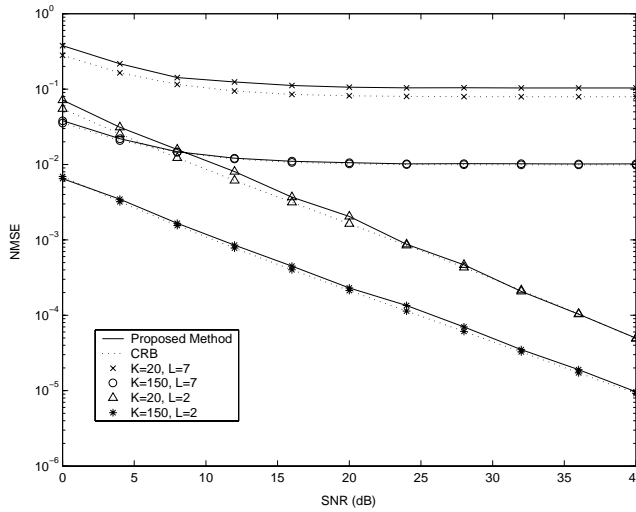


Figure 4.6: Comparison of the simulated NMSE and the CRB vs. SNR for different channel orders when $N_t = 5$ and constant training sequences are used. The results are plotted for two different values of K , namely 20 and 150

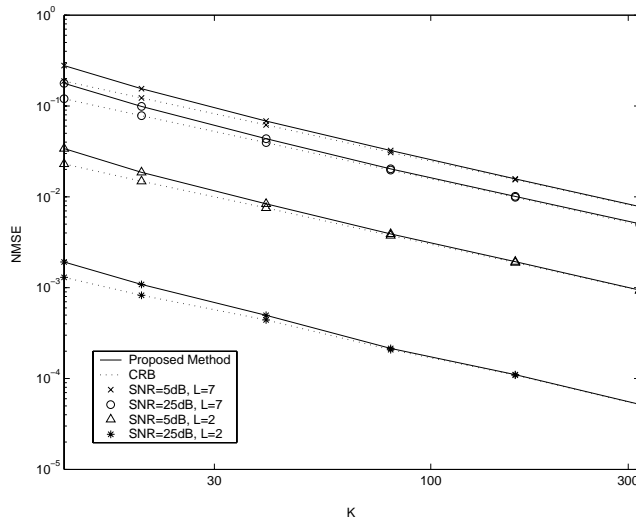


Figure 4.7: Comparison of the simulated NMSE and the CRB vs. K for different channel orders when $N_t = 5$ and constant training sequences are used. The results are plotted for two different values of the SNR, namely 5 and 25dB.

form ML channel estimate in this case. Note that there is no point in using the iterative method in this context, since the closed form channel estimate corresponds to its convergence point, which is confirmed by experimental results (not shown in the figures).

Experimental Comparison

We have seen that there is some difference between the theoretically predicted behavior of the ML estimates and the results that are effectively achieved. Hence, it is worthwhile verifying the conclusions that were drawn from the compared CRB performance (**Fig. 4.1**) of the constant and changing training sequence algorithms still hold when the actual closed form channel estimates are used. The results are presented in **Fig. 4.8** where the first experiment is repeated but the presented curves now represent simulated results rather than the CRB. The experiments are performed for two different values of K . The results show that the previous conclusion of improved performance for the changing training sequence situation still holds except in the specific situation where the channel order is very small and the number of observed blocks, K , is small. In that case, the constant training sequence ML channel estimate outperforms the changing training sequence approximate ML channel estimate for high values of the SNR. However, if we use the iterative method and perform

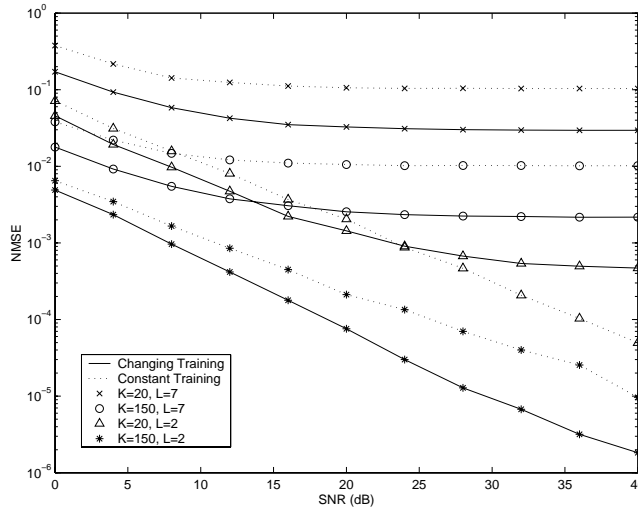


Figure 4.8: Comparison of the simulated NMSE of constant and changing training sequences vs. the SNR for different channel orders and different values of K when $N_t = 5$.

a couple of iterations (not shown in the figure), the advantage of the changing training sequences scheme over the constant training scheme is restored.

4.8.2 Comparison with Existing Methods

We compare the proposed method with two other methods: a classical training-based ML approach and a more advanced semi-blind method based on the EM algorithm presented in [70].

Classical training-based ML channel estimation techniques solely rely on the part of the received symbols that only contain contributions from the known training symbols. They simply discard the received samples that are corrupted by contributions from the unknown data symbols. Such symbols can be observed at the receiver only when $N_t \geq L + 1$. In this case, one can define $\mathbf{T}'_k = \mathbf{T}_k(L + 1 : nt, :)$ and $\mathbf{u}'_k = \mathbf{u}_k(L + 1 : N_t)$. The ML channel model is then known to be the LS fit of $\mathbf{T}'_k \mathbf{h}$ to \mathbf{u}'_k :

$$\mathbf{h}'_{ML} = \left(\sum_{k=1}^K \mathbf{T}'_k{}^H \mathbf{T}'_k \right)^{-1} \sum_{k=1}^K \mathbf{T}'_k \mathbf{u}'_k. \quad (4.45)$$

When constant training sequences are used, the solution becomes $\mathbf{h}'_{ML} =$

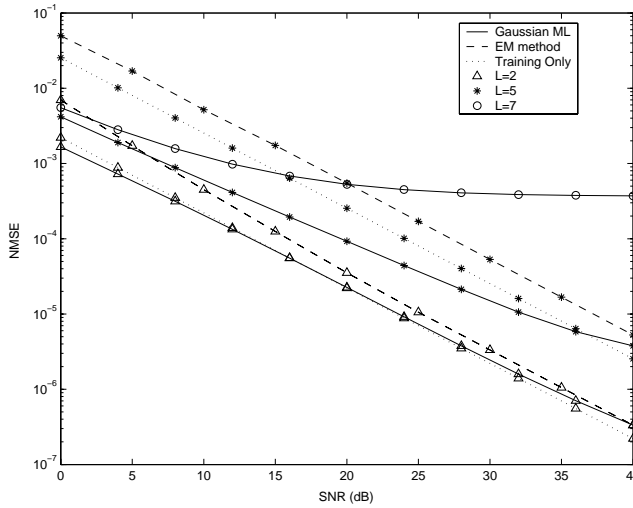


Figure 4.9: Simulated NMSE vs. SNR for the proposed Gaussian ML method and existing channel estimation techniques for different channel orders when changing training sequences of length $N_t = 6$ are used and the number of observed blocks is set to $K = 100$.

$\left(\sum_{k=1}^K \mathbf{T}'_k \mathbf{T}'_k\right)^{-1} \sum_{k=1}^K \mathbf{T}'_k \mathbf{u}'_k$, and only exists when $N_t \geq 2L + 1$. Note that the conditions on the training sequence length N_t are much more stringent than in our proposed method.

We also compare our algorithm with the semi-blind procedure described in [70]. We modify the method in order to estimate the channel relying on the channel output samples contained in the set of \mathbf{u}_k vectors, as exploiting all the channel output samples would yield an intractable complexity. Moreover, as the method cannot be readily adapted to jointly consider all \mathbf{u}_k 's, we compute a channel estimate for each received vector. Considering that the resulting modeling error has a Gaussian distribution, the final ML channel estimate is obtained after computing the mean of the K available channel estimates. Note that, since we consider SISO channels rather than SIMO channels as in [70], the training sequences must have a minimal length of $N_t \geq L + 1$. If that condition is not fulfilled, the method cannot be applied because of rank deficiency problems. Hence, it will not be possible to use this method when the channel order gets too large.

We can now compare the results obtained using these methods with the proposed Gaussian ML estimates. The NMSE of the different methods is presented as a function of the SNR in **Fig. 4.9**. We consider changing training sequences

and compare the results for different channel orders when the length of the training sequences is set to $N_t = 6$.

The proposed Gaussian ML method outperforms existing ones in all situations, except for short channels and at high SNRs where other methods perform slightly better. We know however that increasing the number of observed blocks K or performing a few iterations would restore the advantage of our Gaussian ML method. When the channel order increases, the advantage of the new method increases, especially at low SNR. When the channel order increases and $N_t < L$, the new method still provides reliable channel estimates whilst traditional methods cannot be applied anymore. Finally note that the proposed Gaussian ML method has a complexity that is comparable to the complexity of the classical training-based method, whilst the semi-blind EM-based method has a significantly higher complexity. Indeed, the EM-based method of [70] requires up to 200 iterations to reach convergence, each of these iterations having approximately the same computational complexity as the proposed closed-form GML expressions.

4.9 Channel Estimation in Block Transmission

As they organize the transmitted signals in different ways, the presented block transmission techniques are not equally suited for channel identification. In this section, we present different channel identification techniques that are suited to the different block transmission schemes and discuss the approaches that are used in practical system implementations.

4.9.1 Channel Estimation in CP Transmission

As **OFDM** transmission has been used already in many applications and given the practical importance of accurate channel estimation, a large variety of algorithms and techniques has been proposed in the past years that can be applied for channel identification in both OFDM and SC-CP. We discuss below several approaches that are commonly used and refer to some recently proposed methods that fit in these approaches.

In CP transmission systems, a first property that can be used towards blind or semi-blind channel identification is the CP-induced *cyclostationarity* of the transmitted data, relying on the fact that the cyclostationarity is also present in the received signal for this signal property is not affected by linear filtering (i.e. by the transmission channel). A method proposed by Giannakis et al. [78] directly exploits the cyclostationarity of the transmitted signal. Other methods (e.g. [79] or [80]), use the cyclostationarity of the transmitted signal in order

to blindly identify the channel coefficients from the received symbols and the data model using subspace-based algorithms.

These methods, however, have a quite high computational complexity and the resulting channel models can differ significantly from the true channel since the properties of the transmitted signals are not very well suited for channel identification. Therefore, most existing CP schemes rely on specific training schemes in order to ease the channel estimation task. These training schemes can take various forms (pilot tones or explicit training sequences) but always harm the effective system throughput, as known symbols must be transmitted instead of information symbols to enable the derivation of sufficiently accurate channel estimates

In Hiperlan2 and IEEE802.11a, a preamble appended at the beginning of each transmitted packet contains two pilot OFDM blocks that can be viewed as long *training sequences* that allow channel estimation. These can be exploited to directly estimate the frequency-domain channel matrix \mathbf{H}_{diag} . Let \mathbf{s}_1 and \mathbf{s}_2 be the two OFDM training symbols. A first channel estimate $\hat{\mathbf{H}}_{diag}$ can be obtained by averaging the observed frequency-response over these 2 symbols: $\hat{\mathbf{H}}_{diag}(i, i) = 0.5(z_1[i]/s_1[i] + z_2[i]/s_2[i])$. Though very simple, this method is far from optimal as the noise is averaged over 2 samples only and hence the resulting channel estimate is not very reliable, especially at low SNRs. Moreover, the correlation between the different sub-channel coefficients is not exploited. Several techniques that we do not detail here perform frequency-domain filtering or interpolation on the obtained FD channel estimates in order to exploit the correlation between the subcarriers. The quality of the estimate can be easily improved (see e.g. [81]) if one performs a time-domain estimation rather than a direct frequency-domain identification. Exploiting the fact that the transmission channel is constrained to a (supposedly known) finite length $L + 1$ which is much smaller than the number of subcarriers, it is possible to obtain a more accurate channel estimate:

$$\hat{\mathbf{h}} = \mathbf{X}^\dagger [\mathbf{y}_1^T, \mathbf{y}_2^T]^T \quad (4.46)$$

where \mathbf{X} is a tall $2(N_s + \mu) \times (L + 1)$ Toeplitz matrix with $[\mathbf{x}_1^T, \mathbf{x}_2^T]^T$ in the first column and zeros in the upper triangular part (note that \mathbf{x}_k and \mathbf{y}_k are defined as in *Ch. 2*). The diagonal frequency-domain matrix is obtained by performing a P -point DFT of this time-domain channel estimate: $\hat{\mathbf{H}}_{diag} = \sqrt{P} \text{diag}(\mathcal{F}_P[\hat{\mathbf{h}}^T, 0, \dots, 0]^T)$. This time-domain approach yields much more accurate results than the previous one. *This is the approach that will be used for the experiments presented later in this chapter.* Note that the computational complexity of this channel estimation method is significantly smaller than the computational complexity of the GML method presented above in the context of KSP. Note also that this approach can be used for the **SC-CP** transmission scheme.

On top of these training symbols, *pilot tones* (i.e., tones on which known sym-

bols are modulated rather than unknown data symbols) are often inserted during the data transmission phase. These are traditionally used in order to estimate and compensate carrier phase and frequency offsets. Some recently proposed methods that we do not detail here exploit the knowledge of these pilots in order to estimate the channel. Some solely rely on the knowledge of these pilots in order to produce the channel estimate (see e.g. [82], [83], [84], [85]). Others combine pilot tones and training sequences [86]. In the Hiperlan2 and IEEE 802.11a standards, only 4 pilots are inserted in each OFDM symbol. In this context, these methods will only marginally improve the channel estimate and they are therefore not considered in the experimental comparison we present here.

4.9.2 Channel Estimation in KSP Transmission

In KSP transmission, the unknown symbols of the CP are replaced by short sequences of known symbols (zero or non-zero) that have the same length as the CP. These padded sequences can be viewed as short training sequences, possibly allowing accurate channel identification without requiring the insertion of explicit (long) training sequences. Below, we present some channel identification methods that can be used in the context of KSP transmission, including some recent methods that were specifically designed for the KSP context.

A first method was proposed in [87] for the special case of **ZP-KSP**, where the presence of the zeros of the padded sequence in the data model allows the estimation of the channel using subspace-based algorithms. The proposed method is a blind deterministic one and allows to identify the channel up to a complex scaling factor.

When the padded sequence is not set to zero (**NZP-KSP**), the knowledge of this sequence can be exploited at the receiver to perform a better estimate of the channel. In [7], Leus et al. proposed a method, inspired by [87], that is suited for the case where the padded sequence is non-zero. This method is deterministic and semi-blind. It combines the blind algorithm of [87] together with an additional constraint that comes from the knowledge of the padded symbols. This NZP-KSP method outperforms the proposed ZP-KSP method, even when the uncertainty resulting from the unknown complex scaling factor of the ZP-KSP method is artificially removed. Note that this method requires that the padded sequence remains the same from block to block.

More general channel identification methods can be used as well. We have shown earlier that semi-blind methods are best suited to the KSP context. We have introduced some existing methods and developed a new Gaussian ML semi-blind method that is well suited to this KSP context. This new method has a tractable computational complexity, outperforms existing semi-

blind methods, and asymptotically achieves the CRB and is thus the best suited of the existing semi-blind methods for this specific context.

Note that the CRB study presented in *sec. 4.4* applies to all the channel identification methods that can be considered in the KSP context. This study indicates that the composition of the padded sequences has a major impact on the best achievable accuracy of the channel estimate, independent of the method that is used. This study shows that the use of non-constant padded sequences yields a significant advantage in terms of channel estimation accuracy in comparison to the situation where the same padded sequence is used for all blocks, especially for intermediate channel orders. Hence, the semi-blind methods, which are able to cope with such non-constant padded sequences, are expected to outperform the specific KSP methods [7] and [87], which require the use of constant padded sequences. Experimental results presented in [88] confirm this analysis. *Hence, the method that seems best suited for the KSP context is the GML method presented in this chapter and it is the one we will use in the presented simulation results.*

Note that these requirement on the padded sequences for accurate channel estimation are opposed to the requirements for the use of low-complexity equalizers. Indeed, the computationally less demanding FD equalizers can be used only when the equivalent transmission channel is cyclic, which requires all padded sequences to be equal in the existing KSP scheme.

4.10 Shifted Known Symbol Padding (S-KSP)

We have seen in *sec. 3.4.1* that the use of FD KSP equalizers reduces the computational complexity approximately by a factor 10. The study of the experimental BER performance presented in *sec. 3.3* shows as well that FD and optimal linear KSP equalizers yield similar performance for BPSK and QPSK. For higher order constellations, the FD scheme has a lower but still acceptable performance. The BDFE KSP, which has the highest computational complexity is the best performer in all situations. Hence, a flexible system design should format the transmitted data symbols such that the three equalization schemes are possible. It is then up to the receiver to decide which equalization scheme will be used depending on the required performance, available processing power, channel conditions, and constellation of the transmitted signal. Current KSP schemes require the use of constant padded sequences in order to enable the direct use of FD equalizers.

If one wants to perform accurate channel identification exploiting the knowledge of the padded sequences however, the CRB study presented in *sec. 4.4* indicates that the use of changing training sequences results in significantly more accurate channel estimates. The two desirable goals of flexible receiver

design and accurate channel estimation hence pose contradictory requirements on the composition of the padded sequences.

In this section, we propose a new KSP scheme, namely *Shifted KSP (S-KSP)*, that offers the possibility of describing the transmission system with a circulant channel matrix whilst using non-constant training sequences ($\mathbf{t}_k \neq \mathbf{t}_l, \forall k \neq l$), thereby simultaneously allowing low-complexity frequency domain equalization and accurate KSP-based channel estimation. In contrast, other KSP schemes allowing low-complexity FD equalization require the use of constant training sequences, thereby limiting the accuracy of the resulting channel estimate and hence the final system performance.

In the proposed S-KSP scheme, the requirement on the length of the padded sequences is changed from the classical $N_t \geq L$ into $N_t \geq L + 1$, i.e., the length of the padded sequences is increased by 1. These padded sequences are picked from a unique vector of known symbols $\boldsymbol{\tau}$ of size $(K + N_t - 1) \times 1$ (K is the total number of blocks in a packet). The k^{th} block of known symbols is chosen as:

$$\mathbf{t}_k = \boldsymbol{\tau}(k : k + N_t - 1), \quad (4.47)$$

such that the last $N_t - 1$ symbols of \mathbf{t}_k are equal to the $N_t - 1$ first symbols of \mathbf{t}_{k+1} , i. e.

$$\mathbf{t}_k(2 : N_t) = \mathbf{t}_{k+1}(1 : N_t - 1) = \boldsymbol{\tau}(k + 1 : k + N_t - 1). \quad (4.48)$$

When the unique vector of known symbols $\boldsymbol{\tau}$ is chosen randomly, experimental results show that the channel estimates that are obtained with the GML method presented above have the same accuracy as the channel estimates obtained with purely random padded sequences.

When BDFE or optimal linear KSP equalization is performed at the receiver, the classical receiver structure and equalizer structures can be used without any modification.

If the receiver wants to use the FD equalization scheme, discarding the last sample of each received block will allow to express the effects of the channel with a circulant matrix. To do so, the preprocessing matrix of the receiver is changed into

$$\mathbf{R} = [\mathbf{I}_{N_s+N_t-1} | \mathbf{0}_{(N_s+N_t-1) \times 1}]. \quad (4.49)$$

The S-KSP block transmission scheme is then expressed on a block level as

$$\mathbf{z}_k = \mathbf{H}'_{KSP} \mathbf{s}_k + \mathbf{H}'_{t0} \mathbf{t}_k(1 : N_t - 1) + \mathbf{H}'_{t1} \mathbf{t}_{k-1}(2 : N_t) + \boldsymbol{\eta}_k, \quad (4.50)$$

where \mathbf{H}'_{KSP} , \mathbf{H}'_{t0} and \mathbf{H}'_{t1} are defined as the \mathbf{H}_{KSP} , \mathbf{H}_{t0} and \mathbf{H}_{t1} defined in (2.32), but with modified dimensions ($N_t - 1$ is used instead of N_t in the definitions).

Since the last $N_t - 1$ symbols of \mathbf{t}_{k-1} are equal to the first $N_t - 1$ symbols of \mathbf{t}_k , (4.50) can be rewritten as

$$\mathbf{z}_k = \left[\mathbf{H}'_{KSP} \mid \mathbf{H}'_{t1} + \mathbf{H}'_{t0} \right] \begin{bmatrix} \mathbf{s}_k \\ \mathbf{t}'_k \end{bmatrix} + \boldsymbol{\eta}_k = \mathbf{H}'_{circ} \mathbf{x}'_k + \boldsymbol{\eta}_k, \quad (4.51)$$

where \mathbf{t}'_k and \mathbf{x}'_k are shorthand notations for $\mathbf{t}_k(1 : N_t - 1)$ and $\mathbf{x}_k(1 : N_x - 1)$.

At the cost of an extra redundant symbol, the transmission scheme can again be described with a circulant channel matrix, thereby allowing the use of low-complexity FD equalization schemes, whilst the quality of the channel estimates resulting from non-constant training sequences is preserved.

4.11 System Performance with Estimated Channel Models

In this section, we experimentally analyze the impact of the channel estimation possibilities of the different block transmission schemes on their BER performance. The chosen experimental setup is the same as in *Ch. 3*, i.e. we consider a Hiperlan2 / IEEE 802.11a context with either the standard OFDM or alternatively SC-CP, KSP and S-KSP transmission. In the presented experiments, the receiver first estimates the channel and then computes the equalizers relying on the estimated channel impulse response. The channel estimates are obtained with the method presented in (4.46) for CP-systems and with the closed-form GML channel estimate for KSP systems. The channel estimates for CP systems rely on the two OFDM training symbols of the preamble appended at the beginning of each packet as proposed in the Hiperlan2 and IEEE.811a standards. For KSP systems, no preamble is inserted and the channel estimate relies solely on the knowledge of the padded sequences. The channel estimates are obtained relying on 50 blocks of data symbols. Longer packets of data would improve the quality of the channel estimates in KSP. However, the considered channel estimation method requires that the whole packet is received and stored in a buffer before the channel estimate is computed. We thus chose a relatively small packet size that corresponds to a realistic buffer size of 3600 samples. We consider short padded sequences ($N_t = 8$, which corresponds to the optional short CP length in Hiperlan2, the standard length being $N_t = 16$).

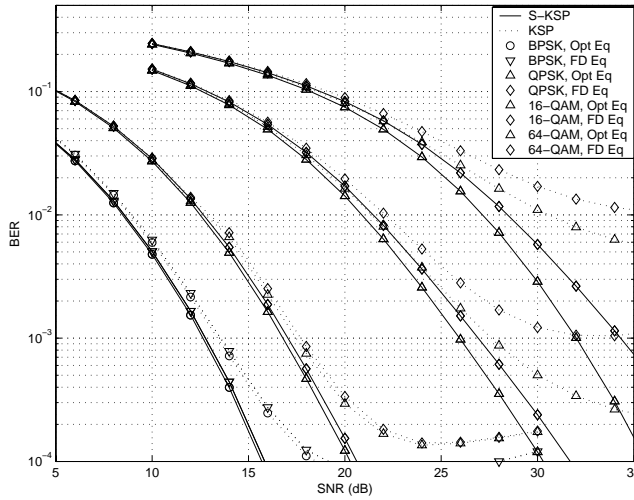


Figure 4.10: BER performance vs. SNR for S-KSP and classical KSP transmission when the equalizers are computed with the channel estimates obtained with the presented GML method. Different constellations and different equalization schemes (FD and optimal linear KSP equalizers) are considered.

The presented results are averaged over 400 random realizations of a Rayleigh fading channel of order $L = 5$ and approx. $2 \cdot 10^5$ bits of data are transmitted over each channel realization. Perfect synchronization and perfect linearity of the amplifiers are considered in the presented experiments.

In a *first experiment*, we compare the proposed S-KSP scheme to classical KSP transmission. The considered classical KSP scheme allows the receiver to perform low-complexity FD equalization, i.e. the padded sequence remains constant from block to block. The channel is estimated relying on the knowledge of the padded sequences only for both schemes. The simulation results are presented in **Fig. 4.10**, where the uncoded BER of the two schemes are compared for different constellations considering both optimal and FD equalization schemes. The use of constant padded sequences in classical KSP transmission yields channel estimates of relatively poor quality. This results in the appearance of a floor in the BER performance of the classical KSP scheme. Thanks to the more accurate channel estimates, S-KSP clearly outperforms the classical KSP transmission scheme. Finally note that the low-complexity FD S-KSP equalizers outperform the more complex optimal linear equalizers of classical KSP transmission.

This first experiment highlights the advantage yielded by the use of the S-KSP scheme. Hence, we only consider S-KSP in the experiments presented below

where KSP is compared with OFDM.

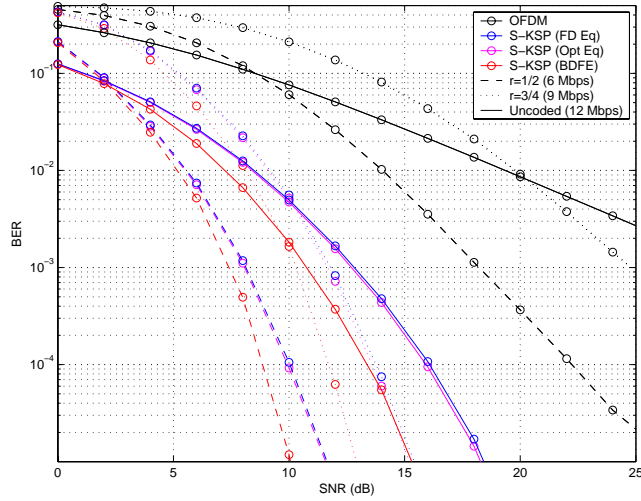
In a *second experiment*, we compare the proposed S-KSP scheme with OFDM. The results are presented in **Fig. 4.11** and **Fig. 4.12**. In the context of this experiment, the channel modeling error is small when the proposed method is used for the S-KSP channel estimation and the impact of the channel modeling error on the final BER performance of the optimal linear KSP equalization schemes is hardly noticeable. The channel modeling error has an impact on the BER performance of the FD S-KSP equalizer, but this impact is limited. In contrast, the poorer quality of the available channel estimates in OFDM transmission harms the BER performance significantly and the final performance remains several dBs away from the perfect channel knowledge situation. The advantage of S-KSP transmission, which already appeared clearly in the perfect channel knowledge case, is thus increased when realistic channel estimates are considered. Finally note that the performance of the FD S-KSP equalizer is comparable to that of the optimal linear S-KSP equalizer in most cases, except when 64-QAM constellations are used. The system performance is further improved at the cost of an increased equalization complexity when the BDFE is used in this S-KSP context.

In **Fig. 4.13**, we summarize the performance of the different block transmission techniques when realistic channel estimates are considered. The achievable datarate is presented as a function of the SNR for a fixed target BER for the different techniques. These results show that S-KSP outperforms OFDM by several dBs for both coded and uncoded transmission. These results also show the interest of using uncoded S-KSP transmission when the SNR is not too low so that high datarates can be delivered.

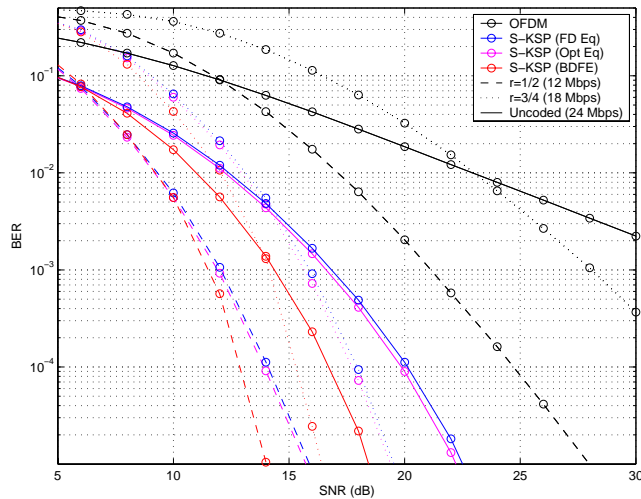
4.12 Conclusions

In this chapter, we have first presented a new Gaussian ML channel identification method where the training sequences can be shorter than the channel impulse response length. The main advantage of the proposed method is that it exploits all the channel output samples that contain contributions from the known symbols, including those that are corrupted by the unknown surrounding data symbols. We analyzed two situations: the situation where the same training sequence is repeated at the end of each data block (constant training sequence case) and the situation where this training sequence is changed at the end of each data block (changing training sequence case). We first proposed an iterative ML method and then derived approximate closed form expressions for the ML channel estimates.

The proposed closed-form expressions have a low complexity and effectively achieve the CRB in most practical situations. In the few situations where the

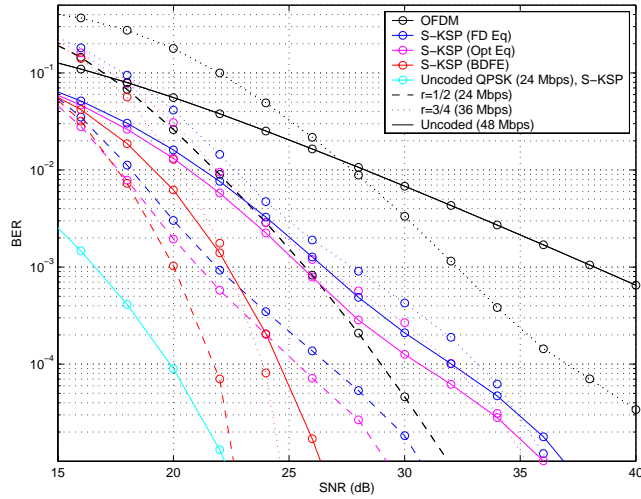


(a)

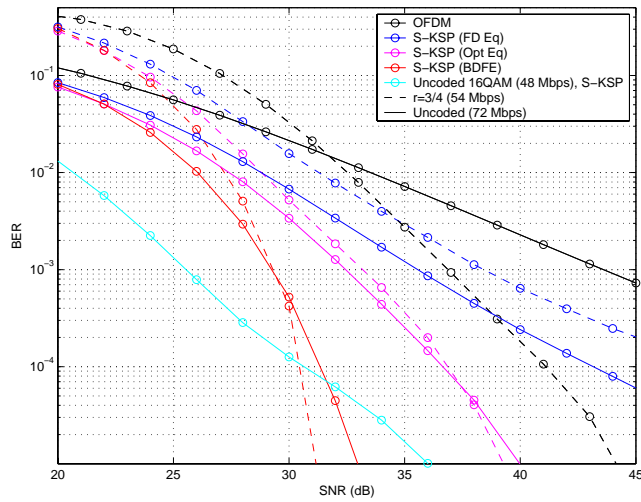


(b)

Figure 4.11: BER performance vs. SNR for different coding rates (1/2, 3/4, Uncoded) and different equalization schemes (BDFE, optimal and FD linear equalizers) in BPSK (a) and QPSK (b) modulation. The equalizers are computed relying on the channel estimates obtained with the proposed GML method.

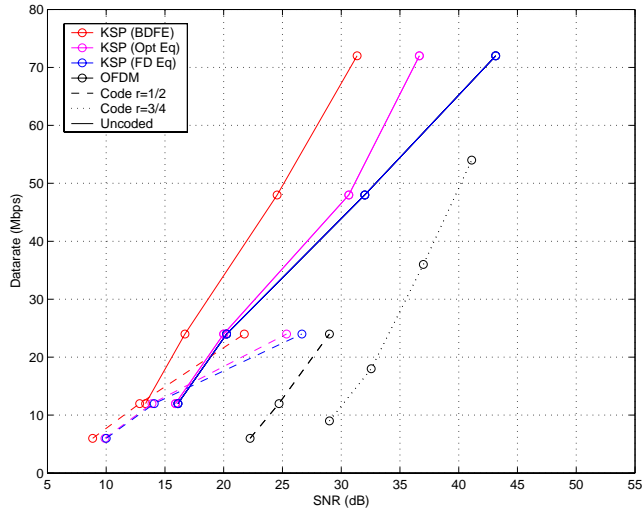


(a)

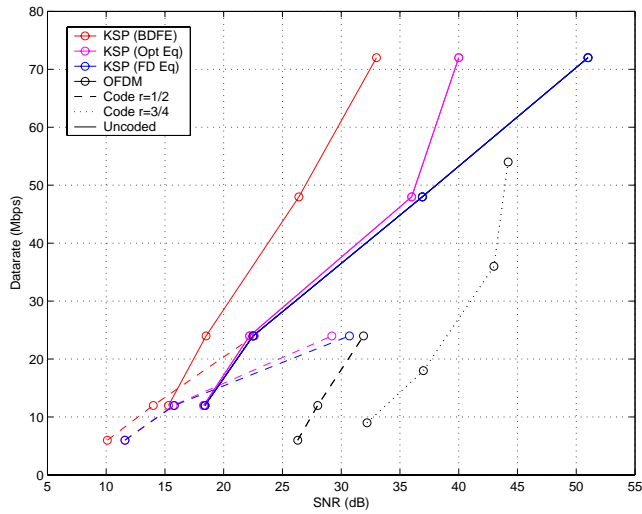


(b)

Figure 4.12: BER performance vs. SNR for different coding rates (1/2, 3/4, Uncoded) and different equalization schemes (BDFE, optimal and FD linear equalizers) in 16-QAM (a) and 64-QAM (b) modulation. The equalizers are computed relying on the channel estimates obtained with the proposed GML method.



(a)



(b)

Figure 4.13: Achievable data rates (Mbits/s) as a function of the SNR for target BER of 10^{-4} (a) and 10^{-5} (b) for different transmission schemes. Each curve is obtained with a fixed equalization scheme and a fixed coding rate but with different constellations. The equalizers are computed relying on the channel estimates obtained with the proposed GML method.

CRB is not achieved (i.e. low channel order, large SNR and small number of training sequences), the iterative method can be used and will achieve the CRB in a few iterations. The proposed method can be used in white as well as in colored noise conditions and clearly outperforms existing semi-blind methods and classical training-based methods that solely rely on the channel output samples that only depend on the known symbols. Finally, amongst existing training-based and semi-blind methods, the proposed channel identification method currently appears to be the best suited candidate to achieve near-CRB channel estimates in this KSP context and this with the advantage of a low computational complexity.

The CRB analysis proposed in this chapter shows that there is a significant advantage in terms of achievable accuracy of the channel estimate when the padded sequences are changed after each block. We introduced in this chapter a new KSP method, S-KSP, that jointly allows to estimate the channel accurately and to perform low-complexity FD equalization. Experimental results show the superiority of S-KSP over classical KSP in terms of BER performance when the channel is estimated relying on the knowledge of the padded sequences.

Finally, we performed an experimental comparison between the different block transmission techniques in a Hiperlan2 / IEEE 802.11a context relying on the channel estimates that can be obtained in the considered transmission schemes. This experiment shows that the advantage of S-KSP over OFDM that was already apparent in perfect channel knowledge conditions (see *Ch. 3*) is further increased in this case thanks to the improved accuracy of the channel estimates obtained in the S-KSP context.

Appendix A

In this appendix, we show that (see *sec. 4.5*)

$$\mathbf{Q}_{ML}(\mathbf{h}) = K^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}_k \mathbf{h}) (\mathbf{u}_k - \mathbf{T}_k \mathbf{h})^H. \quad (4.52)$$

The proof is adapted from [72, pp.201-202]. Let us first define the sample covariance matrix:

$$\mathbf{R}_K(\mathbf{h}) \triangleq K^{-1} \sum_{k=1}^K (\mathbf{u}_k - \mathbf{T}_k \mathbf{h}) (\mathbf{u}_k - \mathbf{T}_k \mathbf{h})^H.$$

Using this definition, the log-likelihood function (4.5) can be re-expressed as

$$-\mathcal{L} = \frac{K}{2} (\text{tr}(\mathbf{R}_K(\mathbf{h})\mathbf{Q}^{-1}) + \ln|\mathbf{Q}|) + \text{cst} \quad (4.53)$$

The proposed ML estimate defined in (4.52) is equivalent to $\mathbf{Q}_{ML}(\mathbf{h}) = \mathbf{R}_K(\mathbf{h})$. Claiming that (4.53) is minimized with respect to \mathbf{Q} for $\mathbf{Q} = \mathbf{R}_K(\mathbf{h})$ is equivalent to claiming that

$$\text{tr}(\mathbf{R}\mathbf{Q}^{-1}) + \ln|\mathbf{Q}| \geq \text{tr}(\mathbf{R}\mathbf{R}^{-1}) + \ln|\mathbf{R}|, \quad \forall \mathbf{Q} = \mathbf{Q}^H > 0.$$

where \mathbf{R} is a shorthand notation for $\mathbf{R}_K(\mathbf{h})$. The following equivalences can be derived:

$$\begin{aligned} & \text{tr}(\mathbf{R}\mathbf{Q}^{-1}) + \ln|\mathbf{Q}| \geq \text{tr}(\mathbf{I}_{N_t+L}) + \ln|\mathbf{R}| \\ \Leftrightarrow & \text{tr}(\mathbf{R}\mathbf{Q}^{-1}) + \ln(|\mathbf{Q}|/|\mathbf{R}|) \geq N_t + L \\ \Leftrightarrow & \text{tr}(\mathbf{R}\mathbf{Q}^{-1}) - \ln|\mathbf{R}\mathbf{Q}^{-1}| \geq N_t + L. \end{aligned} \quad (4.54)$$

It is clear from its definition that \mathbf{R} can be factorized in a full rank square matrix and its complex conjugate transpose: $\mathbf{R} = \mathbf{G}\mathbf{G}^H$. Define next the matrix $\mathbf{F} \triangleq \mathbf{G}^H\mathbf{Q}^{-1}\mathbf{G}$. This matrix \mathbf{F} is symmetric and positive definite. Its eigenvalues $\lambda_1 \dots \lambda_{N_t+L}$ clearly satisfy $\lambda_i > 0$. Using these new definitions, we can proceed and rephrase (4.54) as:

$$\begin{aligned} & \text{tr}(\mathbf{G}\mathbf{G}^H\mathbf{Q}^{-1}) - \ln|\mathbf{G}\mathbf{G}^H\mathbf{Q}^{-1}| \geq N_t + L \\ \Leftrightarrow & \text{tr}(\mathbf{G}^H\mathbf{Q}^{-1}\mathbf{G}) - \ln|\mathbf{G}^H\mathbf{Q}^{-1}\mathbf{G}| \geq N_t + L \\ \Leftrightarrow & \text{tr}(\mathbf{F}) - \ln|\mathbf{F}| \geq N_t + L \\ \Leftrightarrow & \sum_{i=1}^{N_t+L} \lambda_i - \ln \prod_{i=1}^{N_t+L} \lambda_i \geq N_t + L \end{aligned}$$

$$\Leftrightarrow \sum_{i=1}^{N_i+L} \lambda_i - \ln \lambda_i - 1 \geq 0$$

Since $\forall \lambda > 0$, $\ln \lambda \leq \lambda - 1$, the last inequality holds. This concludes the proof.

Appendix B

In this appendix, we show that (see *sec. 4.6.1*)

$$\left| \mathbf{I} + \hat{\mathbf{Q}}^{-1} (\mathbf{Th} - \bar{\mathbf{u}}) (\mathbf{Th} - \bar{\mathbf{u}})^H \right| = 1 + (\mathbf{Th} - \bar{\mathbf{u}})^H \hat{\mathbf{Q}}^{-1} (\mathbf{Th} - \bar{\mathbf{u}}).$$

Let \mathbf{x} be a $n \times 1$ column vector and \mathbf{y} a row vector of the same length and define the rank-1 matrix $\mathbf{A} = \mathbf{xy}$ of size $n \times n$.

Since \mathbf{A} has rank 1, it has a single non-zero eigenvalue, λ_1 . Moreover, its null space has dimension $n - 1$, and the characteristic polynomial is divisible by t^{n-1} . Since λ_1 is a root of the characteristic polynomial, we can express it as $t^n - \lambda_1 t^{n-1}$.

Observing that $\mathbf{Ax} = (\mathbf{xy})\mathbf{x} = \mathbf{x}(\mathbf{yx})$, we know that \mathbf{x} is an eigenvector of \mathbf{A} with eigenvalue $\lambda_1 = \mathbf{yx}$. The characteristic polynomial can thus be developed into:

$$|t\mathbf{I} - \mathbf{A}| = t^n - \mathbf{yx}t^{n-1}. \quad (4.55)$$

Taking $t = -1$ yields

$$\begin{aligned} |-\mathbf{I} - \mathbf{xy}| &= (-1)^n (1 + \mathbf{yx}) \\ \Leftrightarrow (-1)^n |\mathbf{I} + \mathbf{xy}| &= (-1)^n (1 + \mathbf{yx}) \\ \Leftrightarrow |\mathbf{I} + \mathbf{xy}| &= 1 + \mathbf{yx}. \end{aligned} \quad (4.56)$$

The sought property results from the following choices for \mathbf{x} and \mathbf{y} : $\mathbf{x} = \hat{\mathbf{Q}}^{-1} (\mathbf{Th} - \bar{\mathbf{u}})$ and $\mathbf{y} = (\mathbf{Th} - \bar{\mathbf{u}})^H$.

Let us first perform an eigenvalue decomposition of the following product:

$$\hat{\mathbf{Q}}^{-1} (\mathbf{Th} - \bar{\mathbf{u}}) (\mathbf{Th} - \bar{\mathbf{u}})^H = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H, \quad (4.57)$$

where \mathbf{U} is a unitary matrix and $\mathbf{\Sigma}$ is a diagonal matrix with the eigenvalues λ_i in decreasing order on the main diagonal.

Because of the term $(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H$, (4.57) has only rank one and there is a single non-zero eigenvalue. Since (4.57) is positive definite, this only eigenvalue is positive, hence it is λ_1 . Since the trace of a matrix is the sum of the eigenvalues of this matrix, λ_1 is equal to the trace of the initial product (4.57), or equivalently:

$$\lambda_1 = (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H \hat{\mathbf{Q}}^{-1} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}}). \quad (4.58)$$

It follows that:

$$\left| \mathbf{I} + \hat{\mathbf{Q}}^{-1} (\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})(\mathbf{T}\mathbf{h} - \bar{\mathbf{u}})^H \right| = |\mathbf{I} + \boldsymbol{\Sigma}| = 1 + \lambda_1. \quad (4.59)$$

This concludes the proof.

Chapter 5

Direct Symbol Estimation

After having discussed the approach where the receiver estimates the channel and constructs an equalizer relying on the available channel estimate in order to reconstruct the transmitted sequence of data symbols, we consider in this chapter the case where the receiver does not rely on an explicit channel model in order to recover the transmitted data. This approach is simply an alternative to the classical approach (channel ID followed by equalizer design) in NZP-KSP or in CP transmission when training or pilots are inserted in the stream of transmitted data. In ZP-KSP or in CP transmission without training or pilots however, the receiver cannot estimate the transmission channel accurately with the channel estimation techniques described earlier in this text and therefore, the direct symbol estimation methods discussed in this chapter are the only possible approaches for the estimation of the transmitted data.

The direct approaches that will be investigated here aim at maximizing the likelihood of the transmitted sequence exploiting the finite alphabet properties of the transmitted signals. These techniques are known as Maximum Likelihood (ML) techniques. Their main advantage is their performance; when optimal ML decoding is done, the lowest achievable error probability for the estimation of the transmitted sequence is obtained. Their major impediment though is their computational complexity since most of them use exhaustive or trellis searching techniques in order to estimate the transmitted sequence. Sub-optimal techniques aim at approaching the ML detector with a reduced computational complexity.

In this chapter, we propose a new sub-optimal blind ML technique for the direct estimation of data symbols transmitted over an unknown convolutive channel that is suited for the KSP and SC-CP transmission schemes.

The problem of ML estimation of data transmitted over an unknown channel

and some recently proposed techniques are briefly presented in *sec. 5.1*. In *sec. 5.2*, we present the data model that will be used and the ML detector in the context of SC-CP and KSP. In *sec. 5.3*, we propose an ILSP algorithm suited for this context whose performance approaches the ML detector. In *sec. 5.4*, we describe an alternative to the proposed method that is suited to the context of OFDM transmission, namely [89]. In *sec. 5.5*, we present more detailed simulation results, and we finally draw some conclusions in *sec. 5.6*.

5.1 Blind Maximum Likelihood Sequence Estimation Techniques

The problem of blindly estimating a sequence that was transmitted over a convolutive channel \mathbf{h} of order L is given as follows. Let \mathbf{x} be the transmitted sequence and \mathbf{y} the received sequence. The goal is to find among all possible combinations of \mathbf{x} and \mathbf{h} the pair $\hat{\mathbf{x}}$ and $\hat{\mathbf{h}}$ that yields the highest probability of observing the received sequence \mathbf{y} :

$$\max_{\hat{\mathbf{x}}, \hat{\mathbf{h}}} p(\mathbf{y} | \hat{\mathbf{x}}, \hat{\mathbf{h}}) \quad (5.1)$$

This maximization problem has to be solved under the constraint that the elements of the transmitted sequence belong to a finite alphabet, while the channel coefficients are unconstrained. When the noise at the receiver is AWGN, this is equivalent to a joint Least Squares (LS) minimization problem:

$$\min_{\hat{\mathbf{x}}, \hat{\mathbf{h}}} \|\mathbf{y} - \hat{\mathbf{x}} \otimes \hat{\mathbf{h}}\|^2. \quad (5.2)$$

Since this problem is separable in its two variables $\hat{\mathbf{x}}$ and $\hat{\mathbf{h}}$, it can be solved in two steps. In a first step we compute the LS channel estimate for every possible transmitted sequence \mathbf{x} . In a second step, the value of the cost function (5.2) is computed for all possible candidates. The joint ML sequence and channel estimator chooses the pair \mathbf{x}, \mathbf{h} with the smallest cost function.

This exhaustive search procedure is optimal but computationally prohibitive, its cost being exponential in the length of the transmitted sequence. Cheaper optimal decoding techniques have been proposed. A first family of solutions proposes iterative algorithms that rely upon an initial channel estimate. They find the ML transmitted sequence under the hypothesis that the true channel is equal to the available channel estimate. The channel is then re-estimated based on the decoded sequence. These algorithms proceed iteratively until convergence. The different proposed solutions differ in the way they estimate the ML sequence. The K-means Algorithm [90] uses the Viterbi algorithm to

decode the sequence, while the EM algorithm [91] computes the likelihoods of all possible data sequences conditioned on the available channel estimate.

In [92], non-iterative blind methods are proposed. A first proposed optimal technique performs decoding using Viterbi-related trellis search techniques on several hypothesized trellises. The computational cost of this optimal technique remains prohibitive enough to discount its use in practical systems. A sub-optimal blind trellis searched technique whose complexity is comparable to adaptive Viterbi decoding with exact channel knowledge was proposed in [92] as well. Here, instead of retaining a single survivor into each state as is done in classical Viterbi, the M best paths into a state are retained. The channel model is then updated for each of the retained paths as in adaptive Viterbi. Each survivor path is thus coupled to a channel estimate. The channel model associated to the path with the lowest metric converges to the true channel model whatever the initial channel estimate. Here again the computational cost, which is exponential in the channel order, remains problematic especially in broadband communications where the channels have long impulse responses.

When no ISI is present in the transmission system, the problem of blindly estimating the transmitted sequence clearly becomes much less complex thanks to the absence of memory in the channel. In [74], iterative methods exploiting finite alphabet properties of the transmitted signals have been proposed in a MIMO context for such systems. One of these proposed methods, known as Iterative Least Squares with Projection (ILSP), approaches Maximum Likelihood (ML) detection of the transmitted sequence whilst maintaining the complexity at an acceptably low level. When the transmitted signal is received through an antenna array, it is possible to identify it blindly using subspace-based approaches that exploit the spatial structure of the antenna array (see [93] for the ISI-free case and [94], [95] or [96] for OFDM based extensions of this last method to frequency-selective channels)

In this chapter, we present a new sub-optimal blind MLSE (Maximum Likelihood Sequence Estimation) in the context of ZP KSP transmission over multipath channels whilst maintaining the computational complexity at an affordable level. The presented method is readily extended to the SC-CP transmission scheme. When ZP KSP is used, the channel matrix is circulant, which creates flat-fading channels in the frequency-domain. This property allows us to use simpler MLSE methods that are designed for flat-fading channels, but these methods have to be applied in the frequency-domain instead of in the time-domain as is usually done.

To approach MLSE, we propose a novel ILSP algorithm that exploits both the cyclic prefixed structure of the transmitted signals and their finite alphabet properties. The idea of combining block transmission techniques with MLSE techniques designed for flat fading-channels is not new. It has already been used in [89], where a modified ILSP procedure for OFDM systems is proposed.

This method reconstructs the frequency-domain channel matrix from a time-domain estimate of the channel with a limited number of taps. In the algorithm proposed here, this limited channel order constraint on the frequency-domain channel matrix is relaxed. Experimental results show that this difference is crucial for the performance of the algorithm, allowing our method to perform much better in terms of bit error rates (BER) than the existing one. The reasons for this improved performance are discussed in more details in *sec. 5.5*.

5.2 Data Model and Maximum Likelihood Sequence Estimation

In this section, we introduce the proposed method in the specific context of ZP-KSP. However, the presented data model as well as the proposed method are applicable to SC-CP, S-KSP or more general KSP transmission schemes with minor modifications. The data model of ZP-KSP transmission presented in *Ch. 2* is expressed on a per block basis. We adapt it here in order to take the whole packet of data symbols into account. Define the matrix of transmitted data symbols as:

$$\mathbf{X} = \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_K \\ \mathbf{0}_{N_t \times 1} & \cdots & \mathbf{0}_{N_t \times 1} \end{bmatrix}, \quad (5.3)$$

and the matrix of channel outputs as

$$\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_K]. \quad (5.4)$$

The matrix \mathbf{N} of the AWGN is defined similarly. Since the padded sequence is constant in this ZP context, the transmission scheme can be expressed with a circulant channel matrix:

$$\mathbf{Y} = \mathbf{H}_{circ} \mathbf{X} + \mathbf{N}, \quad (5.5)$$

or, equivalently,

$$\mathbf{Y} = \mathcal{I}_P \mathbf{H}_{diag} \mathcal{F}_P \mathbf{X} + \mathbf{N}, \quad (5.6)$$

where (\mathcal{I}_P) \mathcal{F}_P is the P -point (I)DFT matrix ($P = N_s + N_t$) and \mathbf{H}_{diag} is a diagonal matrix with the frequency-domain description of the channel on the main diagonal.

If we want to focus on the unknown data symbols, the data model is expressed as

$$\mathbf{Y} = \mathcal{I}_P \mathbf{H}_{diag} \mathcal{F}'_P \mathbf{S} + \mathbf{N}, \quad (5.7)$$

where $\mathbf{S} = \mathbf{X}(1 : N_s, :)$ and \mathcal{F}'_P is the partial DFT matrix defined as $\mathcal{F}'_P = \mathcal{F}_P(:, 1 : N_s)$

The data model shows that the received signals can be modeled as *deterministic sequences corrupted by AWGN*. The log likelihood function of the transmitted data is thus:

$$\mathcal{L} = -\alpha - \beta \ln(\sigma^2) - \frac{1}{\sigma^2} \sum_{k=1}^K \|\mathbf{y}_k - \mathcal{I}_P \mathbf{H}_{diag} \mathcal{F}'_P \mathbf{s}_k\|^2, \quad (5.8)$$

where α and β are constants. The ML detector maximizes \mathcal{L} with respect to the unknown parameters \mathbf{H}_{diag} and \mathbf{s}_k , $n = 1 \cdots K$ under the following constraint set:

Constraint Set 1 (CS1):

- The elements of \mathbf{S} belong to a finite alphabet,
- $\mathbf{H}_{diag} = \mathcal{F}_P \mathbf{H}_{circ} \mathcal{I}_P$ where \mathbf{H}_{circ} is a circulant matrix built from any unconstrained L^{th} order channel \mathbf{h} .

The ML problem is equivalent to the following minimization problem:

Cost Function 1 (CF1):

$$\min_{\mathbf{H}_{diag}, \mathbf{S}} \|\mathbf{Y} - \mathcal{I}_P \mathbf{H}_{diag} \mathcal{F}'_P \mathbf{S}\|^2, \quad (5.9)$$

under the same constraints. Multiplying both terms by \mathcal{F}_P and using the notations $\mathbf{S}_f = \mathcal{F}'_P \mathbf{S}$ and $\mathbf{Y}_f = \mathcal{F}_P \mathbf{Y}$ yields the following alternative formulation of the Maximum Likelihood problem:

Cost Function 2 (CF2):

$$\min_{\mathbf{H}_{diag}, \mathbf{S}_f} \|\mathbf{Y}_f - \mathbf{H}_{diag} \mathbf{S}_f\|^2, \quad (5.10)$$

with the same constraint on \mathbf{H}_{diag} , and where \mathbf{S}_f is constrained by the finite alphabet structure of \mathbf{S} . This problem can be solved using exhaustive search procedures as described in *sec. 5.1*, the cost of this method being totally prohibitive.

We know that the constraint on \mathbf{H}_{diag} gives it the structure of a diagonal matrix with $L + 1$ degrees of freedom. In the sequel, we investigate what happens if we give \mathbf{H}_{diag} P degrees of freedom, that is if we allow all elements of its diagonal to vary freely¹. The constraint set becomes:

Constraint Set 2 (CS2):

- The elements of \mathbf{s} belong to a finite alphabet,
- \mathbf{H}_{diag} is a diagonal matrix with unconstrained elements on the diagonal.

The MLSE problem in our context is thus equivalent to:

- solving a minimization problem equivalently expressed by CF1 or CF2
- under the constraint set CS1 or under the constraint set CS2

We focus on the solution of this problem under the constraint set CS2. In the noiseless case, the solution to this modified problem (CS2) and the initial ML problem (CS1) are the same. When the received sequence is noisy, we cannot guarantee that the derived solution will be equal to the ML solution under CS1. However, we still find the solution of an ML-like problem, but with a modified constraint set (CS2). We therefore propose to classify this method as a modified MLSE method.

Since the problem is separable in its two variables, the optimization can be carried out in two steps [74], using the two equivalent formulations (5.9) and (5.10). We first solve (5.10) with respect to \mathbf{H}_{diag} . Exploiting the new condition on \mathbf{H}_{diag} (CS2), we force it to have the desired diagonal structure by splitting the problem into P independent sub-problems:

$$\min_{\mathbf{H}_{diag}(j,j)} \|\mathbf{Y}_f(j, :) - \mathbf{H}_{diag}(j, j)\mathbf{S}_f(j, :)\|^2, \quad j = 1 \cdots P, \quad (5.11)$$

the solution of which is

$$\hat{\mathbf{H}}_{diag}(j, j) = \mathbf{Y}_f(j, :)\mathbf{S}_f(j, :)^H (\mathbf{S}_f(j, :)\mathbf{S}_f(j, :)^H)^{-1}, \quad (5.12)$$

Note that this solution would not be achievable under the initial constraint on \mathbf{H}_{diag} (CS1), all the sub-problems then being linked to one another. We thus see that using CS2 allows us to reduce the complexity of the problem. Inserting these results that only depend on \mathbf{S}_f and \mathbf{Y}_f into (5.9) allows us to find the optimum sequence by enumerating over all the possible values for the

¹This choice of allowing full freedom to the elements of the diagonal of \mathbf{H}_{diag} is further discussed in *sec. 5.5*.

matrix \mathbf{S} , computing the value of the cost function, and choosing the matrix that minimizes this cost function. This includes the computation of $\hat{\mathbf{H}}_{diag}$ (and thus \mathbf{S}_f) for every possible combination of the inputs. The computational cost of this method, which is exponential in P , K , and the alphabet size, again severely limits its practical interest.

5.3 Iterative Least Squares with Projection

In the previous section, we derived an expression for a modified ML detector. The associated minimization problem is separable in its continuous and discrete variables. In this section, we apply an Iterative Least Squares with Projection algorithm inspired by [74], that uses this separation property to approach the solution of the modified ML problem by iteratively minimizing the cost functions (5.9) and (5.10) for one variable and then for the other.

Assume we have an initial channel estimate $\hat{\mathbf{H}}_{diag}$. We first minimize the cost function (5.9) with respect to \mathbf{S} with this fixed $\hat{\mathbf{H}}_{diag}$:

$$\min_{\mathbf{S}} \|\mathbf{Y} - \mathcal{I}_P \hat{\mathbf{H}}_{diag} \mathcal{F}'_P \mathbf{S}\|^2. \quad (5.13)$$

This is equivalent to computing a soft estimate of the transmitted symbols, implicitly performing a classical frequency-domain equalization on the received symbols:

$$\hat{\mathbf{S}} = \mathcal{I}'_P \hat{\mathbf{H}}_{diag}^{-1} \mathcal{F}_P \mathbf{Y}, \quad (5.14)$$

where $\mathcal{I}'_P = \mathcal{I}_P(1 : N_s, :)$. The finite alphabet constraint is then exploited by making hard decisions on $\hat{\mathbf{S}}$, yielding $\check{\mathbf{S}}$

We then use this estimate to minimize the cost function (5.10) with respect to \mathbf{H}_{diag} , where $\check{\mathbf{S}}_f = \mathcal{F}'_P \check{\mathbf{S}}$ is derived from the previous step:

$$\min_{\mathbf{H}_{diag}} \|\mathbf{Y}_f - \mathbf{H}_{diag} \check{\mathbf{S}}_f\|^2. \quad (5.15)$$

Again we split this into P parallel independent problems to force \mathbf{H}_{diag} to have a diagonal structure:

$$\min_{\mathbf{H}_{diag}(j,j)} \|\mathbf{Y}_f(j, :) - \mathbf{H}_{diag}(j, j) \check{\mathbf{S}}_f(j, :)\|^2, \quad j = 1 \cdots P, \quad (5.16)$$

the solution of which is

$$\hat{\mathbf{H}}_{diag}(j, j) = \mathbf{Y}_f(j, :) \check{\mathbf{S}}_f(j, :)^H (\check{\mathbf{S}}_f(j, :) \check{\mathbf{S}}_f(j, :)^H)^{-1}. \quad (5.17)$$

Since \mathbf{H}_{diag}^{-1} is used in the iterations to estimate \mathbf{S} , we can alternatively (but not equivalently) estimate it directly by solving the following minimization

problem:

$$\min_{\mathbf{H}_{diag}^{-1}(j,j)} \|\mathbf{H}_{diag}^{-1}(j,j)\mathbf{Y}_f(j,:) - \check{\mathbf{S}}_f(j,:)\|^2, \quad j = 1 \cdots P, \quad (5.18)$$

this is a simple least squares problem whose solution is:

$$\hat{\mathbf{H}}_{diag}^{-1}(j,j) = \check{\mathbf{S}}_f(j,:)\mathbf{Y}_f(j,:)^H (\mathbf{Y}_f(j,:)\mathbf{Y}_f(j,:)^H)^{-1}. \quad (5.19)$$

Note that this modification implies that we perform an implicit MMSE equalizer estimate rather than a direct channel estimate, which also has the benefit of avoiding noise enhancement when the next \mathbf{S} estimate is computed. This \mathbf{H}_{diag}^{-1} is used to compute a new $\check{\mathbf{S}}$, and we then proceed iteratively. The iterations are stopped when two consecutive finite alphabet estimates of the transmitted sequences are identical. The procedure converges to a minimum of the cost function associated to the modified ML problem and can be initialized with a random channel estimate. If this initial estimate is too distant from the true ML channel model, the convergence point can be an irrelevant local minimum of the cost function. In that case, we restart the iterations with another channel estimate that is sufficiently distant from the first estimate. Experimental results show that the iterative procedure must be restarted 2 or 3 times at most before the global minimum is reached.

An important advantage of this ILSP method is its low computational complexity. At every iteration, one symbol estimation step and one channel estimation step are performed. The complexity of the symbol estimation step is $\mathcal{O}(KP \log(P))$. The complexity of the channel estimation step is $\mathcal{O}(KP \log(P))$ including the computation of \mathbf{Y}_f from \mathbf{Y} . The total complexity of the proposed method is thus $\mathcal{O}(\log(P))$ per symbol as well. Unlike existing methods, the complexity of the proposed method is independent of the channel order and grows only linearly with the length of the transmitted sequence.

The method presented above was specifically introduced in the framework of ZP-KSP transmission. It is however straightforwardly adapted to the other single-carrier block transmission techniques.

In the *NZP-KSP context*, the method can be easily adapted when the channel model can be expressed using a circulant channel matrix, i.e. in S-KSP or with constant padded sequences. In that case, only a few modifications are needed to take into account the presence of the padded sequences, i.e. the zeros at the bottom of the matrix \mathbf{X} are replaced by the appropriate padded sequences; the matrices \mathbf{S} and \mathbf{S}_f that are used in the iterative procedure are replaced by the equivalent matrices \mathbf{X} and \mathbf{X}_f that take the padded sequences into account. Finally, the partial (I)DFT matrices \mathcal{F}'_P (\mathcal{I}'_P) are replaced by the full (square) (I)DFT matrices. Modifying the data model in that way allows to use the method as outlined above. The presence of the known symbols improves the convergence properties of the iterative methods and experimental results

show that the method systematically converges to the global minimum in this configuration without having to restart the iterative procedure.

In the SC-CP context, the data model is slightly changed as well in order to enable the use of the proposed ILSP method, i.e. no zeros are appended at the end of the matrix \mathbf{X} and the partial (I)DFT matrices \mathcal{F}_P (\mathcal{I}_P) are replaced by the full (I)DFT matrices that are now of size $P = N_s$. The properties of the iterative method in the SC-CP context do not differ from the ZP-KSP situation and experimental results show that the final performance is the same.

5.4 ILSP for OFDM Systems

As outlined above, the ILSP method presented in this chapter can be easily used for all the single-carrier block transmission techniques. The only block transmission technique to which it cannot be directly tailored is OFDM. In OFDM indeed, using the iterative method with the constraint CS2 would be equivalent to performing a direct FD equalization with the channel model available at the initialization step of the procedure, making the whole iterative performance pointless. We detail below an existing iterative procedure that is suited to the OFDM context.

In [89], an ILSP procedure for joint blind estimation of channel and data symbols in OFDM systems is proposed. Like in our method, the data are organized in blocks of N_s symbols padded with N_t zeros with the condition $N_t \geq L$. An IDFT is then computed on the data blocks (of size $N_t + N_s$) before transmission and no CP is added. This transmission scheme makes it possible to avoid Inter Block Interference by removing N_t samples at the receiver. The system can be described with a matrix formulation similar to (5.5) where \mathbf{H}_{circ} is replaced by some appropriate matrix \mathbf{A} . This channel matrix \mathbf{A} is directly derived from the time-domain model of the channel \mathbf{h} . The proposed detection algorithm starts with an arbitrary estimate of the channel matrix \mathbf{A} . Based on this channel estimate, a soft estimate of the transmitted sequence is computed, followed by a hard decision. Based on this sequence estimate, the channel matrix \mathbf{A} is re-estimated. One could decide to start the next iteration with this new channel matrix estimate. However, the fact that the channel matrix is derived from an L^{th} order channel model imposes a constraint on the structure of \mathbf{A} . To fulfill this constraint, [89] seeks the L^{th} order channel model that is the best LS fit to the estimated channel matrix. A new channel matrix is then computed based on this estimate of \mathbf{h} . The iterations proceed with this last channel matrix estimate.

Note that this iterative method has a quite high complexity as the modified OFDM scheme that is used does not rely on the traditional low-complexity OFDM equalizers, but on specific equalizers whose complexity is similar to the

optimal linear KSP equalization scheme.

5.5 Simulation Results

Introductory Experiment In order to get a first insight into the performance of the proposed method, we performed simulations comparing the proposed ILSP method with the existing OFDM one and with a modified version of our algorithm that uses the initial constraint set CS1 rather than the modified one, CS2. The constraint set CS1 can be enforced either by directly computing the time-domain channel model \mathbf{h} , or by first estimating \mathbf{H}_{diag} as usually and then find the channel model \mathbf{h} that best fits \mathbf{H}_{diag} in the LS sense. We then use the FD description $\hat{\mathbf{H}}_{diag}$ to equalize the received symbols and find the next candidate input sequence.

The results are presented in **Fig. 5.1**, where the achieved BERs are shown for different SNRs. The experiments were performed on a set of 200 different realizations of a first order Rayleigh fading channel using BPSK signaling. We ran the three proposed methods on the same data sequences and noise realizations. The experiments were performed with a packet length $K = 100$ and 64 sub-carriers. The upper line is the BER obtained with the existing OFDM method. The second curve is obtained under the constraint set CS1. The lower curve is obtained with the proposed ILSP method where \mathbf{H}_{diag} is only constrained to a diagonal matrix structure (CS2).

The modified method (with CS1) only slightly outperforms the existing ILSP method for OFDM and does not show any diversity advantage. Knowing the potential performance advantage of KSP transmission over OFDM in this uncoded context, this is quite a poor achievement.

In contrast, when the constraint on \mathbf{H}_{diag} is relaxed, only forcing it to have a diagonal structure results in much better BER performance. The blind OFDM scheme is now largely outperformed and the diversity advantage of KSP transmission is restored.

As mentioned earlier, an important difference between this OFDM method and the one we propose resides in the constraints that are put on the channel matrix. In the OFDM method, the \mathbf{A} channel matrix is forced to be derived from an L^{th} order channel model, allowing only $L + 1$ degrees of freedom to the channel matrix. If this constraint is relaxed, \mathbf{A} has N_s^2 degrees of freedom and the convergence properties of the iterative algorithm become erratic because the problem is under-constrained. In the modified method that uses CS1, we put an equivalent constraint on \mathbf{H}_{diag} by forcing it to be the frequency-domain description of a general L^{th} order channel, allowing $L + 1$ degrees of freedom to the channel matrix. However, in order to reduce the computational complexity

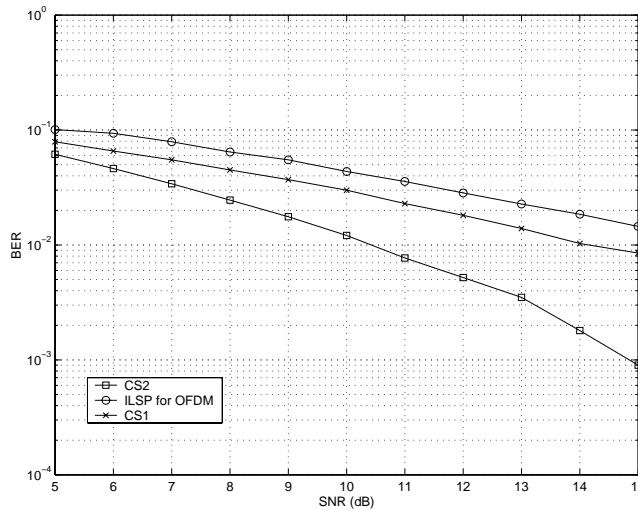


Figure 5.1: BER vs. SNR for the OFDM ILSP method (upper line), our method under the traditional constraint set (CS1, middle line) and our method with the modified constraint set (CS2, lower line).

of the algorithm, we have relaxed this constraint on \mathbf{H}_{diag} , replacing it by the constraint of having a diagonal structure. This increases the number of degrees of freedom on \mathbf{H}_{diag} to P instead of the $L+1$ offered by the previous constraint.

Experimental results show that relaxing this L^{th} order time-domain equivalent channel constraint on the channel matrix yields spectacular improvements in terms of BER. We give next an intuitive explanation for this improved performance.

When blind MLSE is performed by exhaustive search (see *sec. 5.1*), the cost function (5.2), (5.9) or (5.10) is minimized for every possible input sequence. Existing iterative methods avoid this computationally prohibitive exhaustive search by only considering a few selected candidates among all the possible sequences. These iterative procedure all use the same framework: the iterations are started with an initial sequence candidate; the LS channel model associated with this sequence candidate is then computed (this LS channel model is also the ML channel model). The next step consists in selecting a new sequence candidate under the hypothesis that the provided channel model is exact. Finally, the best proposed sequence is selected (the one with the lowest cost function). The technique that is used to select new candidate sequences is crucial for the performance of the iterative methods. Only the candidates selected by this technique will be considered as possible estimates for the transmitted sequence. For instance, the K-means Algorithm [90] performs ML detection

using the Viterbi algorithm under the hypothesis that the provided channel model is exact in order to select new candidates. The selection procedure used in the CS1 method implicitly performs a zero-forcing FD equalization of the estimated channel. The best possible candidate is obtained performing such a ZF equalization with the real channel model. However, in our KSP context, ZF-FD equalizers are known to have a poor performance for channels exhibiting a high frequency selectivity. This poor selection method severely limits the performance that could possibly be achieved by such methods. The proposed method behaves a bit differently. As we let the iterations run, \mathbf{H}_{diag} converges to the frequency-domain diagonal matrix that will best project the received sequence onto the known finite alphabet of the transmitted sequence. The \mathbf{H}_{diag} that is computed at each step is thus a sort of frequency-domain best fit between the hypothesized input and the received noisy channel output, instead of the ML channel model for the corresponding input sequence that is usually computed. The noise effects are implicitly included in the computed \mathbf{H}_{diag} since its estimate is based on noisy data observations. We end up with a frequency-domain MMSE equalizer that generates much better candidate sequences than the ZF equalizer that is obtained when \mathbf{H}_{diag} is constrained according to CS1. *Using CS2 rather than CS1 thus not only allows to reduce the complexity of the method, it also allows to find better candidate sequences and thereby increases the performance of the algorithm.*

Convergence:

Using the constraint set CS2 makes the problem under-constrained, the convergence properties of the algorithm get much worse than what they are when CS1 is used. Indeed, when CS1 is used, the method converges in 2 or 3 iterations independently of the initial channel estimate and the resulting channel model is close to the true channel. In contrast, when CS2 is used many more iterations and initial channel guesses are required to reach the optimum. However, when a fairly good initial channel estimate is available, the convergence properties are improved and the method converges rapidly to its minimum. Several approaches are possible to obtain this initial channel estimate. First, if NKP-KSP transmission is considered, the channel estimation method presented in *Ch. 4* can be used yielding a very fast convergence given the high accuracy of the initial channel estimate. In the ZP-KSP or SC-CP situation the simple channel identification method proposed earlier cannot be used anymore and a first solution would be to initialize the iterations with a blind channel estimation method, with the drawback of an increased complexity. An interesting alternative is to run a couple of iterations under the fast converging constraint set CS1 and use the resulting channel estimate as starting point for the iterations under the constraint set CS2. This last approach is the one we used in the simulations and it yields satisfactory results, at least for low-order constellations.

Regarding the convergence properties, it is interesting to note that the ZP-

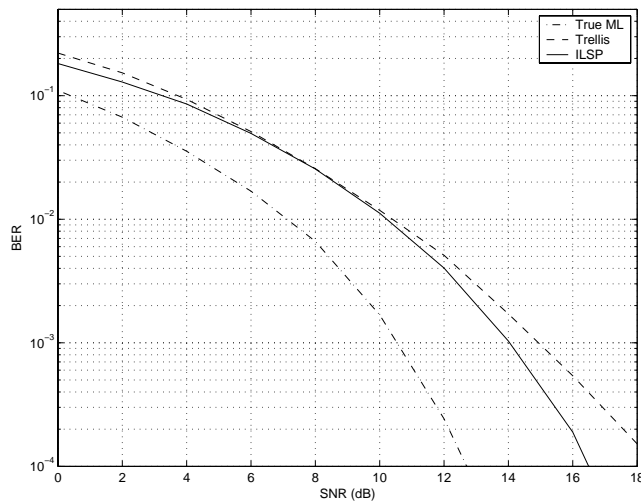


Figure 5.2: BER vs. SNR for the trellis searched method (full line) and our iterative method (dashed line).

KSP and SC-CP have similar convergence speeds, whereas NZP-KSP converges much faster as the knowledge of the padded sequences improves the quality of the channel estimates. Another interesting aspect is that the performance of the convergence point is the same for these three block transmission techniques

Comparison with the Blind Trellis Searched Method In the previous section, we have seen how relaxing the constraint on \mathbf{H}_{diag} allows our method to reach high performance. In this section, we investigate how the resulting iterative method compares with existing sub-optimal blind MLSE methods. We chose to compare our method with a trellis searched blind equalization method that was first proposed in [92], which approaches blind ML decoding whilst keeping the computational complexity at a reasonable level. This method was already mentioned in *sec. 5.1*. The M best paths into a state are retained instead of one as done in the classical Viterbi algorithm and a channel estimate is associated to each of these metrics. A feature of interest which is offered by the known padded sequence is that it forces the trellis initialization and termination. At the beginning and the end of a transmitted block, the channel memory is filled with the padded zeros. This allows us to perform a block per block detection without having to retain the trellis metrics and surviving paths between different blocks.

The results are presented in **Fig. 5.2**, where the BER achieved for different SNRs are shown. The experiments are performed on a set of 200 different

realizations of a third order Rayleigh fading channel using BPSK. We ran the different methods on the same data sequences and noise realizations. The experiments were performed with a packet length $K = 100$ and 64 sub-carriers. The proposed ILSP scheme slightly outperforms the trellis method, especially in the high SNR region. Moreover, the proposed method requires much less computing power than the trellis searched one. As the channel order increases, the complexity of our method remains the same whilst the complexity of the trellis method grows exponentially. Finally note that both methods remain at a respectable distance from the performance of the true ML solution, i.e. the performance that is obtained when the true channel model is used for Viterbi decoding.

Extended Simulation Results In this last section, we investigate the performance of the proposed method for higher order channels. We also study the impact of the packet length K both in terms of performance and complexity. We present simulation results for Rayleigh-fading channels with $L = 5$, $N_t = 5$ and $P = 64$ using BPSK symbols. We used 400 randomly generated Rayleigh-fading channels, and approximately estimated 10^7 data symbols for each point of the graph. Note that the results of the iterative method are always obtained in a ZP-KSP with 2 iterations performed with CS1 for the initialization.

Fig. 5.3 shows the BER for the proposed method as a function of the SNR. The different curves show the results for different values of the packet length K . The lower curve shows the theoretical BER floor for ML decoding of BPSK signals that are transmitted over an L^{th} order Rayleigh-fading SISO channels [97, p 955]. The upper curve shows the BER obtained with a classical ZF FD KSP equalizer in the frequency-domain assuming perfect channel knowledge.

The performance of the method increases with K , which is probably because the curve of the cost function gets smoother as the number of data symbols increases. For large values of K , the BER curve follows the true ML with a 3dB loss in SNR.

Fig. 5.4 shows the average number of iterations that were required for the algorithm to converge. This number is maintained at a quite low level thanks to the good initial channel estimator. This is especially true for high SNRs. It should be noted that large K 's, which yield better BER performance, require more iterations, especially for low SNRs.

Fig. 5.6 shows the evolution of the BER performance as the number of iterations increases. The upper curve shows the performance after one iteration, etc. The experiments are performed on 3rd order channels with a packet length K set to 100. We see from these experiments that most of the gain is achieved after 4 iterations. Further iterations still improve the BER but only in a marginal way. This shows thus that a very small number of iterations is required in order

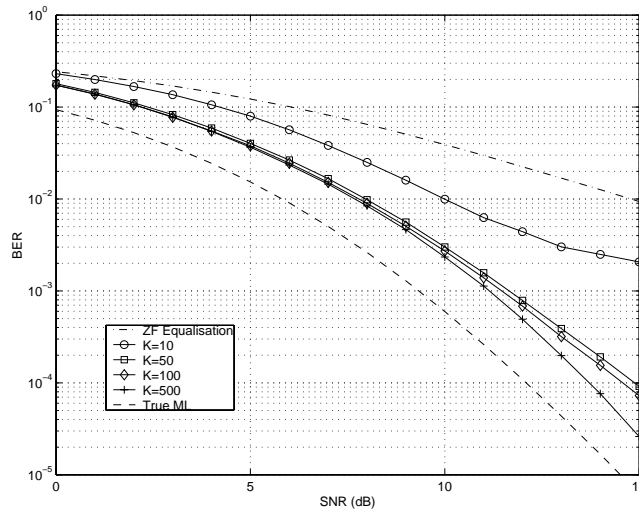


Figure 5.3: BER vs. SNR for different Packet Lengths K .

to approach closely the optimal BER for this method.

Finally, we compare in **Fig. 5.5** the proposed ILSP scheme to different MMSE KSP equalizers obtained with channel estimates obtained with the closed-form GML channel estimation technique presented in *Ch. 4* (note that S-KSP is considered for the accuracy of the channel estimates). The true ML performance is again indicated as a benchmark. The performance of the proposed ILSP method is near that of the optimal linear KSP equalizer. This shows that the indirect approach (channel estimation followed by an equalization step) yields more accurate results in NZP-KSP. However, that indirect approach cannot be applied in ZP-KSP or SC-CP schemes without explicit training. In that case, the presented ILSP method is an interesting alternative to estimate the transmitted data without having to rely on an explicit data model. It is also interesting to note that the performance of the BDFE is pretty close to the ML.

5.6 Conclusions

The method presented in this chapter proposes a cheap and efficient solution for direct sequence estimation in a KSP or SC-CP context. It performs almost as well as the optimal linear MMSE KSP equalizer, at least in low-order constellations, even when it is not possible to obtain a reliable channel model (in the absence of training or in ZP-KSP). It also outperforms existing blind trellis

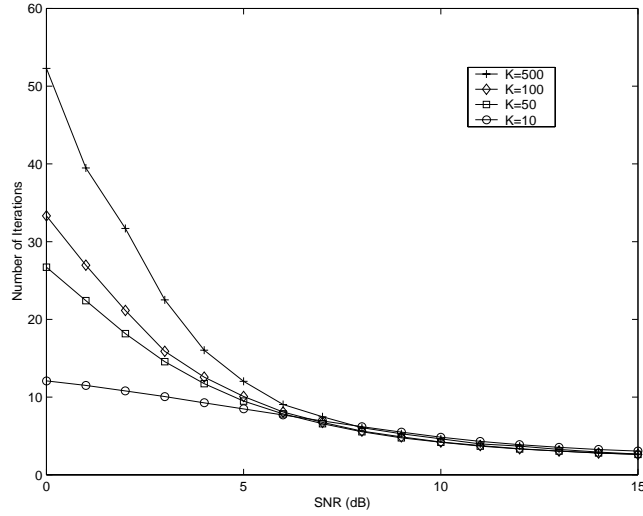


Figure 5.4: Num. of It. vs. SNR for different Packet Lengths K .

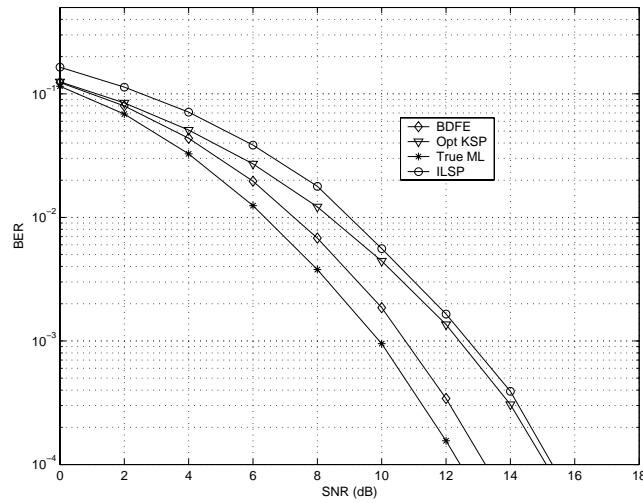


Figure 5.5: BER vs. SNR for the proposed method, different KSP equalizers and MLSE.

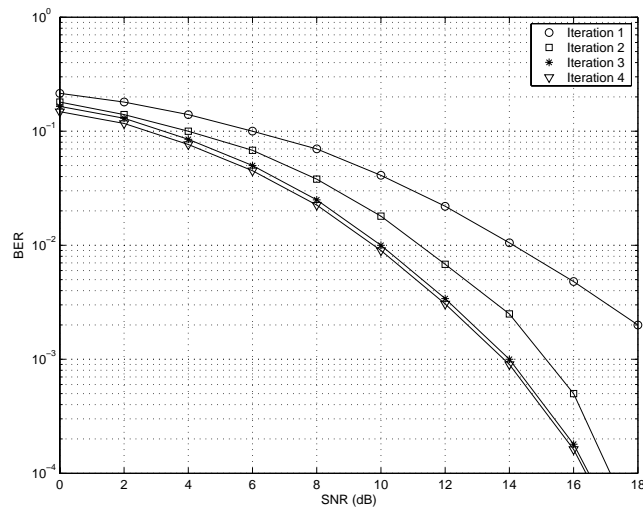


Figure 5.6: BER vs. SNR for increasing numbers of iterations.

searched techniques in the same context. The main advantage of this method is its low computational complexity that is independent of the channel order.

Chapter 6

KSP in Doubly-Selective Channels

Until now, we have restricted our attention to broadband channels that have a relatively large coherence time. This framework corresponds to applications where a high data rate is delivered to users that are fixed or have a low mobility, such as WLAN or fixed wireless access applications. Wireless networks of the next generation will aim at delivering similar data rates to mobile users possibly experiencing high mobility. The delivery of the desired data rates will require the use of broadband communication channels. Besides the frequency-selectivity of these channels, these systems will have to cope with the high user mobility that causes the transmission channel to change rapidly in time. Such channels experiencing both frequency and time selectivity are termed *doubly selective* channels.

The block transmission techniques that have been studied in previous chapters have been shown to cope efficiently with the frequency selectivity of the channels as long as these remain stationary. Moreover, this stationarity eases the channel identification process and enables the derivation of accurate channel models. The goal of this chapter is to analyze how the presented block transmission techniques can be adapted when the mobility is increased and the channel becomes time-selective and to analyze whether these schemes are appropriate candidates to cope efficiently with the doubly selective nature of the channels.

The first consequence of the fast channel variations is the performance degradation that occurs in most block transmission schemes when the mobility is high, and the coherence time of the channel becomes smaller than the block length. In that case, all the presented frequency-domain equalization schemes

(i.e. OFDM, SC-CP and FD KSP equalizers) suffer from *Inter Carrier Interference (ICI)*. Indeed, as the channel changes during the block transmission, the channel matrix is not circulant anymore and the diagonalization property of the channel matrices does not hold. Hence the FFT operations performed by these FD equalizers does not result in a frequency-domain interference-free model and ICI occurs, which severely limits the performance of these equalization schemes. The situation is somewhat different when the BDFE or optimal linear KSP equalizers are used. Indeed, these do not rely on the diagonalization of circulant channel matrices and, if the full time-varying channel model is known, it is possible to slightly modify existing equalizers to maintain the performance level whilst the computational complexity of the equalizers remains unchanged.

Another consequence of the time-selectivity is the increased difficulty of the channel estimation process. In low mobility conditions, it is reasonable to assume a constant channel that can be identified by appending training sequences at the beginning of the packet of data symbols (e.g. Hiperlan2 or IEEE802.11a), or by jointly considering all the known symbols appended by the KSP scheme (see e. g. *Ch. 4*), which allows to obtain quite accurate channel estimates. When the channel varies in time, it is a more challenging task to track its evolution. The identification scheme must be able to regularly update the channel estimates with a sufficient accuracy whilst only a limited number of channel output samples can be exploited to compute these estimates. Moreover, the placement of the pilot symbols is important to enable an accurate channel estimation. It has been proved in [13] and [98] that the optimal placement in order to track the channel variations consists in equi-spaced training symbol sequences of constant length, which matches the KSP transmission scheme that has been described in previous chapters.

For these two reasons (limited performance degradation of the classical block equalizers and optimal training placement), KSP transmission seems to be a natural candidate for data transmission over doubly selective channels. In contrast, the presented CP techniques do not appear to be very well suited for such channels. Indeed, the use of low-complexity FD equalizers, which are their main advantage in stationary channels, is not possible in doubly selective channels and alternative equalization schemes of a higher complexity have to be developed in this context. Besides this, classical CP systems are not well suited for the identification of the time-varying channel and the introduction of numerous pilot tones that harm the effective data throughput seems to be necessary to track the channel evolution with a decent accuracy. Hence, it seems that KSP, which already outperformed other block transmission techniques in the context of stationary channels, appears to be the best suited block transmission scheme in the context of this chapter. Therefore, we limit our investigations to the KSP transmission scheme.

In this chapter, we adapt the KSP data model to the context of doubly selec-

tive channels and show how classical KSP equalizers can be modified to suit such channels. Furthermore, we investigate how the knowledge of the padded sequences can be optimally exploited in order to accurately estimate doubly selective transmission channels. We propose three new methods towards that goal. These channel identification methods are the main contributions proposed in this chapter.

Note that the experimental investigations performed in this chapter focus on WLAN-like setup with increased velocity. Given the relatively small size of the transmitted block in this setup, we end up in moderate time-selectivity scenarios. Indeed, the variations of the channel are essentially felt during the transmission of a whole packet and the main challenge consists in tracking the variation of this channel across the block of data. The time variation of the channel also causes some ICI. However, as the channel variation inside one block of data is limited in the considered scenarios, the impact of the ICI, though visible, is limited. Alternative setups with larger block sizes (DVB standards for instance) could be considered for the study of transmission schemes with a strong ICI.

The *first channel estimation technique* is directly inspired by the Gaussian Maximum Likelihood channel estimation technique in the context of KSP transmission over stationary frequency-selective channels that we presented in *Ch. 4*. We propose here an adaptive version (exponentially weighted RLS scheme) of this method, which makes it possible to track the time variations of the transmission channel with an acceptable accuracy.

The two other techniques rely on a recently proposed approach for modeling doubly selective channels, namely the Basis Expansion Model (BEM) [99], [100] and [101]. This new model has attracted a lot of attention recently for it allows for an accurate representation of doubly selective channels with a limited number of complex exponentials and allows for cheap and efficient channel equalization schemes [102], [103] [100]. The problem of identifying the BEM parameters of the transmission channel through training has already been discussed in [98] and [104]. These methods only exploit the channel output samples that solely contain contributions from the training symbols, neglecting the channel output samples containing mixed contributions from the training symbols and the unknown surrounding data symbols. It has been proved in [98] that, when these estimation methods are used, the optimal training sequences for a fixed training power consist of $2Q + 1$ equispaced packets of $2L + 1$ pilot symbols with a single non-zero element placed in the middle ($2Q + 1$ represents the number of complex exponentials used in the BEM and L is the channel order). However, it is not clear whether or not these training sequences are optimal if all the channel output samples containing contributions from the training symbols are considered for channel estimation. Note as well that when the channel order increases, inserting training sequences of length $2L + 1$ might harm the available bandwidth if these sequences need to be repeated on a regular ba-

sis. Moreover, concentrating the whole training energy in one single symbol results in large peaks in the transmitted signal, which results in an important PAPR problem, leading to problems that are similar to those encountered in OFDM transmission (non-linearities, higher quantization noise, ...). If the implementational constraint is not the total training power but the power per transmitted symbol, shorter training sequences with a constant modulus (or at least reduced PAPR) are a better choice than the training scheme proposed in [98]. It is therefore interesting to develop alternative channel identification methods that are able to optimally exploit training sequences that differ from the proposed ones. This is what we do in this chapter where we propose two new BEM channel estimation techniques that take into account all the channel output samples containing contributions from both the training symbols and the unknown surrounding data symbols. Another important problem of existing methods is that they assume that the period of the BEM is equal to the interval over which the channel is identified, which generally leads to large modeling errors at the edges of the interval. The proposed methods are able to cope with arbitrary periods for the BEM model (the period of the BEM can be longer than the modeling interval), which significantly reduces these BEM-induced modeling errors.

The *first BEM method* relies on the channel estimates produced by the proposed adaptive method and finds the set of BEM coefficients that fits best (in a least squares sense) to the adaptive solution. The *second BEM method* directly identifies the BEM coefficients and outperforms the other methods. This last method appears to be the most efficient of the three. We compare the proposed methods with a modified version of the existing methods [98] and [104]: we adapt them to enable arbitrarily long periods for the BEM model, which leads to reduced modeling errors and significantly improved performance for these methods. When the $2L + 1$ -long training sequences containing a single non-zero element are used, the proposed method for the direct identification of the BEM parameters has the same performance as the modified version of [98] and [104]. For all other training sequences, the proposed methods significantly outperform existing ones.

Note that this chapter only reviews a small part of the methods and techniques that have been proposed to cope with doubly selective channels. It is not our ambition here to cover exhaustively this research area and we have restricted our field of investigations to some recently proposed techniques that are well suited to the block transmission techniques, which are the main investigation subject of this text. However, for the reader who is interested in a deeper exploration of the literature on time varying channels, we have included references to some interesting papers proposing alternative approaches than those that are studied here.

The structure of this chapter is as follows. In *sec. 6.1*, we present the different data models that enable us to cope efficiently with the doubly-selective

channels. We present the channel models in *sec. 6.1.1*, and adapt the KSP data models and equalizers in *sec. 6.1.2* and *sec. 6.1.3*. The proposed channel estimation techniques are presented in *sec. 6.2* and the resulting system performances are analyzed in *sec. 6.3*.

6.1 Data Model

In this section, we first focus on the description of doubly selective channels. We present a general data model for data communications over such channels and, relying on the physical description of the transmission channel, we then show how realistic channels can be simulated within the presented transmission model. Subsequently, we introduce the Basis Expansion Model (BEM) that is used to model doubly selective channels with a limited number of parameters. We then derive the data model of a KSP transmission scheme and adapt classical KSP equalizers to the situation of doubly selective channels.

6.1.1 Doubly Selective Channel Model

Time-Varying Channels

The baseband model 2.7 that we presented in *Ch. 1* needs to be modified if we are to cope with time-varying channels. In the time-selective context, the channel can still be accurately modeled by a FIR filter, but the coefficients of the filter become now time-dependent. We thus propose the following model to describe the transmission of data symbols over a doubly selective channel. Let $x[n]$ be the sequence of transmitted data symbols. Sampling the receive antenna at the symbol rate, the sequence of received samples $y[n]$ can without loss of generality be described by:

$$y[n] = \sum_{\nu=-\infty}^{+\infty} h[n; \nu] x[n - \nu] + \eta[n], \quad (6.1)$$

where $h[n; \nu]$ accounts for the effects of the transmission channel and the transmit and receive filters ($h[n; \nu]$ is thus the complex multiplicative channel coefficient that accounts for the contribution of the $(n - \nu)^{th}$ transmitted data symbol into the n^{th} received sample), and $\eta[n]$ is the additive noise, that we will consider to be white and Gaussian distributed as previously.

Physical Channel Model

In practical situations, the channel parameters do not vary randomly as they are linked to the physical properties of the transmission channel. The proposed physical channel model allows us to parameterize the channel coefficients as a function of the physical transmission channel. Consider a multipath propagation channel where C clusters, each consisting in R reflected or scattered rays, arrive at the receiver. Considering that the transmission interval is short enough such that the number of rays and clusters does not change during the transmission, and the time-variation of the channel is negligible during the time-span of the receive filter, the transmission channel can be described as:

$$h[n; \nu] = \sum_{c=1}^C \psi(\nu T_s - \tau_c) \sum_{r=1}^R G_{c,r} e^{j2\pi f_{c,r} n T_s}, \quad (6.2)$$

where T_s is the symbol period and $\psi(t)$ is the total impulse response of the transmit and receive filters, τ_c is the delay of the c^{th} cluster, $G_{c,r}$ and $f_{c,r}$ are respectively the complex gain and the frequency offset of the r^{th} ray of the c^{th} cluster. The frequency offset is caused by the relative motion between the receiver and the scatterer and is the source of the time-variation of the channel coefficients. Jakes' model [105], which is often proposed to simulate time-varying transmission channels, is a special case of the presented physical channel model.

The *Doppler spread* f_{max} of the channel is the maximum absolute value of all these frequency offsets. This Doppler spread is usually determined by the speed of the mobile user terminal. If the transmitted symbols are modulated on a carrier of frequency f_c and v is the speed of the mobile terminal, the Doppler spread can generally be obtained as $f_{max} = v f_c / c$, where c is the speed of light.

The coherence time of the transmission channel defined as $\tau_{coh} = 1/(2\pi f_{max})$ is a rough measure of the time during which the channel coefficients remain relatively correlated. When $NT_s \geq \tau_{coh}$, N being the total number of symbols in the transmitted packet, the channel coefficients undergo significant changes during the transmission and the channel is labeled as time-varying, which is the situation we will consider here. A more refined measure of the channel variations was introduced by Jakes [105] and is called the *envelope correlation coefficient*:

$$\rho(\Delta_t) = J_0^2(2\pi f_{max} \Delta_t), \quad (6.3)$$

where J_0 is the 0^{th} order Bessel function of the first kind. It is a measure of the correlation between two channel realizations separated by a time interval of Δ_t seconds.

Similarly to what happens in the stationary channel model, the finite character of the physical channel's delay spread combined with the limited time span of

the transmit and receive filters results in channel orders of (constant) limited order L , i.e. $h[n; \nu] = 0$, when $\nu < 0$ or $\nu > L$.

The transmission scheme (6.1) can then be re-expressed as:

$$y[n] = \sum_{\nu=0}^L h[n; \nu] x[n - \nu] + \eta[n]. \quad (6.4)$$

The physical channel model presented here, though very handy for simulating realistic time-varying transmission channels, still contains many parameters which makes it impractical to use for channel estimation/equalization applications. In the rest of the text, we adopt the notation \mathbf{H} defined as $\mathbf{H}(i, j) = h[i; j - 1]$, where the channel coefficients $h[i; j - 1]$ are derived from (6.2) to characterize the evolution of the channel during the transmission of \mathbf{x} .

Basis Expansion Model (BEM)

Exploiting the limited Doppler spread of physical channels, the Basis Expansion Model (BEM), which has been proposed recently (see e.g. [99], [100] or [101]), models each tap of the time-varying channel with a limited number of complex exponentials. This approach allows us to represent the channel with a reduced number of parameters, namely the multiplicative coefficients of the complex exponentials, referred to as the BEM parameters. The true channel $h[n; \nu]$ is approximated over the transmission interval ($n = 1, \dots, N$) by the BEM as:

$$h[n; \nu] \approx \sum_{l=0}^L \delta[\nu - l] \sum_{q=-Q}^Q h_{q,l} e^{j2\pi qn/N_{mod}}. \quad (6.5)$$

Each channel tap is modeled as the sum of $2Q+1$ complex exponentials and the whole channel is described with a limited number of $(2Q+1)(L+1)$ parameters, namely the $h_{q,l}$ coefficients. The parameters Q and N_{mod} should be selected carefully in order to allow for an accurate approximation of the true channel over the packet interval. The Doppler spread of the channel's BEM (which is equal to its highest frequency component) is equal to $Q/(N_{mod}T_s)$. Hence, Q and N_{mod} should be chosen such that the Doppler spread of the BEM is approximately equal to the Doppler spread of the true channel. Furthermore, the BEM is periodic with a period N_{mod} . Therefore, as the true channel is generally not periodic, N_{mod} should at least be as large as N ; the match of the BEM to the true channel over the packet interval gets tighter as N_{mod} increases. However, increasing N_{mod} forces us to increase Q in order to fulfill the Doppler spread requirement. The optimal choice of these parameters Q and N_{mod} when the final goal is the identification of the transmission channel is a trade-off between two antinomic effects: the modeling error and the identification error.

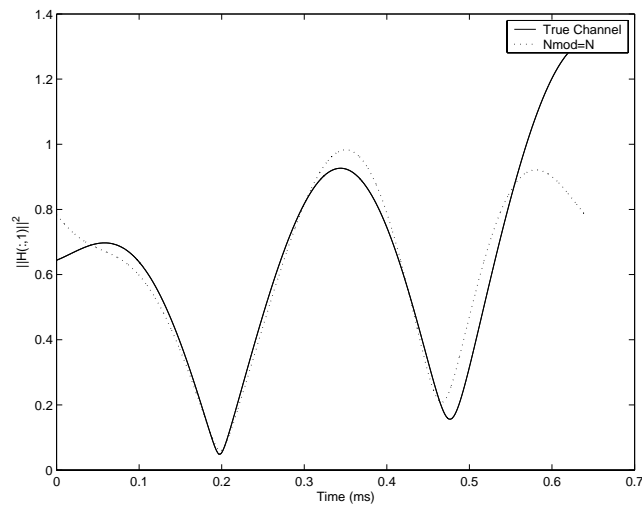
The modeling error is the difference between the best possible BEM (for the chosen Q and N_{mod}) and the true channel. The identification error is the error that is made when the coefficients of the BEM are estimated from the received signals using some identification procedure. Choosing a too small N_{mod} will yield a large modeling error and, as Q , and thus the number of parameters to estimate, becomes smaller in that case, the identification error is relatively small. The mismatch of the identified channel model is then dominated by the BEM-related modeling errors. Choosing a too large N_{mod} will yield a near-zero modeling error, but, as Q increases, the number of parameters that need to be identified from the finite set of observations increases as well and relatively large identification errors become inevitable.

A good empirical rule for the experimental setups that we consider later in this chapter is to choose $N_{mod} = 3N$ and then choose Q according to the Doppler spread rule: $Q = \lceil f_{max} N_{mod} T_s \rceil$, which yields a sufficiently tight match of the BEM with a limited number of parameters. When the channel varies slowly and $1/(3NT_s) \gg f_{max}$, the above procedure yields $Q = 1$ but the Doppler Spread of the BEM will be significantly larger than the true Doppler spread, yielding a poor match of the BEM. In this case, increasing N_{mod} in order to make the true Doppler spread equal to the BEM Doppler spread largely improves the accuracy of the BEM: $N_{mod} = \lceil 1/(T_s f_{max}) \rceil$. Existing methods [98], [104] do not take such considerations into account and simply take $N_{mod} = N$ as the period of the BEM, resulting in large modeling errors that are mainly located at the edges of the considered interval.

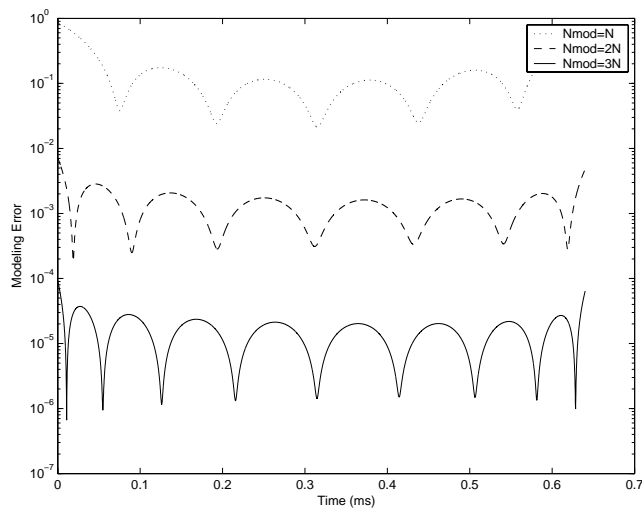
In **Fig. 6.1**, we analyze the modeling error that results from different choices of N_{mod} , i. e. $N_{mod} = N, N_{mod} = 2N$ and $N_{mod} = 3N$. The experimental setup used to generate the channel resembles the Hiperlan setup with an increased mobility. The detailed description of this setup can be found in *sec. 6.3*. The norm of the true channel tap is shown in **Fig. 6.1(a)** together with the norm of the different BEMs. The difference between the true channel and the BEM is not visible when $N_{mod} = 2N$ or $N_{mod} = 3N$, but appears clearly when $N_{mod} = N$. The modeling error in these different configurations is shown in **Fig. 6.1(b)**. The modeling error is approximately divided by 100 each time N_{mod} is increased by N . The different channel identification methods that are presented later in this chapter can lead to relative errors in the range $10^{-3} \sim 10^{-4}$. Hence, the BEM-related modeling error is unacceptably high when $N_{mod} = N$, slightly too high when $N_{mod} = 2N$ and becomes unnoticeable only for $N_{mod} = 3N$.

Defining the $(2Q + 1) \times (L + 1)$ matrix \mathbf{H}_{BEM} of the BEM coefficients as $\mathbf{H}_{BEM}(i, j) = h_{i-(Q+1), j-1}$ and the $(N + L) \times (2Q + 1)$ matrix \mathbf{C} of complex exponentials as $\mathbf{C}(m, n) = e^{j2\pi(n-(Q+1))m/N_{mod}}$, the channel matrix \mathbf{H} is approximated by the BEM as

$$\mathbf{H} \approx \mathbf{C}\mathbf{H}_{BEM}. \quad (6.6)$$



(a)



(b)

Figure 6.1: Norm of the first tap of the true channel and of the channel BEM (a) and of the BEM-induced modeling error (b) vs. time for different values of the modeling interval N_{mod} at a speed of 400 km/h.

The optimal (in the least squares sense) matrix of BEM coefficients is obtained as

$$\mathbf{H}_{BEM} = \mathbf{C}^\dagger \mathbf{H}. \quad (6.7)$$

This optimal set of BEM coefficients is simply called the BEM parameters of the channel in the rest of the text. When the design parameters Q and N_{mod} are chosen following the procedure described above, the difference between the true channel and its BEM representation is negligible in front of the identification errors that will be encountered later in this chapter. Hence, the channel is equivalently described by its $(2Q + 1)(L + 1)$ BEM parameters \mathbf{H}_{BEM} as by the $(N + L)(L + 1)$ parameters of the channel matrix \mathbf{H} .

Using the BEM, the input-output relationship (6.4) of the transmission channel over the packet interval can be written as:

$$y[n] = \sum_{l=0}^L \sum_{q=-Q}^Q h_{q,l} e^{j2\pi qn/N_{mod}} x[n-l] + \eta[n]. \quad (6.8)$$

Note that alternative approaches that allow to model doubly selective channels with a limited number of parameters have been proposed in the literature. A low-complexity basis function approach is proposed in [106]. The time-variations of the transmission channel can also be modeled using wavelet decompositions (see e.g. [107], [108] or [109]). An increased modeling accuracy is obtained when a polynomial description of the channel is used (see e. g. [110] for an overview of polynomial channel fitting and estimation, [111] for a comparison of the polynomial approach with the wavelet approach in a CDMA context, [112] for a combined polynomial-spline approach and [113] for a polynomial approach in an OFDM context). When the parameters Q and N_{mod} are chosen properly, the BEM allows to model doubly selective channels with a similar or higher accuracy than these methods. Another important advantage of the BEM is the existence of high performance low-complexity equalization schemes, with the additional advantage of a remarkably low computational complexity for the design of the equalizers themselves. These low-complexity schemes are enabled by the inherent presence of the complex exponentials in the BEM of the channel (see [102], [114], [115] or [103] for examples of such low-complexity BEM equalization schemes).

6.1.2 KSP in Doubly Selective Channels: Data Model

In this section, we adapt the classical data model for KSP transmission over stationary frequency-selective channels to the context of doubly selective channels.

The sequence of transmitted data symbols remains the same as the one presented in *Ch. 2*, i.e. a block $\mathbf{s}_k(N_s \times 1)$ of data symbols is defined as a column

vector of size N_s , $k = 1 \cdots K$ being the block index:

$$\mathbf{s}_k = [s_k[1], \cdots, s_k[N_s]]^T, \quad (6.9)$$

and a block $\mathbf{t}_k(N_t \times 1)$ of training symbols is similarly defined as:

$$\mathbf{t}_k = [t_k[1], \cdots, t_k[N_t]]^T. \quad (6.10)$$

The k^{th} block of transmitted symbols, \mathbf{x}_k of length $N_x = N_s + N_t$ is obtained with the same precoding matrices as in (2.31) and is thus defined as:

$$\mathbf{x}_k = [\mathbf{s}_k^T, \mathbf{t}_k^T]^T. \quad (6.11)$$

The transmitted sequence is the concatenation of all these blocks: $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \cdots, \mathbf{x}_K^T]^T$ and the total length of the transmitted packet is $N = KN_x$. The received sequence \mathbf{y} , which is derived from (6.4), can be organized in blocks $\mathbf{z}_k(N_x \times 1)$ of received symbols corresponding in the same way as in the stationary channel case, i. e:

$$\mathbf{z}_k = [y[(k-1)N_x + 1], \cdots, y[kN_x]]^T, \quad (6.12)$$

The resulting input-output relationship of the channel can be expressed on a block level in a similar fashion as in (2.32), but the involved matrices now become block-dependent and we adopt the subscript k to denote the block index:

$$\mathbf{z}_k = \mathbf{H}_{KSP,k} \mathbf{s}_k + \mathbf{H}_{t0,k} \mathbf{t}_k + \mathbf{H}_{t1,k} \mathbf{t}_{k-1} + \boldsymbol{\eta}_k, \quad (6.13)$$

where $\boldsymbol{\eta}_k$ is the same AWGN vector as previously and

$$\mathbf{H}_{KSP,k} = \begin{bmatrix} h[n_{k,1};0] & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ h[n_{k,L};L] & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & h[n_{k,N_x-N_t};0] \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h[n_{k,N_x-N_t+L};L] \end{bmatrix}, \quad (6.14)$$

$\mathbf{0}_{(N_t-L) \times N_s}$

$$\mathbf{H}_{t1,k} = \begin{bmatrix} 0 & \cdots & 0 & h[n_{k,1};L] & \cdots & h[n_{k,1};1] \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & h[n_{k,L};L] \\ \vdots & & & & \vdots & 0 \\ \vdots & & & & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & 0 \end{bmatrix}, \quad (6.15)$$

and

$$\mathbf{H}_{t0,k} = [\mathbf{H}_{t0,k}^L \mid \mathbf{H}_{t0,k}^R], \quad (6.16)$$

with

$$\mathbf{H}_{t0,k}^L = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & & \vdots \\ h[n_{k,N_x-N_t+1};0] & \ddots & \vdots \\ \vdots & \ddots & 0 \\ \vdots & & h[n_{k,N_x-L};0] \\ h[n_{k,N_x-N_t+L+1};L] & & \vdots \\ 0 & \ddots & \vdots \\ 0 & 0 & h[n_{N_x};L] \end{bmatrix}, \quad (6.17)$$

and

$$\mathbf{H}_{t0,k}^R = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & & \vdots \\ h[n_{N_x-L+1};0] & \ddots & \vdots \\ \vdots & \ddots & 0 \\ h[n_{N_x};L-1] & \cdots & h[n_{N_x};0] \end{bmatrix}. \quad (6.18)$$

where $\mathbf{H}_{t0,k}^L$ has $(N_t - L)$ columns and $n_{k,j}$ is a shorthand notation for the time index of the j^{th} element of the k^{th} received block \mathbf{z}_k : $n_{k,j} = (k-1)N_x + j$.

6.1.3 KSP in Doubly Selective Channels: Equalizers

Based on the above data model, the conventional KSP equalizers can be adapted to the doubly selective channel situation. We focus on Minimum Mean Squared Error (MMSE) equalization and assume that the receiver has perfect channel knowledge. Note that, since the channel is time-varying, the equalizers will differ from block to block, we thus use the subscript k to denote the block index of the equalizers.

Optimal Linear KSP Equalizers

As in the stationary channel case, the contribution of the known padded sequences is subtracted from the received blocks before the equalization step:

$$\mathbf{z}'_k = \mathbf{z}_k - \mathbf{H}_{t0,k} \mathbf{t}_k - \mathbf{H}_{t1,k} \mathbf{t}_{k-1}. \quad (6.19)$$

The data model then reduces to $\mathbf{z}'_k = \mathbf{H}_{KSP,k}\mathbf{s}_k + \boldsymbol{\eta}_k$, and the MMSE Equalizer is expressed as

$$\mathbf{W}_k = (\mathbf{H}_{KSP,k}^H \mathbf{H}_{KSP,k} + \sigma^2 \mathbf{I}_{N_s})^{-1} \mathbf{H}_{KSP,k}^H, \quad (6.20)$$

where σ^2 is the noise variance.

This optimal linear MMSE equalizer has performs better than its FD counterpart, but it does not always fully exploits the delay diversity of the channel and channel-irrespective symbol recovery. Indeed, because of the channel variations, it may happen that $\mathbf{H}_{KSP,k}$ is not full column rank, in which case the channel-irrespective symbol recovery is not preserved. Its drawback though is its computational complexity. Indeed, besides the complexity of the equalization procedure, which is the same as in the stationary channel situation, a specific equalizer must be computed for each received block, which requires the inversion of an $N_s \times N_s$ matrix. For similar reasons, the KSP block decision feedback equalizer is not considered here as the need to compute a matrix inverse and perform a Cholesky decomposition for each transmitted block, combined with the inherent complexity of the equalization scheme makes its use computationally intractable. However, for the reader interested in low-complexity decision-feedback equalizer structures that are suited to the BEM, we refer to [114] where such structures are discussed and where the authors show that the resulting system performance is equivalent to what would be obtained with the full-complexity BDFE approach.

Frequency-Domain KSP Equalizers

When the coherence time of the channel is larger than the duration of a KSP block ($1/f_{max} \gg N_x T_s$), the channel coefficients can be considered as constant during the transmission of such a KSP block: $||h[n; \nu] - h[n'; \nu]|| \ll ||h[n; \nu]||$, $\forall (n - n') \leq N_x$. In this case, if constant training sequences are used ($\mathbf{t}_k = \mathbf{t}$, $\forall k$), the data model can be expressed with a circulant channel matrix (note that we do not discuss the S-KSP scheme in this section for reasons that are highlighted in the next section where channel identification is discussed):

$$\mathbf{z}_k = \left[\mathbf{H}_{KSP,k} \mid \mathbf{H}_{t0,k} + \mathbf{H}_{t1,k} \right] \begin{bmatrix} \mathbf{s}_k \\ \mathbf{t} \end{bmatrix} = \mathbf{H}_{circ,k} \mathbf{x}_k, \quad (6.21)$$

where $\mathbf{H}_{circ,k}$ is circulant with $[h[kN_x; 0], \dots, h[kN_x; L], 0, \dots, 0]^T$ on the first column and $[h[kN_x; 0], 0, \dots, 0, h[kN_x; L], \dots, h[kN_x; 1]]$ on the first row. Relying on the diagonalization property of such circulant matrices, one can derive block-dependent low-complexity frequency-domain equalizers. The FD MMSE equalizer in this case is expressed as

$$\mathbf{W}_k = \mathcal{I}_{N_x} (\mathbf{H}_{diag,k}^H \mathbf{H}_{diag,k} + \sigma^2 \mathbf{I}_{N_x})^{-1} \mathbf{H}_{diag,k}^H \mathcal{F}_{N_x}(1 : N_s, \cdot), \quad (6.22)$$

where $\mathbf{H}_{diag,k} = \text{diag}(\mathcal{F}_{N_x}[h[kN_x; 0], \dots, h[kN_x; L], 0, \dots, 0]^T)$.

The orthogonality between the virtual carriers exploited by this equalizer is maintained as long as the training sequences are constant and the channel coefficients do not change significantly within one block of symbols.

The main advantage of this equalizer is its low computational complexity, but it yields sub-optimal performance as it does not fully exploit the diversity offered by the delay spread of the channel and does not guarantee channel-irrespective symbol recovery when $\sigma^2 = 0$. Moreover, although one can expect the hypothesis of constant channel coefficients to hold in moderate velocity scenarios, the channel coefficients can vary significantly within a KSP block as the speed increases. In that case, the channel matrix is not circulant anymore and the orthogonality between the carriers is lost. This causes Inter Carrier Interference (ICI) and significantly reduces the performance of the proposed equalizer. Note that some techniques have been developed recently to reduce the ISI resulting from an increased mobility in an OFDM context (see [116]). These techniques linearly combine neighboring tones in order to counteract the ICI and could be adapted to the SC-CP context but are not analyzed here. Note as well that the optimal linear KSP equalizers presented above do not suffer from increased mobility as the channel variations inside a KSP block are accounted for in their data model.

The above discussed equalization schemes are direct adaptations of the classical KSP equalizers. Other equalization strategies than those presented here have been developed that exploit the specific possibilities offered by the BEM approach. High-performance equalizers whose low complexity is enabled by the use of the BEM framework are introduced in [102], [103], [114] and [115] for instance. Note that other interesting approaches have been presented recently for sequence detection in doubly selective channels and we indicate here some of them for the interested reader. A maximum likelihood sequence detector adapted from [117] and relying on a polynomial description of the channel is presented in [118]. Other receiver structures in the polynomial channel description framework are presented in [119]. A thorough study of maximum a posteriori receivers is presented in [120] and several receiver structures adapted to the wavelet context are presented in [107] and [109].

6.2 Estimation of Doubly Selective channels in KSP Transmission

In this section, we analyze how doubly selective channels can be estimated relying on the known symbols inserted by the KSP transmission scheme. As opposed to most channel estimation methods that only rely on the channel

output samples containing contributions only from the known symbols (and thus do not consider the channel output samples containing mixed contributions from the training symbols and the unknown surrounding data symbols), we aim at estimating the channel relying on all the channel output samples containing contributions from the known symbols.

The $(N_t + L) \times 1$ vector of received symbols containing contributions from \mathbf{t}_k is defined as:

$$\mathbf{u}_k = [y[kN_x - N_t + 1], \dots, y[kN_x + L]]^T. \quad (6.23)$$

We propose below three new methods that aim at estimating the doubly selective transmission channel relying on the set of vectors \mathbf{u}_k . Note that these methods are presented in the framework of KSP transmission considered here but it is straightforward to adapt them to other training schemes.

6.2.1 Adaptive Implementation of the Gaussian Maximum Likelihood Method

We propose here an adaptive version of the Gaussian Maximum Likelihood method for stationary channel identification that was presented in *Ch. 4*. This adaptive method is suited for time-varying frequency-selective channels, but can also be used for as a practical implementation of the original method in stationary channel conditions simply by tuning the parameters of the RLS scheme appropriately.

Data Model

Assuming that the coherence time of the channel is significantly larger than $N_t + L$, the channel can be considered constant during the reception of \mathbf{u}_k . We define the vector \mathbf{h}_k of the approximately constant channel coefficients during this time interval as:

$$\mathbf{h}_k \triangleq [h[kN_x; 0], \dots, h[kN_x, L]]^T \triangleq [h_k[0], \dots, h_k[L]]. \quad (6.24)$$

Relying on this definition, the received vector \mathbf{u}_k can be expressed as:

$$\mathbf{u}_k = \mathbf{T}_k \mathbf{h}_k + \boldsymbol{\epsilon}_k. \quad (6.25)$$

- The first term, $\mathbf{T}_k \mathbf{h}_k$ is a deterministic term where \mathbf{T}_k is defined as in *sec. 4.2*, i. e. it is an $(N_t + L) \times (L + 1)$ Toeplitz matrix with $[\mathbf{t}_k^T, 0, \dots, 0]^T$ as its first column and $[t_k[1], 0, \dots, 0]$ as its first row.
- The second term, $\boldsymbol{\epsilon}_k$ is stochastic and represents the contributions from the unknown surrounding data symbols and the AWGN:

$$\boldsymbol{\epsilon}_k = \underbrace{\left[\mathbf{H}_{s,k}^L \mid \mathbf{H}_{s,k}^R \right]}_{\mathbf{H}_{s,k}} \check{\mathbf{s}}_k + \boldsymbol{\eta}_k, \quad (6.26)$$

where $\boldsymbol{\eta}_k$ is the AWGN vector and the vector $\check{\mathbf{s}}_k$ of the unknown data symbols contributing to \mathbf{u}_k is defined as in (4.3). The $(N_t + L) \times 2L$ matrix $\mathbf{H}_{s,k}$ gathers the channel coefficients that multiply these data symbols and is the concatenation of two matrices:

$$\mathbf{H}_{s,k}^L = \begin{bmatrix} h_k[L] & \cdots & h_k[1] \\ & \ddots & \vdots \\ \mathbf{0} & & h_k[L] \\ \hline \mathbf{0}_{(N_t \times L)} \end{bmatrix}, \mathbf{H}_{s,k}^R = \begin{bmatrix} \mathbf{0}_{(N_t \times L)} & & \\ h_k[0] & & \mathbf{0} \\ \vdots & \ddots & \\ h_k[L-1] & \cdots & h_k[0] \end{bmatrix}. \quad (6.27)$$

Keeping the same hypotheses as in *Ch. 4* on the statistics of the noise and data symbols, the first and second order statistics of $\boldsymbol{\epsilon}_k$ (assuming also $N_s \geq 2L$) are readily derived as:

$$\begin{aligned} \mathbb{E}\{\boldsymbol{\epsilon}_k\} &= \mathbf{0}, \\ \mathbb{E}\{\boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_l^H\} &= \delta_{k,l} \mathbf{Q}_k, \quad \forall k, l, \\ \mathbf{Q}_k &= \lambda^2 \mathbf{H}_{s,k} \mathbf{H}_{s,k}^H + \sigma^2 \mathbf{I}. \end{aligned} \quad (6.28)$$

Proposed Algorithm

We have shown in *Ch. 4* that when the channel is stationary ($\mathbf{h}_k = \mathbf{h}$ and $\mathbf{Q}_k = \mathbf{Q}$, $\forall k$), the closed-form GML channel estimate of \mathbf{h} is obtained as:

$$\hat{\mathbf{h}} = \left(\sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}^{-1} \mathbf{T}_k \right)^{-1} \sum_{k=1}^K \mathbf{T}_k^H \hat{\mathbf{Q}}^{-1} \mathbf{u}_k, \quad (6.29)$$

where $\hat{\mathbf{Q}}$ can be seen as an estimate of the noise correlation matrix \mathbf{Q} based on the received data.

We propose here an exponentially weighted RLS adaptation of this technique to track the time variations of the channel between successive blocks:

$$\hat{\mathbf{h}}_k = \left(\sum_{n=1}^k \lambda^{2(k-n)} \mathbf{T}_n^H \hat{\mathbf{Q}}_n^{-1} \mathbf{T}_n \right)^{-1} \sum_{n=1}^k \lambda^{2(k-n)} \mathbf{T}_n^H \hat{\mathbf{Q}}_n^{-1} \mathbf{u}_n, \quad (6.30)$$

where λ is the forget factor whose value is adapted depending on the variation speed of the channel and the noise conditions. Adopting the notations $\mathbf{S}_1(k)$

and $\mathbf{S}_2(k)$ for the summations present in (6.30), the following equalities are readily derived:

$$\begin{aligned}\mathbf{S}_1(k+1) &= \lambda^2 \mathbf{S}_1(k) + \mathbf{T}_{k+1}^H \hat{\mathbf{Q}}_{k+1}^{-1} \mathbf{T}_{k+1}, \\ \mathbf{S}_2(k+1) &= \lambda^2 \mathbf{S}_2(k) + \mathbf{T}_{k+1}^H \hat{\mathbf{Q}}_{k+1}^{-1} \mathbf{u}_k, \\ \hat{\mathbf{h}}_{k+1} &= (\mathbf{S}_1(k+1))^{-1} \mathbf{S}_2(k+1).\end{aligned}\quad (6.31)$$

A recursive expression for the channel estimate is then obtained as:

$$\hat{\mathbf{h}}_{k+1} = \hat{\mathbf{h}}_k + (\mathbf{S}_1(k+1))^{-1} \mathbf{T}_{k+1}^H \hat{\mathbf{Q}}_{k+1}^{-1} (\mathbf{u}_{k+1} - \mathbf{T}_{k+1} \hat{\mathbf{h}}_k), \quad (6.32)$$

where the explicit computation of $\mathbf{S}_2(k+1)$ is not needed anymore. Note that an estimate of the noise correlation matrix $\hat{\mathbf{Q}}_{k+1}$ is needed for this method to work. In the stationary channel case, the optimal estimate of \mathbf{Q} yielding the ML channel estimate was directly derived from the received symbols. This approach is not possible here as \mathbf{Q} cannot be estimated from a single vector \mathbf{u}_k . We thus propose to derive $\hat{\mathbf{Q}}_{k+1}$ directly from its definition, assuming that $\hat{\mathbf{h}}_k$ is an acceptable approximation of \mathbf{h}_{k+1} . Hence, we use $\hat{\mathbf{h}}_k$ to construct the matrix $\hat{\mathbf{H}}_{s,k+1}$ from its definition (6.27) and, relying on that estimate $\hat{\mathbf{H}}_{s,k+1}$, we can construct $\hat{\mathbf{Q}}_{k+1}$ from its definition (6.28). The resulting $\hat{\mathbf{Q}}_{k+1}$ can then be used in (6.32).

This method has the advantage of a low computational complexity and a low latency as the channel is estimated ‘on the fly’, allowing to directly equalize the received samples without keeping them in a receive buffer. Its performance is very sensitive to the value of the parameter λ . Low values of λ allow to track fast channel variations, but make the channel estimates more noisy. Higher values of λ allow to filter out the noise, but fast channel variations are harder to follow and the channel estimate becomes biased. The optimal choice of λ results from a tradeoff between channel tracking and noise filtering. The method can be initialized with $\hat{\mathbf{h}}_0 = \mathbf{0}$ and $\hat{\mathbf{Q}}_1 = \mathbf{I}$ but will yield more reliable estimates for the first KSP blocks if a reliable initial channel estimate $\hat{\mathbf{h}}_0$ resulting from a long training sequence is provided.

Using this method, channel models are available for K time instants only (we assume that $\hat{\mathbf{h}}_k$ approximates the channel for the last sample of \mathbf{z}_k , i.e. $\hat{\mathbf{h}}_k \approx [h[kN_x; 0], \dots, h[kN_x; L]]^T$). When the full channel model is needed (computation of the optimal linear KSP equalizers for instance), we use a simple linear interpolation between these points.

The idea of tracking the evolution of the doubly selective channel with adaptive methods such as this RLS scheme is not new. In [121], the authors study an RLS implementation of a relatively simple LS scheme and derive optimal values for the forget factor. The RLS scheme we propose here differs from existing ones by the underlying method that is recursively implemented, i.e. a GML channel identification method.

Finally note that the RLS scheme proposed here can be successfully used in stationary channel conditions as well. If the value of the forget factor is chosen as $\lambda = 1$, the resulting solution is close to the closed-form solution¹ that would have been obtained if all the received blocks had been considered jointly. Hence, using this implementation, it becomes possible to refine the channel estimate at each received block and adapt the equalizers that can then be used directly for the incoming blocks of data symbols. Hence, if this RLS approach is used, it is not needed to keep all the data symbols of the packet in a buffer before processing the whole packet after its reception, which was the case with the closed-form solution.

6.2.2 Adaptive BEM Method

Alternatively to the above proposed linear interpolation, it is possible to fit the channel estimates produced by the adaptive method into a BEM as soon as $K \geq 2Q + 1$ (which is generally the case). In this case, as the BEM has only $2Q + 1$ degrees of freedom, it is generally not possible to interpolate between the identified channel points and we have to seek the BEM that best fits them in the least square sense.

Defining the $K \times N$ selection matrix $\tilde{\mathbf{S}}$ as

$$\begin{aligned}\tilde{\mathbf{S}}[i, N_x i] &= 1, \quad \forall i = 1 \cdots K, \\ \tilde{\mathbf{S}}[i, j] &= 0 \quad \text{elsewhere,}\end{aligned}$$

and the matrix \mathbf{H}_{adap} of the channel coefficients identified by the adaptive method (6.32) as:

$$\mathbf{H}_{adap} = [\hat{\mathbf{h}}_1, \cdots, \hat{\mathbf{h}}_K]^T, \quad (6.33)$$

we are seeking the set of BEM coefficients $\hat{\mathbf{H}}_{BEM}$ that yield the best match (in a LS sense) to the adaptively identified channel coefficients for the K considered time instants. Using the selection matrix and the previously defined matrix \mathbf{C} of complex exponentials, this translates into:

$$\tilde{\mathbf{S}}\mathbf{C}\hat{\mathbf{H}}_{BEM} \approx \mathbf{H}_{adap}, \quad (6.34)$$

which is solved in a least squares sense as

$$\hat{\mathbf{H}}_{BEM} = (\tilde{\mathbf{S}}\mathbf{C})^\dagger \mathbf{H}_{adap}. \quad (6.35)$$

The parameters Q and N_{mod} of the BEM are designed as outlined earlier, assuming that the Doppler spread is known. By doing so, the dynamics of the

¹It is not exactly equal to the true GML solution though, as the matrix $\hat{\mathbf{Q}}$ that is used here does not correspond to the one used in the closed-form ML solution. Hence, the obtained solution will differ from the ML solution.

BEM are in line with those of the true channel and the noise-induced variations of the adaptively identified channel coefficients cannot be modeled. The BEM modeling step actually filters out a part of the noise that is present in the adaptively identified channel coefficients. Therefore, the optimal choice of λ is different when the BEM approach is used. Lower values of λ should yield better results as it allows an improved tracking of the channel variations, the increased noise in the channel estimates being filtered out by the BEM modeling step.

6.2.3 Direct Estimation of the BEM parameters

In this section, we propose a new approach in order to directly identify the BEM parameters of the channel, rather than the indirect approach of the previous section where the channel was partly identified before to seek the BEM coefficients that best suit the partly identified channel. The proposed approach directly relies on the knowledge of the training symbols appended by the KSP transmission scheme and blindly filters out the contributions of the surrounding data symbols. It is thus a semi-blind method, as the GML method presented in *Ch. 4* Existing methods for the direct BEM estimation of doubly selective channels are either purely training-based (see [98] or [104]) and only exploit the channel output samples that solely contain contributions from the training sequence \mathbf{t}_k , discarding all the channel output samples containing mixed contributions from \mathbf{t}_k and the unknown data symbols \mathbf{s}_k , or they are purely blind [122] and are not able to incorporate a-priori knowledge of the transmitted symbols. The method we present here exploits all the channel output samples that contain contributions from \mathbf{t}_k , including those which contain contributions from both unknown data symbols and training symbols (i.e. we rely on the set of vectors \mathbf{u}_k defined in (6.23)) and thus has the potential to outperform existing methods.

Data Model

Re-arranging the BEM expression of the transmission scheme (6.8), we obtain the following data model that is well suited for the identification of the channel's BEM parameters:

$$\mathbf{u}_k = \mathcal{T}_k \mathbf{h}_{BEM} + \boldsymbol{\epsilon}_k. \quad (6.36)$$

- The first term, $\mathcal{T}_k \mathbf{h}_{BEM}$ is a deterministic term where \mathbf{h}_{BEM} is the $(2Q + 1)(L + 1)$ -wide vector of the channel's BEM coefficients: $\mathbf{h}_{BEM} = [h_{-Q,0}, \dots, h_{-Q,L}, \dots, h_{Q,L}]^T$ and \mathcal{T}_k is an $(N_t + L) \times ((2Q + 1)(L + 1))$ matrix accounting for the contributions of the complex exponentials of the BEM and the training sequences, which has the following structure:

$$\mathbf{T}_k = \left[\begin{array}{c} \boxed{\mathbf{T}_{k,0}} \\ \boxed{\mathbf{T}_{k,1}} \cdots \boxed{\mathbf{T}_{k,L}} \end{array} \right], \quad (6.37)$$

with $\mathbf{T}_{k,l} = \text{diag}(\mathbf{t}_k) \mathbf{C}'_{k,l}$, where $\mathbf{C}'_{k,l}$ accounts for the BEM's complex exponentials multiplying the $h_{q,l}$ coefficients: $\mathbf{C}'_{k,l}(x, y) = e^{j2\pi(y-(Q+1))\frac{kN_x - N_t + l + x - 1}{N_{\text{mod}}}}$.

• The second term, ϵ_k can be expressed as in (6.26), but the time variations of the channel during the reception of \mathbf{u}_k are taken into account and so the left and right parts of $\mathbf{H}_{s,k}$ are redefined as:

$$\mathbf{H}_{s,k}^L = \left[\begin{array}{ccc} h[n_{k,1}; L] & \cdots & h[n_{k,1}; 1] \\ & \ddots & \vdots \\ \mathbf{0} & & h[n_{k,L}; L] \\ \hline & \mathbf{0}_{(N_t \times L)} & \end{array} \right], \quad (6.38)$$

$$\mathbf{H}_{s,k}^R = \left[\begin{array}{ccc} & \mathbf{0}_{(N_t \times L)} & \\ h[n_{k,N_t}; 0] & & \mathbf{0} \\ \vdots & \ddots & \\ h[n_{k,N_t+L}; L-1] & \cdots & h[n_{k,N_t+L}; 0] \end{array} \right], \quad (6.39)$$

where $n_{k,l}$ is a shorthand notation for the index of the l^{th} element of \mathbf{u}_k : $n_{k,l} = kN_x - N_t + l$.

Proposed Algorithms

The statistics of ϵ_k defined in (6.28) remain valid in this situations under the same assumptions that were needed to derive them. We directly rely on these statistics to propose the following methods.

LS Channel Estimate

Relying on the first-order statistics of ϵ_k , a simple Least Squares (LS) approach

provides us with an unbiased estimator of \mathbf{h}_{BEM} :

$$\hat{\mathbf{h}}_{LS} = \left(\sum_{k=1}^K \mathcal{T}_k^H \mathcal{T}_k \right)^{-1} \sum_{k=1}^K \mathcal{T}_k^H \mathbf{u}_k. \quad (6.40)$$

Because of the presence of the complex exponentials, the inverse of the sum will always exist as soon as $K(N_t + L) \geq (2Q + 1)(L + 1)$.

WLS Channel Estimate

Since ϵ_k is not white, the LS approach is not optimal. A Weighted Least Squares (WLS) approach taking into account the color of ϵ_k would yield an improved estimate of the channel parameters. *Assuming that all the \mathbf{Q}_k 's are known* (see also next paragraph), the WLS estimate of \mathbf{h}_{BEM} can be computed as:

$$\hat{\mathbf{h}}_{WLS} = \left(\sum_{k=1}^K \mathcal{T}_k^H \mathbf{Q}_k^{-1} \mathcal{T}_k \right)^{-1} \sum_{k=1}^K \mathcal{T}_k^H \mathbf{Q}_k^{-1} \mathbf{u}_k. \quad (6.41)$$

The presence of the AWGN term in \mathbf{Q}_k ensures the existence of its inverse and the inverse of the sum exists under the same conditions as for the LS estimate.

Iterative WLS Channel Estimate

Unfortunately, \mathbf{Q}_k is not known at the receiver for it depends on the sought channel. The WLS approach can thus not be straightforwardly adopted. We propose below an iterative approach that allows to cope with the dependence of \mathbf{Q}_k on the channel.

Assume a channel estimate $\hat{\mathbf{h}}_{BEM}^{(i)}$ is available at the receiver (i^{th} iteration). Exploiting (6.5) and the definition of $\mathbf{H}_{s,k}$, it is possible to construct its estimate $\hat{\mathbf{H}}_{s,k}^{(i)}$ from $\hat{\mathbf{h}}_{BEM}^{(i)}$, for $k = 1 \cdots K$. Relying on the parametric definition of \mathbf{Q}_k and assuming that σ^2 is known, we construct the K estimates $\hat{\mathbf{Q}}_k^{(i)}$ of the color of the different ϵ_k 's. This estimate is used to produce a refined estimate $\hat{\mathbf{h}}_{BEM}^{(i+1)}$ of the channel model with a WLS approach:

$$\hat{\mathbf{h}}_{BEM}^{(i+1)} = \left(\sum_{k=1}^K \mathcal{T}_k^H \hat{\mathbf{Q}}_k^{(i)-1} \mathcal{T}_k \right)^{-1} \sum_{k=1}^K \mathcal{T}_k^H \hat{\mathbf{Q}}_k^{(i)-1} \mathbf{u}_k. \quad (6.42)$$

The iterative procedure is stopped when there is no significant difference between two consecutive channel estimates. If the starting point is sufficiently accurate, this iterative procedure converges to a solution which is close to the true WLS estimate.

The iterative procedure can be initialized with the LS channel estimate of (6.40): $\hat{\mathbf{h}}_{BEM}^{(0)} = \hat{\mathbf{h}}_{LS}$, which is equivalent to choosing $\hat{\mathbf{Q}}_k^{(0)} = \mathbf{I}$, $\forall k$. Experimental results show that this choice allows the iterative procedure to converge in 2 or 3 steps. The experimental results presented below are obtained with 2 iterations of the iterative procedure.

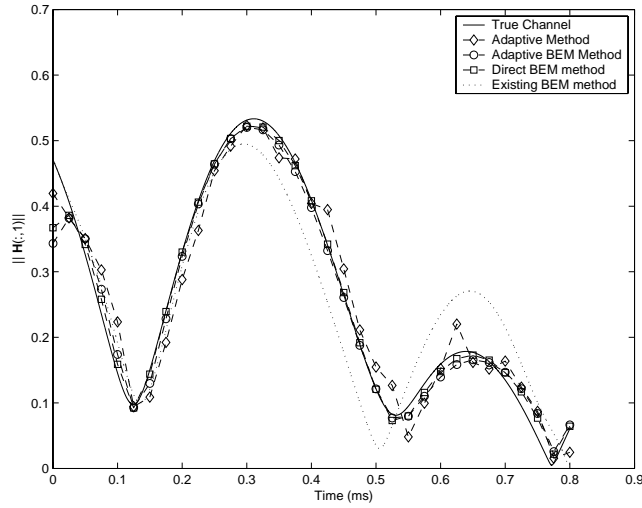
It is interesting to note the similarity between the WLS channel estimate and the closed-form GML channel estimate (4.30) proposed in *Ch. 4*. It is possible to show that, under the hypothesis that the true channel and its optimal set of BEM coefficients (6.7) match each other perfectly, using the GML estimate (or equivalent appropriate estimations) of the different \mathbf{Q}_k matrices would yield the GML channel estimate. However, as there is one such matrix to identify per received block of training, the problem is under determined and it is not possible to find directly the GML expression of the BEM channel estimate. The proposed iterative method tries to overcome this problem using the parametric definition of \mathbf{Q}_k , and, though it is not a purely ML method, it will approach the ML BEM channel estimate.

Note as well that we did not discuss the proposed S-KSP scheme in this section. The reason therefore is that there is no advantage here in terms of achievable channel estimation accuracy between the constant and changing padded sequence situations. Indeed, the different \mathcal{T}_k matrices differ from one another regardless of the training sequences composition. Hence, using constant training sequences in time varying channels simultaneously allows for accurate channel estimation and enables the use of low-complexity FD equalizers (if the Doppler spread is not too important though).

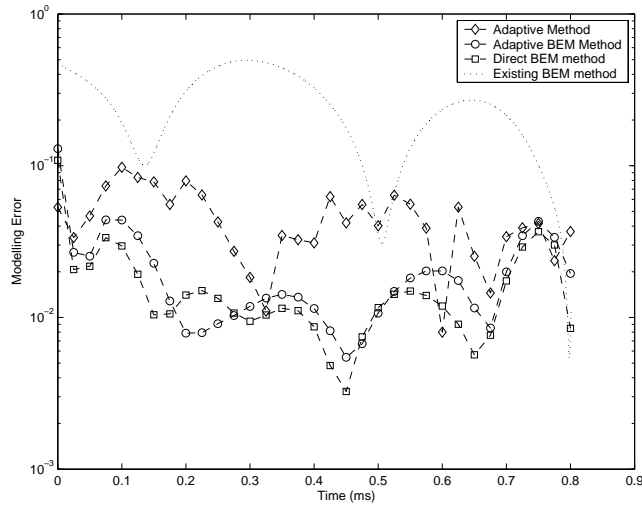
Finally note that alternative methods have been proposed recently for the identification of doubly selective channels when alternative parameterizations of the channel are used. The problem of estimating the transmission channel in the wavelet framework is addressed in [108] or [109]. Several approaches towards the estimation of the polynomial channel model can be found in [111], [110] or [119].

6.3 Simulation Results

In this section, we assess the performance of the proposed channel estimation methods and the BER performance of the proposed KSP equalizers when the proposed channel estimates are used for their design. The considered system setup is similar to the Hiperlan2 standard for WLAN applications, except that KSP transmission is considered rather than OFDM transmission. We also modify the sampling time which is increased by a factor 2: $T_s = 0.1\mu s$ instead of the proposed $0.05\mu s$, which increases the effects of mobility. The other



(a)



(b)

Figure 6.2: Results of different channel estimation methods at a speed of 400 km/h and with a SNR of 20 dB. The norm of the first tap of the channel estimates are shown together with the true channel in (a) and the corresponding estimation error is shown in (b).

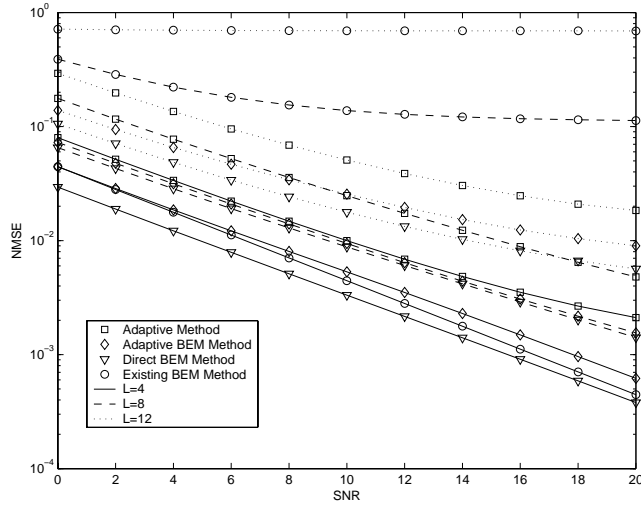
parameters of the transmission scheme are as prescribed by the standards: $f_c = 5.5\text{GHz}$, guard band duration (N_t in our scheme) of 16 samples, useful symbol duration (N_s) of 64 samples. The padded sequences of the KSP schemes are randomly picked QPSK symbols. The constellations used for the mapping of the binary data are Gray-mapping BPSK, QPSK, 16-QAM and 64-QAM. When coded transmissions are considered, the convolutional mother code of rate $r = 1/2$ of the standard is used. No preambles for synchronization or channel estimation are inserted before data transmission and we assume perfect timing and carrier frequency synchronization. The number of KSP blocks transmitted (K in the previously adopted notation) is set to 100. The transmission channels used for the simulations are random realizations of a physical channel model with 10 clusters of 100 rays each with a channel order set to $L = 8$ for most experiments (we explicitly state so when the considered channel order differs from 8).

The considered speed of the mobile terminals ranges from 25 km/h (low mobility in an urban environment) to 400 km/h (high speed train), yielding Doppler spreads in the interval $127\text{Hz} \leq f_{max} \leq 2\text{kHz}$. The corresponding interval for the coherence time of the channel is $78\mu\text{s} \leq \tau_{coh} \leq 1.3\text{ms}$. The total duration of a packet in the considered scenario is 0.8 ms, whereas the duration of a KSP block is 8 μs . The values of the coherence time indicate that we are indeed in a moderate mobility scenario, i.e. the channel changes significantly during the transmission of a data block, but the variation inside a KSP block are limited. Analyzing the envelope correlation coefficient confirms this. The interval for this coefficient evaluated over the duration of a KSP block for the considered speed interval is $0.995 \leq \rho(\Delta_{t,bl}) \leq 1$, where $\Delta_{t,bl}$ is the duration of a KSP block. Similarly, the interval for the envelope correlation coefficient over the duration of the whole packet is $0.0614 \leq \rho(\Delta_{t,pack}) \leq 0.811$, where $\Delta_{t,pack}$ is the duration of the full packet. Note also that the lower bound on $\rho(\Delta_{t,bl})$ shows that the variation of the channel inside a KSP block is noticeable but of limited importance for the highest considered speed.

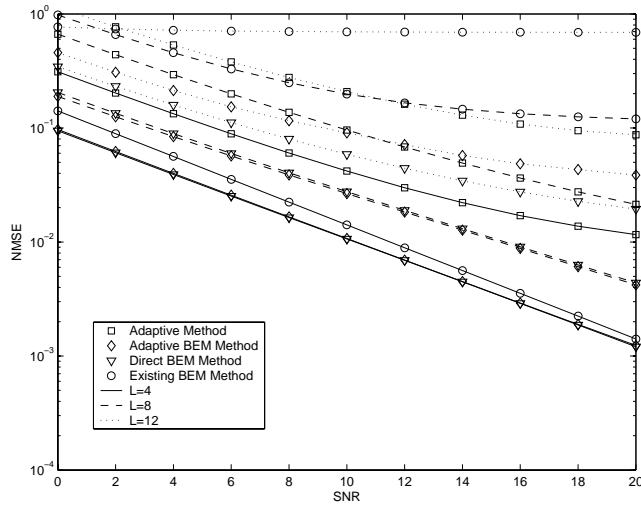
The number of complex exponentials ($2Q + 1$) required to accurately track the channel variations in this scenario ranges from 3 ($Q = 1$) at 25 km/h to 9 ($Q = 4$) at 400 km/h and the period N_{mod} of the BEM is derived from the procedure described in *sec. 6.1.1*. The forget factor of the adaptive method ranges from $\lambda = 0.95$ at 25 km/h to $\lambda = 0.75$ at 400 km/h. For the proposed adaptive BEM method, we pick $\lambda = 0.1$ as experimental results indicate that it yields the best accuracy.

6.3.1 Channel Estimation

In a first experiment, we compare the proposed channel estimation methods (BEM and adaptive) with the method proposed in [98] and [104] (note that



(a)



(b)

Figure 6.3: Statistical averaging of the Normalized Mean Square Error obtained with the different channel estimation techniques when increasingly large channel orders are considered with a fixed length $N_t = 16$ for the padded sequences. The considered user terminals speeds are 25 km/h (a) and 400 km/h (b).

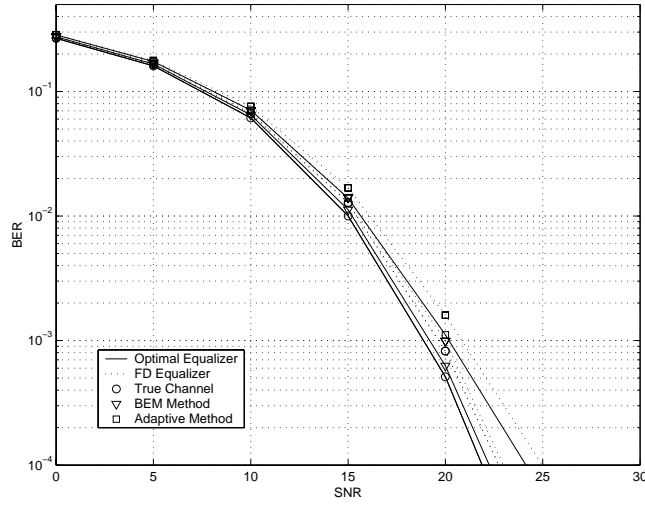
we generalize the method to handle arbitrary BEM periods, which significantly improves their original performance). In **Fig. 6.2**, we show how the different methods track the evolution of a given channel tap with a mobile terminal speed of 400 km/h under an SNR of 20 dB. The proposed BEM methods seem to match the true channel better than the others with a slight advantage for the direct BEM method over the adaptive BEM method. The channel estimates of the purely adaptive method (initialized with $\mathbf{h}_0 = \mathbf{0}$) vary less smoothly in time than those of the existing BEM method [98] and [104] but are closer to the true channel.

The performance of the different channel estimation schemes is further assessed in **Fig. 6.3**, where the Normalized Mean Square Error (NMSE= $\|\mathbf{H} - \hat{\mathbf{H}}\|^2/\|\mathbf{H}\|^2$) of the different channel estimation methods is presented as a function of the SNR for different channel orders ($L=4, 8$ and 12) for the extremes of the considered user terminal speeds. The proposed BEM methods clearly outperform the purely adaptive scheme as well as the existing BEM method. The performances of the proposed BEM methods are relatively close to each other with some advantage for the direct BEM method. For low channel orders only, the existing BEM method outperforms the purely adaptive one, but for higher channel orders, the existing BEM method shows poor performance and is outperformed by the two proposed schemes. The experiment further shows that the accuracy of all channel estimates is reduced as the speed of the mobile terminal increases. A drawback of the purely adaptive method is that the choice of the forget factor λ results from a tradeoff: when λ approaches 1, the noise is filtered out and its impact on the channel estimates is limited, but the channel variations cannot be tracked accurately. When λ is smaller, the channel variations can be tracked accurately but the noise has an increased impact on the modeling error and the final channel estimate becomes noisy. In contrast, the BEM methods average out the noise over the duration of the considered packet and models the channel variations with its complex exponentials. No tradeoff between channel tracking and noise reduction has to be made as both are done optimally, which results in improved channel estimates.

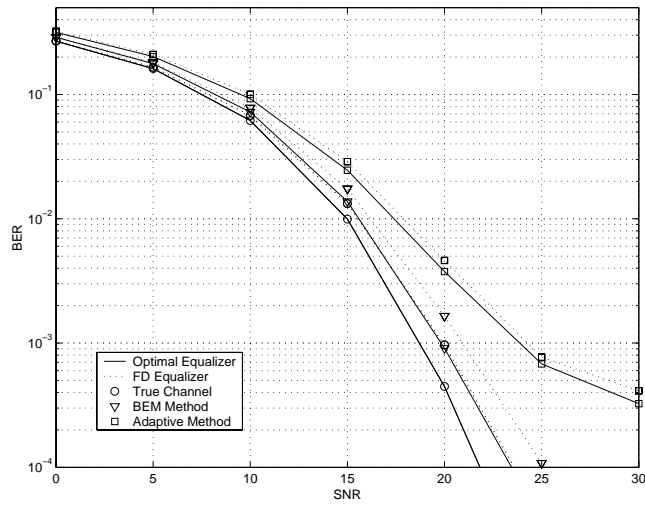
6.3.2 Channel Equalization

In a second experiment of which the results are presented in **Fig. 6.4**, we analyze the BER performance of the presented KSP equalizers (both the optimal linear equalizer and less complex FD equalizer are considered) for QPSK transmission when two of the proposed channel estimation methods and the true channel model are used². The considered channel identification methods are the direct BEM approach that yields the most accurate channel model

²When BEM channel identification is used, the FD equalizer for the k^{th} received block is computed using the BEM-derived channel model for the middle sample of the considered block.



(a)



(b)

Figure 6.4: BER performance obtained relying on the channel estimates obtained with the purely adaptive and the direct BEM channel estimation methods. Both optimal and FD linear KSP equalizers are considered the transmitted sequence is uncoded and mapped onto a QPSK constellation. The considered user terminals speeds are 25 km/h (a) and 400 km/h (b).

and the purely adaptive method that has the smallest computational complexity. The performance of the optimal linear KSP equalizer with perfect channel knowledge appears to be speed-independent. When QPSK constellations are used, we have seen earlier that the optimal and FD linear KSP equalizer yield similar performance at the considered SNRs when the channel is stationary. In the presented experiments, it appears that the use of the FD equalizer yields a performance degradation of approximately 0.75 dB at 25 km/h and 2 dB at 400 km/h for a target BER of 10^{-4} . This performance degradation is caused by the time variations of the channel inside a KSP block, resulting in the loss of orthogonality between the carriers and uncompensated ICI.

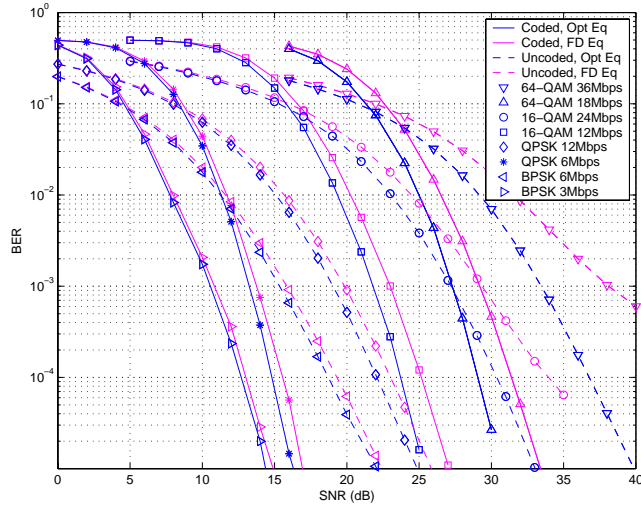
When the channel models obtained with the direct BEM method are used for the design of the equalizers, the BER performance matches quite closely the perfect channel knowledge situation and the performance degradation is quite limited.

When the channel estimates obtained with the adaptive method are used, a relatively good match is observed at low speeds but, as the speed increases and the channel model becomes less accurate, we observe the appearance of a floor in the BER performance of the system.

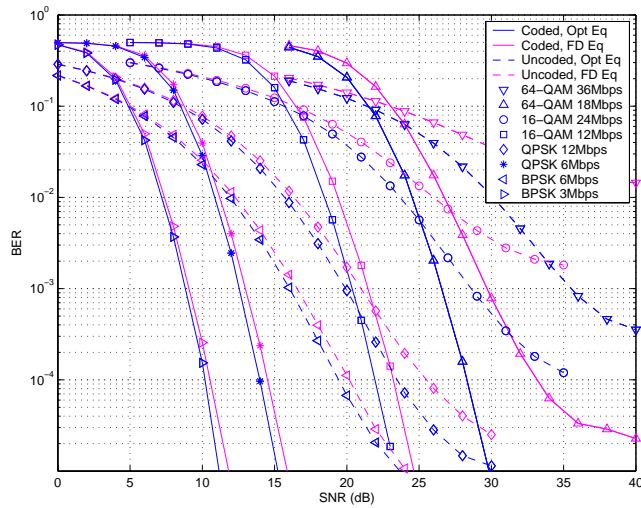
This experiment shows that the higher accuracy of the direct BEM method results in significantly improved BER performance than when other channel identification methods are used. It shows as well that the mobility has no impact on the performance of the optimal linear KSP equalizers. The FD equalizers are affected by the effects of the mobility but the performance degradation is not dramatic in the considered setup. The main source of performance degradation in this context are the channel modeling errors.

In a third experiment, we further investigate the BER performance of the system when the direct BEM identification method is used and the optimal and FD linear KSP equalizers are used. As proposed in the Hiperlan2 Standard, the effective data rates are varied through the use of different constellations and coding schemes. We assess the system performance for BPSK, QPSK, 16-QAM and 64-QAM constellations for coded and uncoded transmission. In coded transmission, we only analyze the use of the rate $r = 1/2$ binary convolutional encoder as described in the Hiperlan2 and IEEE 802.11a standards. The combination of the different constellations and coding rates allows 6 different data rates ranging from 3 to 36 Mbps.

It is important to note the possible advantage yielded by the use of the error-correcting codes in this context of time-varying channels. We have seen earlier that, when the channel is stationary, the optimal linear KSP equalizer allows to fully exploit the delay diversity of this channel and the only performance advantage brought by the coding schemes is their intrinsic coding and diversity gains. The maximal diversity advantage offered by the channel is equal to the



(a)



(b)

Figure 6.5: BER performance obtained relying on the channel estimates obtained with the direct BEM channel estimation method. Both optimal and FD linear KSP equalizers are considered. Both uncoded and $r = 1/2$ coded transmission are considered, and this for different constellation sizes. The considered user terminals speeds are 25 km/h (a) and 400 km/h (b).

number of independent copies of the transmitted signal that are delivered by the channel. In stationary channels, this diversity order is thus equal to $L + 1$ when the channel taps are uncorrelated or less if they are correlated. When the channel changes in time however, the maximum achievable diversity order increases. A diversity study presented in [123] shows that the total diversity order is proportional to the total number of degrees of freedom that the channel exhibits in the time and frequency dimensions, i.e. it is proportional to $(2Q + 1)(L + 1)N/N_{mod}$ when the channel taps are uncorrelated in the time. The KSP equalizers are able to fully exploit the channel diversity within a single block of data, which is the frequency-domain diversity and will hardly exceed $L + 1$ in the considered context. However, they are not able to exploit the time-domain diversity of the channel as the data symbols are not spread across different channel realizations. This problem is discussed in [123] where block precoders allowing to fully exploit the time-frequency selectivity of the channel are presented. In this context, the convolutional codes that are present in the standards represent an interesting opportunity to exploit this diversity. Just like they enabled OFDM to benefit from the frequency-domain diversity by spreading across independent tones information that would otherwise have been confined to a single tone, these codes can enable the KSP scheme to exploit the time-domain selectivity by spreading across different blocks (for which the channels are as uncorrelated as possible) information that would otherwise have been transmitted in a single block with a single realization of the transmission channel.

In order to enable the exploitation of the time-domain diversity, the coded scheme is thus used in combination with an interleaver that differs from the block-interleaver of the standard. In order to spread the information across independent channel realizations, the proposed interleaver operates on a packet level. It is designed such that 2 consecutive data bits are always encoded in different KSP blocks with a minimum distance of 15 blocks between these two blocks.

The BER performance of the system is presented in **Fig. 6.5** for speeds of 25 and 400 km/h. We first discuss the performance of the uncoded scheme and then focus on the coded performance.

For low-order constellations (BPSK and QPSK), the uncoded performance of optimal and FD linear KSP equalizers are close to one another with some advantage for the optimal scheme, and this advantage is increased at higher speeds. For higher-order constellations, using the more complex optimal linear KSP equalizer yields a significant performance gain, especially as the speed increases. The effect of increased speeds results in a performance degradation for both equalization schemes and an error floor appears at high speeds when 16- or 64-QAM constellations are used. This error floor results from the increased modeling errors at high speeds (plus the loss of carrier orthogonality for the FD equalization scheme) and is much higher for the FD equalizers than for the

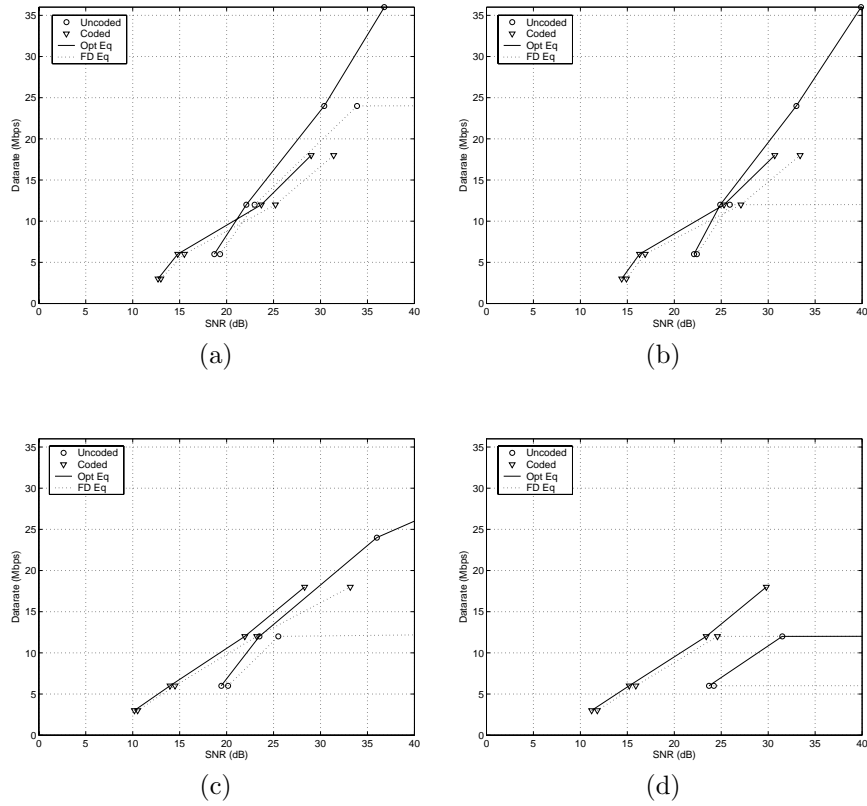


Figure 6.6: Achievable data rates when the direct BEM channel estimation method is used in combination with different coding schemes and different constellations. Both optimal and FD linear KSP Equalizers are considered. The target bit error rate is 10^{-4} in (a) and (c) and 10^{-5} in (b) and (d). The considered mobile terminal speeds is 25 km/h in (a) and (b) and 400 km/h in (c) and (d).

optimal linear equalizers.

In contrast to this, coded transmissions show stable or (slightly) improved performance as the speed increases even though the uncoded performance undergoes a significant degradation. This results from the increased diversity offered by higher terminal speeds that is exploited by the coded schemes. The diversity-related performance gain allows to cancel the channel modeling error-induced performance degradation of the uncoded schemes. The performance gap between optimal and FD linear equalization is largely reduced as coded schemes are used; both schemes offer similar performance for BPSK and QPSK constellations with a small advantage for the optimal linear equalization scheme. For higher order constellations, the optimal linear KSP equalizer offers a performance gain ranging from 1 to 3.5 dB depending on the constellation and the speed of the mobile terminal.

Finally, the achievable data rates of the different schemes (coded and uncoded transmission, FD or optimal linear equalizers) for target bit error rates of respectively 10^{-3} (**Fig. (a)** and **(c)**) and 10^{-4} (**Fig. (b)** and **(d)**) as a function of the SNR are presented in **Fig. 6.6** for speeds of 25 (**Fig. (a)** and **(b)**) and 400km/h (**Fig. (c)** and **(d)**) respectively. The coded scheme always dominates in the low SNR region. At low speeds and for low target bit error rates, the uncoded scheme performs best in the medium and high SNR regions. The difference between coded and uncoded schemes reduces as the speed or BER performance requirement increases. For low target BERs and high speeds, the coded schemes always perform better. The presented results also highlight the more stable performance of the coded scheme for different speeds and target BERs.

6.4 Conclusions

In this chapter, we have introduced three new methods that allow to identify doubly selective channels relying on the knowledge of the known symbols introduced by a KSP transmission scheme. The first method copes with the time-variations of the transmission channel with an adaptive scheme. The second method fits the results of the adaptive model in a BEM framework, which results in improved performance. Finally, the last method directly identifies the BEM parameters of the channel and outperforms others. These methods are able to cope with training sequences of various lengths and compositions. Taking into account the channel output samples that contain contributions from both the training symbols and the unknown surrounding data symbols allows the proposed methods to clearly outperform existing ones, especially for large channel orders. The proposed BEM methods perform significantly better than the adaptive one but have a higher latency and require to buffer the channel

output samples before the channel is identified and equalizers can be designed, whilst the adaptive method estimates the channel on the fly.

Block equalizers for KSP transmission were also described and experimental results show their efficiency. They also show that the above discussed channel estimates are sufficiently accurate to enable reliable channel equalization. Note that the channel estimates obtained here could be used in combination with other, possibly improved equalization schemes. It is possible to use these estimates for BEM-specific equalizer structures, but the full channel model that we obtain can be used in other frameworks such as ML, MAP, polynomial or wavelet-based frameworks.

Finally, experiments where the design of the proposed KSP equalizers relies on the channel estimates provided by the proposed BEM method show that KSP transmission is a suitable candidate in order to deliver high data rates ranging from 3Mbps to 36 Mbps to users experiencing high mobility. When these equalizers are used in combination with a coding scheme allowing to exploit the diversity offered by the time variation of the channel, it is possible to deliver a speed-independent quality of service to the mobile users.

Chapter 7

Conclusions and Future Research

In this chapter, we first summarize the main contributions of this thesis. We then draw some conclusions from the presented results and finally discuss some ideas in relation with the presented results that are worth being further investigated.

7.1 Main Contributions

Block Transmission Techniques: Overview

In this thesis, we first analyzed the achievable performance of different existing block transmission schemes in the WLAN context where the transmission channel is stable and unknown to the transmitter. The study was made under the hypothetical assumption that the transmission channel was perfectly known to the receiver. Single-carrier techniques (SC-CP and KSP) have a clear advantage over OFDM in terms of uncoded BER performance as they better exploit the delay diversity of the channel. When coded schemes are used and hard decoding is done after the decision device, the performance of OFDM is significantly increased as the use of the convolutional codes together with the interleaver allows to exploit the delay diversity as well. Coded OFDM even outperforms coded SC-CP and FD KSP equalizers for high-order constellations. However, *uncoded* single carrier block transmission still outperforms the coded OFDM schemes of equivalent data rates. Amongst the single carrier schemes, KSP has the same performance as SC-CP when FD equalizers are

used. In low-order constellations, this simple FD equalization scheme has the same performance as the more complex optimal linear KSP equalizers. However, when higher order constellations are considered, using the optimal linear KSP equalizers results in improved performance. Finally, the use of the more complex BDFE scheme for KSP significantly improves the performance. The KSP BDFE outperforms all other equalization (FD or Optimal) schemes for KSP as well as for other transmission schemes (OFDM or SC-CP) and this for all the considered coding rates and constellations.

This analysis reveals a high performance potential for the KSP scheme, especially when the BDFE is used. It shows as well that uncoded KSP transmission is an interesting modulation scheme, even when the less complex FD equalizers are used.

Channel Identification

In order to avoid IBI, both CP and KSP schemes introduce some overhead symbols that act as a guard band between consecutive blocks of data symbols. Besides enabling the use of the presented block equalization schemes, these overhead symbols can be exploited towards various goals by the receiver, including the estimation of the carrier frequency offset, the frame synchronization or the estimation of the transmission channel. In the presented research, we mainly focus on the channel estimation possibilities offered by these overhead symbols.

In CP systems (OFDM and SC-CP), the CP-induced cyclostationarity of the transmitted symbols is the only property that can be exploited towards channel estimation. However, this property is not very strong and the resulting channel models are not extremely accurate. Hence, most existing CP systems rely on long preambles acting as training sequences in order to estimate the channel before data transmission.

The situation is somewhat different in KSP where the padded sequences can be seen as short training sequences. A study of the Cramer Rao bound indicates that it is possible to estimate the channel accurately relying on the knowledge of the padded sequences, even for relatively small packet sizes. Several semi-blind methods that could be used in the KSP context have been proposed in the literature, but none of them appears to be fully tailored to the considered situation. We have thus proposed a novel Gaussian ML channel estimation scheme that is well suited to this KSP scheme. We derived closed-form expressions of the GML channel estimates that allow to asymptotically achieve the CRB. Besides its interesting performance, the proposed GML method has a very limited computational complexity. Identifying the channel relying on the knowledge of the padded sequences possibly allows to avoid the insertion of the bandwidth-consuming long preambles that are used towards channel

identification in classical transmission systems.

A New KSP Scheme

The CRB study indicates that there is a clear advantage in terms of achievable channel estimation accuracy when the padded sequences are changed from block to block. This requirement comes in contradiction with the possibility of using the low-complexity FD equalizers that offer attractive performance in low order constellations and still outperform OFDM in higher-order constellations. We proposed a new KSP scheme that is called Shifted Known Symbol Padding that allows to cope simultaneously with these two requirements. At the cost of an extra symbol in the padded sequence, S-KSP indeed offers the possibility of using the FD equalizers whilst the padded sequences are changed from block to block, which largely improved the accuracy of the channel estimates.

Implementability

Several implementational aspects of the different block transmission techniques were discussed as well. A discussion on the computational complexity indicates that the computational complexity of the equalization schemes that have the best performance, namely BDFE and optimal linear KSP equalizers, is significantly higher than that of the FD equalizers for OFDM, SC-CP or KSP. Though significant, this increase in computational complexity is not dramatic and next generation DSPs should enable the use of these equalization techniques at a reasonable cost. This study indicates as well that the complexity of the BDFE is only marginally larger than that of the optimal linear KSP equalizer, whilst the use of the BDFE results in significantly improved performance.

Some implementational issues only have an impact on the OFDM scheme, namely the peak to average power ratio and the problem of out of band noise. The PAPR problem causes a performance degradation for two reasons. First, the orthogonality between the carriers is occasionally lost (in the blocks on which clipping is done), secondly, the power of the transmitted signal is often backed off in order to reduce the probability of clipping, which results in a reduced effective SNR. The out of band noise problem decreases the effective data throughput as the outer tones of wireless OFDM systems are switched off in order to avoid the radiation of power outside the dedicated frequency band.

Two other problems impact all the FD equalization schemes for OFDM, SC-CP and KSP: the carrier frequency offset and the narrow-band interference. The first problem is more important in wireless applications as it destroys the orthogonality of the tones, impacting all the FD equalization schemes that assume ISI-free transmission and whose performance is significantly harmed

when uncompensated ICI arises.

The outcome of this discussion is that the BDFE and optimal linear KSP equalization schemes allow to avoid most practical implementation problems encountered by classical block transmission schemes. Their main drawback though remains their higher computational complexity.

Full System Performance

An experimental study using realistic channel estimates in a Hiperlan2 context allowed us to further assess the performance of the different schemes in practical systems. This study shows that the S-KSP scheme significantly outperforms the traditional KSP scheme thanks to the improved quality of the channel estimates. The S-KSP performance is close to the perfect channel knowledge situation as soon as the length of the considered packets approaches 50 blocks of data. The CP schemes suffer from the poor quality of the channel estimates that can be obtained using the preambles that are foreseen in the standards. Their final performance is several dBs away from the perfect channel knowledge situation. In this context, all KSP schemes, including the low-complexity FD KSP equalizer outperform OFDM, even in coded transmission.

Blind Direct Symbol Estimation

In the ZP KSP situation, the padded sequences are set to zero and the proposed channel estimation scheme cannot be applied. One has to rely on blind schemes in order to estimate the transmitted symbols. We have proposed a blind modified ML sequence estimation method that allows to approach the performance of the MMSE FD-KSP equalizer when low order constellations are used, and this even when no information is available on the transmission channel. Unlike existing ML methods, the proposed method has a low computational complexity as it exploits the diagonalization property of the circulant channel matrix.

High Mobility Scenarios

In the last part of this work, we briefly analyze the situation where mobile terminals are considered, which makes the transmission channels doubly selective. When the channel becomes time-varying, the orthogonality between the carriers is lost and all the FD equalization schemes suffer from the resulting ICI. In contrast, the BDFE and optimal linear KSP equalizers can be used readily without performance degradation provided the full channel model is known at

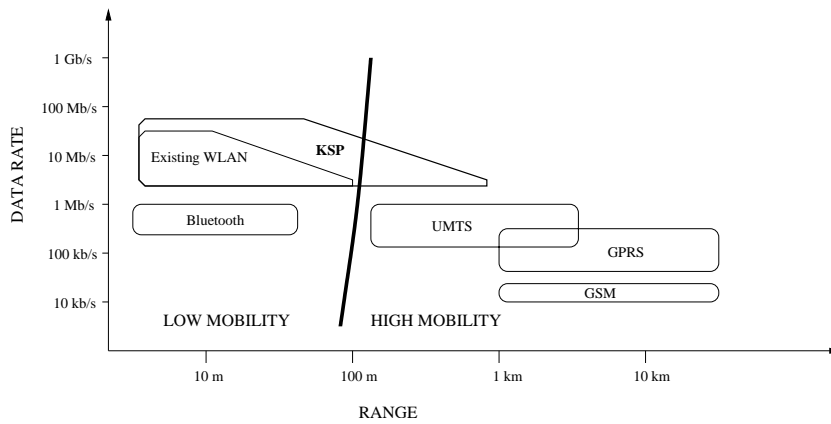


Figure 7.1: Datarate and range of different existing wireless communication systems and of the investigated KSP schemes.

the receiver. Moreover, it has been proved that the optimal placement of training for the identification of time-varying frequency-selective channels consists of equi-spaced pilot symbol sequences of constant length, which matches the KSP transmission scheme.

For these reasons, KSP appears to be a natural candidate for block transmission over doubly selective channels. We thus briefly analyze how KSP equalizers can be adapted to higher mobility scenarios and discuss several methods that allow to track the channel variations relying on the knowledge of the padded sequences. The most interesting approach is proposed in the basis expansion model framework, where the time-varying channel can be accurately described with a limited number of parameters.

Finally, experimental results show that a speed-independent performance can be obtained when coded transmission is used. The uncoded transmission scheme suffers from reduced channel estimation accuracy as the speed increases and a saturation appears in the BER curves when the mobility is too high. However, when the speed increases, the diversity of the channel increases as well. The use of coded transmission in combination with appropriate interleavers allows to exploit this extra diversity and to compensate for the uncoded performance degradation that results from the increased mobility.

7.2 Conclusions

Throughout this study, KSP appears to be an attractive transmission scheme both in the context of stationary and doubly selective channels. It has the implementational advantages of single-carrier applications such as reduced sensitivity to CFO or mobility or reduced dynamic range of the transmitted signals. The proposed S-KSP scheme allows to estimate the channel with a high accuracy giving an important advantage to KSP over CP schemes in stationary channels. Finally, the use of S-KSP allows a flexible receiver design. Depending on the used constellation, the quality of the radio link, the desired system performance and the available processing power, the receiver can choose between different possible equalization schemes of different complexities. When the processing power increases, more complex equalization schemes can be used yielding improved performance.

The computational complexity of the best performing schemes still limits their implementability, but as they are not immoderately complex, the natural evolution of the available processing power should make them suitable candidates in next generation terminals. The complexity of the channel estimation procedures could also be an asset for the immediate implementability of the discussed KSP schemes. Note also that these processing requirements have an impact on the power consumption of the terminal and that the energy savings resulting from the reduced SNR requirements are partly cancelled by this issue. An important asset is the performance of uncoded KSP transmission that shows a higher efficiency than its coded counterparts in high SNR conditions, possibly allowing to avoid the use of resource and bandwidth consuming convolutional codes. Finally, it is possible to handle doubly selective channels in KSP without changing the format of the transmitted data and using similar equalizer structures. The transition between low and high mobility scenarios is thus smooth as the same base stations and terminals can be used in both cases.

In **Fig. 7.1**, we compare the datarates and ranges of existing and upcoming wireless communication schemes with those that would be achievable if KSP transmission was considered. The presented experimental results allow us to re-assess the performance of current WLAN standards (see **Fig. 1.1**), as they show that the highest data rates can only be achieved in a limited range corresponding to high SNR conditions. In the WLAN context, KSP allows a wider range of data rates as uncoded transmission offers attractive performance, which is not the case for current OFDM-based standards. Besides this, the range within which a receiver can be used is largely increased as the SNR requirements are lower in KSP .

The use of KSP in higher mobility scenarios allows to offer a speed-independent quality of service when the coded schemes are considered. Compared to 3G mobile applications (UMTS), the investigated KSP scheme offers a higher through-

put but the range over which the communication can be done appears to be smaller than in 3G applications.

7.3 Open Issues and Future Research

Optimal Training Sequences

The CRB study presented in *Ch. 4* indicates that the expected MSE of the channel estimate depends on the composition of the padded sequences. It would be worth investigating under which conditions the padded sequences allow to minimize the CRB. The use of padded sequences fulfilling that condition would result in an improved channel estimation accuracy.

MIMO Setups

The use of antenna arrays at the transmitter and/or receiver allows to significantly increase the capacity of the transmission channel. In order to offer datarates that truly compete with those of wired systems whilst operating in a limited bandwidth, future wireless setups will probably rely on MIMO channels. Space-time coding techniques allowing to exploit the increased capacity of MIMO channels have been developed mainly for flat-fading channels.

When frequency-selective channels are considered, these space-time coding schemes are readily used on a per tone basis in OFDM transmission. The extension of existing space-time coding schemes to KSP or single-carrier transmission is less straightforward and is an open area of research [124], [125].

Inter Block Interference

In the presented research, we have generally considered the situation where the channel order is smaller than the length of the padded sequence or CP, resulting in an IBI-free situation. However, IBI cannot be excluded in practical systems where the channel order can be larger than foreseen in the initial system design. and ad-hoc schemes must be developed to cope with the resulting IBI.

When IBI arises, the carrier orthogonality is lost and ad-hoc schemes must be developed to avoid a significant performance degradation. Several approaches are possible towards that goal, including the channel-shortening time-domain equalizers (used in ADSL), the per tone equalizers [126], or the use of explicit IBI cancellation schemes that remove the interference from the previous block into the next block.

It would be worth analyzing the impact of these different approaches on the final performance of KSP schemes operating in channels characterized by a large delay spread.

Code Optimization, Joint Equalization and Decoding

In the presented results, channel equalization and error correction are two separated steps. However, channels with known impulse responses can be seen as convolutional encoders as well and a unified approach that jointly equalizes the channel and decodes the encoded stream of data symbols would probably yield improved results. Recently introduced turbo equalization techniques [127], [128] currently begin to explore that approach..

Besides, as discussed in the text, the codes that are foreseen in the current standards are designed for OFDM system. Designing codes that are specifically suited to the KSP would significantly improve the performance of the KSP coded schemes.

In static channel conditions, the OFDM scheme has an important advantage as the soft output of the zero forcing equalizer is also the ML and MAP symbol estimate. Hence, performing sequential ZF equalization and soft error decoding is truly optimal, which is not the case in KSP transmission. The situation differs somehow when time-varying channels are considered as OFDM suffers from ICI and the true ML or MAP solution cannot be obtained with the ZF equalizer.

Improved Equalization Schemes for Doubly Selective Channels

In higher mobility scenarios, we mainly focused on the problem of estimating the time-varying channel relying on the knowledge of the padded sequences. The considered equalization schemes are straightforward extensions of the existing KSP equalizers. More advanced equalization schemes that explicitly exploit the BEM model or the time-varying nature of the transmission channel should yield better performance than the low-complexity FD scheme, possibly similar to the optimal linear KSP equalizer, but with a significantly reduced computational complexity.

Bibliography

- [1] J.A.C. Bingham, “Multicarrier Modulation for Data Transmission: An Idea whose Time has Come,” *IEEE Communications Magazine*, pp. 5–14, May 1990.
- [2] Z. Wang and G.B. Giannakis, “Wireless Multicarrier Communications: Where Fourier meets Shannon,” *IEEE Signal Processing Magazine*, pp. 29–48, May 2000.
- [3] ETSI, *Broadband Radio Access Networks (BRAN); HIPERLAN Type 2 technical specifications Part 1 - Physical Layer*, dts/bran030003-1 edition, Oct 1999.
- [4] R.F.H. Fischer and J.B. Huber, “A New Loading Algorithm for Discrete Multitone Transmission,” in *Proc. of Globecom*, London, England, Nov. 1996, pp. 724–728.
- [5] H. Sari, G. Karam and I. Jeanclaude, “Transmission Techniques for Digital Terrestrial TV Broadcasting,” *IEEE Communications Magazine*, pp. 100–109, Feb. 1995.
- [6] R. Cendrillon, M. Moonen, “Efficient equalizers for single and multi-carrier environments with known symbol padding,” in *Proc. of the Sixth International Symposium on Signal Processing and its Applications (ISSPA 2001)*, Kuala-Lumpur, Malaysia, Aug. 2001, vol. 2, pp. 607–610.
- [7] G. Leus and M. Moonen, “Semi-Blind Channel Estimation for Block Transmission with Non-Zero Padding,” in *Proc. of the Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, California, Nov. 4-7 2001, vol. 1, pp. 762–766.
- [8] L. Deneire, B. Gyselinckx and M. Engels., “Training Sequence vs. Cyclic Prefix: A New Look on Single Carrier Communication,” *IEEE Communication Letters*, vol. 5, no. 7, pp. 292–294, 2001.

- [9] Z. Wang and G. B. Giannakis, "Linearly precoded or coded OFDM against wireless channel fades?," in *Proc. of IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2001)*, Taoyuan, Taiwan R.O.C., March 2001, pp. 267–270.
- [10] P. Stoica and O. Besson, "Training Sequence Design for Frequency Offset and Frequency-Selective Channel Estimation," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1910–1917, March 2003.
- [11] P. Ciblat and L. Vandendorpe, "On the Maximum-Likelihood Based Data-Aided Joint Frequency Offset and Channel Estimates," in *Proc. of EUSIPCO*, Toulouse, France, September 2002, pp. 627–630.
- [12] O. Rousseaux, G. Leus, P. Stoica and M. Moonen, "Generalized Training Based Channel Identification," in *Global Conference on Communications (GLOBECOM 2003)*, San Francisco, California, December 2003, pp. 2432–2436.
- [13] M. Dong, L. Tong and B. M. Sadler, "Optimal Pilot Placement for Semi-Blind Channel Tracking of Packetized Transmission over Time-Varying Channels," *IEEE Transactions on Signal Processing*, 2004, to appear.
- [14] J. H. Manton and Y. Hua, "Affine Precoders for Reliable Communications," in *Proc. of ICASSP*, Istanbul, Turkey, June 2000.
- [15] J. Medbo, H. Hallenberg and J. E. Berg, "Propagation Characteristics at 5 GHz in Typical Radio-LAN Scenarios," in *Proc. of the Vehicular Technology Conference (VTC 99)*, Spring (Houston), 99, pp. 185–189.
- [16] Edward A. Lee and David G. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, 2 edition, 2000.
- [17] G. Leus and M. Moonen, *Handbook on Signal Processing for Communications*, chapter Equalization Techniques for Fading Channels, CRC Press, 2004.
- [18] A. Stamoulis, G. B. Giannakis and A. Scaglione, "Block FIR Decision-Feedback Equalizers for Filterbank Precoded Transmissions with Blind Channel Estimation Capabilities," *IEEE Transactions on Communications*, vol. 49, pp. 69–83, Jan. 2001.
- [19] A. Ginesi, G. M. Vitetta and D. D. Falconer, "Block Channel Equalization in the Presence of a Cochannel Interferent Signal," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1853–1862, 1999.
- [20] A. Duel-Hallen,, "A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple-Access Channels," *IEEE Transactions on Communications*, vol. 43, pp. 421–434, Feb-March-Apr 1995.

- [21] A. Klein, G. Kanas Kaleh and P. W. Baier, "Zero forcing and minimum mean-square-error equalization for multiple-access channels," *IEEE Transactions on Vehicular Technology*, vol. 45, pp. 276–287, May 1996.
- [22] L. Goldfeld, V. Lyandres and D. Wulich, "Minimum BER power loading for OFDM in fading channel," *IEEE Transactions on Communications*, vol. 50, no. 11, pp. 1729–1733, Nov 2002.
- [23] J. Louveaux, L. Vandendorpe and T. Sartenauer, "Cyclic Prefixed Single Carrier and Multicarrier Transmission: Bit Rate Comparison," *IEEE Communications Letters*, vol. 7, no. 4, pp. 180–182, Apr 2003.
- [24] Y. P. Lin and S. M. Phoong, "MMSE OFDM and Prefixed Single Carrier Systems: BER Analysis," in *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Hong Kong, Apr 2003, vol. IV, pp. 229–232.
- [25] A. Scaglione, G. B. Giannakis and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers Part I: Unification and Optimal Designs," *IEEE Transactions on Signal Processing*, vol. 47, no. 7, pp. 1988–2006, July 1999.
- [26] A. Scaglione, G. B. Giannakis and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers Part II: Blind Channel Estimation, Synchronization and Direct Equalization," *IEEE Transactions on Signal Processing*, vol. 47, no. 7, pp. 2007–2022, July 1999.
- [27] Y. Ding, T. N. Davidson and Z. Q. Luo, "Minimum BER Block Precoders for Zero-Forcing Equalization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2410–2423, Sept 2003.
- [28] Z. Wang, X. Ma and G. B. Giannakis, "OFDM or Single-Carrier Block Transmissions?," *IEEE Transactions on Communications*, vol. 52, no. 3, pp. 380–394, March 2004.
- [29] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*, Prentice Hall, 3rd edition, 1996.
- [30] R. van Nee and A. de Wilde, "Reducing the Peak to Average Power Ratio of OFDM," in *Proc. of IEEE Vehicular Technology Conference (VTC 1998)*, 1998, pp. 2072–2076.
- [31] H-G. Ryu, B-l. Jin and I-B. Kim, "PAPR Reduction using Soft Clipping and ACI Rejection in OFDM Systems," *IEEE Transactions on Consumer Electronics*, vol. 48, pp. 17–22, Feb 2002.
- [32] H-G. Ryu, J-E. Lee and J-S. Park, "Dummy Sequence Insertion (DSI) for PAPR Reduction in OFDM Communication Systems," *IEEE Transactions on Consumer Electronics*, vol. 50, pp. 89–94, Feb 2004.

- [33] L. J. Cimini and N. R. Sollenberger, "Peak to Average Power Ratio Reduction of an OFDM Signal using Partial Transmit Sequences," *IEEE Communications Letters*, vol. 4, pp. 86–88, March 2000.
- [34] A. E. Jones, T. A. Wilkinson and S. K. Barton, "Block Coding Scheme for Reduction of Peak to Mean Envelope Power Ratio of Multicarrier Transmission Schemes," *Electronic Letters*, vol. 30, no. 25, pp. 2098–2099, Dec 1994.
- [35] K. G. Paterson, "Generalized Reed-Muller Codes and Power Control in OFDM modulation," *IEEE Transactions on Information Theory*, vol. 46, pp. 104–120, Jan 2000.
- [36] P. Van Eetvelt, G. Wade and M. Tomlinson, "Peak to Average Power Reduction for OFDM Schemes by Selective Scrambling," *Electronic Letters*, vol. 32, pp. 1963–1964, Oct 1996.
- [37] A. R. S. Bahai, M. Singh, A. J. Goldsmith and B. R. Saltzberg, "A New Approach for Evaluating Clipping Distortion in Multicarrier," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1037–1046, May 2002.
- [38] R. Dinis and A. Gusamo, "On the Performance Evaluation of OFDM Transmission Using Clipping," in *Proc. of the Vehicular Technology Conference (VTC 99)*, 1999, vol. 5, pp. 2923–2928.
- [39] N. Dinur and T. Wulich, "Peak to Average Power Ratio in Amplitude Clipped High-Order OFDM," in *Proc. of MILCOM 98*, 1998, vol. 2, pp. 684–687.
- [40] M. Friese, "On the Achievable Information Rate with Peak Power Limited OFDM," *IEEE Transactions on Information Theory*, vol. 46, pp. 2579–2587, July 2000.
- [41] X. Li and L. J. Cimini, "Effects of Clipping and Filtering on the Performance of OFDM," *IEEE Communications Letters*, vol. 2, pp. 131–133, May 1998.
- [42] A. L. Kambanellas and G. B. Giannakis, "Modulo Pre-equalization of Nonlinear Communication Channels," in *Proc. of Signal Processing Advances in Wireless Communications (SPAWC99)*, Annapolis, May 1999, pp. 46–49.
- [43] T. Pollet and M. Moeneclaey, "BER Sensitivity of OFDM Systems to Carrier Frequency Offset and Wiener Phase Noise," *IEEE Transactions on Communications*, pp. 191–193, Feb.- Apr. 1995.
- [44] Jihoon Choi, Changoo Lee, Hae Won Jung and Yong Hoon Lee, "Carrier frequency offset compensation for uplink of OFDM-FDMA systems," *IEEE Communication Letters*, vol. 4, no. 12, pp. 414–416, Dec 2000.

- [45] J-C. Lin, "Maximum-likelihood frame timing instant and frequency offset estimation for OFDM communication over a fast Rayleigh-fading channel," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 4, pp. 1049–1062, July 2003.
- [46] U. Tureli, H. Liu and M. D. Zoltowski, "OFDM blind carrier offset estimation: ESPRIT," *IEEE Transactions on Communications*, vol. 48, no. 9, pp. 1459–1471, Sept 2000.
- [47] X. Ma, G. B. Giannakis and S. Barbarossa, "Non-data-aided frequency-offset and channel estimation in OFDM: and related block transmissions," in *Proc. of the international Conference on Communications (ICC01)*, Helsinki, June 2001, vol. 6, pp. 1866–1870.
- [48] H. Bolcskei, "Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems," *IEEE Transactions on Communications*, vol. 49, no. 6, pp. 988–999, June 2001.
- [49] Q. Shenping, Y. Changchuan, L. Jianfeng and Y. Guangxin, "Pilot-symbol-aided frequency offset estimation and correction for OFDM system," in *Proc. of Personal, Indoor and Mobile Radio Communications, (PIMRC 2003)*, Sept 2003, pp. 593–596.
- [50] X. Ma, H. Kobayashi and S. C. Schwartz, "Joint frequency offset and channel estimation for OFDM," in *Proc. of the Global Telecommunications Conference (GLOBECOM 2003)*, Hong Kong, Dec 2003, vol. 1, pp. 15–19.
- [51] Z. Zhang and M. Zhao, "Joint frame synchronization and frequency offset estimation in OFDM system," in *Proc. of the Global Telecommunications Conference (GLOBECOM 2003)*, Sept 2003, vol. 3, pp. 2249–2253.
- [52] J. Li, G. Liu and G. B. Giannakis, "Carrier frequency offset estimation for OFDM-based WLANs," *IEEE Signal Processing Letters*, vol. 8, no. 3, pp. 80–82, March 2001.
- [53] S. Kapoor and S. Nedic, "Interference Suppression in DMT Receivers using Windowing," in *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2000)*, 2000, pp. 778–782.
- [54] A. Redfern, "Receiver Window Design for Multicarrier Communication Systems," *IEEE Transactions on Selected Areas in Communications*, vol. 20, no. 5, pp. 1029–1036, June 2002.
- [55] G. Cuypers, K. Vanbleu, G. Ysebaert and M. Moonen, "Combined per Tone Equalization and Receiver Windowing in DSL Receivers: WiPTEQ," in *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 02)*, Orlando, 2002, vol. 3, pp. 2341–2344.

- [56] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels (mobile radio)," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, November 1991.
- [57] H. Vikalo, B. Hassibi, B. Hochwald, T. Kailath, "Optimal Training for Frequency-Selective Fading Channels," in *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Salt Lake City, Utah, May 2001, pp. 2105–2108.
- [58] C. Fragouli, N. Al-Dhahir and W. Turin, "Training-Based Channel Estimation for Multiple-Antenna Broadband Transmissions," *IEEE Transactions on Wireless Communications*, vol. 2, no. 2, pp. 384–391, March 2003.
- [59] Jonathan. H. Manton, "On Optimal Channel Identification by Use of Training Sequences and Pilot Tones," in *Proc. of the sixth International Symposium on Signal Processing and its Applications (ISSPA 2001)*, Kuala-Lumpur, Malaysia, 2001, pp. 599 – 602.
- [60] Lang Tong and Sylvie Perreau, "Multichannel Blind Identification: from Subspace to Maximum Likelihood Methods," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1951–1968, Oct. 1998.
- [61] D. Gesbert, P. Duhamel and S. Mayrargue, "Blind Least-Squares Approaches for Joint Data/Channel Estimation," in *Proc. of the IEEE Digital Signal Processing Workshop*. IEEE, 1-4 Sept 1996, pp. 450–453.
- [62] S. Schell and J. Bapat, "Partially Blind Identification of FIR Channels for QAM Signals," in *Proc. of Military Communications Conference*, San Diego, CA, 1995, vol. 2, pp. 592–596.
- [63] Buchoux, V.; Cappe, O.; Moulines, E.; Gorokhov, A., "On the performance of semi-blind subspace-based channel estimation," *IEEE Transactions on Signal Processing*, vol. 48, no. 6, pp. 1750–1759, June 2000.
- [64] de Carvalho, E.; Slock, D., "Deterministic quadratic semi-blind FIR multichannel estimation algorithms and performance," in *Proc. of the 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2000)*, June 5-9 2000, vol. 5, pp. 2553–2556.
- [65] A. Medles, D. T. M. Slock and E. De Carvalho, "Linear Prediction Based Semi-Blind Estimation of MIMO FIR Channels," in *Proc. of Signal Processing Advances In Wireless Communications (SPAWC2001)*, Taiwan ROC, March 2001, pp. 58–61.
- [66] J. Ayadi, E. de Carvalho and D. T. M. Slock, "Blind and Semi-Blind Maximum Likelihood Methods for FIR Multichannel Identification," in *Proc. of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP 98)*, Seattle, May 1998, pp. 3185–3188.

- [67] E. de Carvalho and D. T. M. Slock, "Cramer-Rao Bounds for Semi-Blind, Blind and Training Sequence Based Channel Estimation," in *Proc. of the first IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 97)*, Paris, France, April 1997, pp. 129–132.
- [68] H. Trigui and D. T. M. Slock, "Optimal and Suboptimal Approaches for Training Sequence Based Spatio-Temporal Channel Identification in Coloured Noise," in *Proc. of the 31st ASILOMAR Conference on Signal, Systems and Computers*, Pacific Grove, California, November 1997, pp. 1038–1042.
- [69] E. de Carvalho and D. T. M. Slock, "Maximum Likelihood Blind FIR Multi-Channel Estimation with Gaussian Prior for the Symbols," in *Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP 97)*, Munich, Germany, April 1997, pp. 3593–3596.
- [70] Jyotsna L. Bapat, "Partially Blind Estimation: ML-Based Approaches and Cramer-Rao Bound," *Signal Processing (Elsevier Science)*, vol. 71, pp. 265–277, 1998.
- [71] H. A. Cipran and M. K. Tsatsanis, "Stochastic Maximum Likelihood Methods for Semi-Blind Channel Estimation," *IEEE Signal Processing Letters*, vol. 5, no. 1, pp. 21–24, January 1998.
- [72] T. Söderström, P. Stoica, *System Identification*, International Series in Systems and Control Engineering. Prentice Hall, 1989.
- [73] O. Rousseaux, G. Leus, M. Moonen, "A suboptimal Iterative Method for Maximum Likelihood Sequence Estimation in a Multipath Context," *EURASIP Journal on Applied Signal Processing (JASP)*, vol. 2002, no. 12, pp. 1437–1447, December 2002.
- [74] S. Talwar, M. Viberg and A. Paulraj, "Blind Separation of Synchronous Co-Channel Digital Signals Using an Antenna Array - Part I: Algorithms," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1184–1197, May 1996.
- [75] L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*, Addison-Wesley, Reading, 1991.
- [76] A. Dogandzic and A. Nehorai, "Space-Time Fading Channel Estimation and Symbol Detection in Unknown Spatially Correlated Noise," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 457–473, March 2002.
- [77] G. B. Giannakis, Y. Hua, P. Stoica and L. Tong, *Signal Processing Advances in Wireless and Mobile Communications*, vol. 1: Trends in Channel Estimation and Equalization, Prentice Hall PTR, 2000.

- [78] R. W. Heath Jr and G. B. Giannakis, "Exploiting Input Cyclostationarity for Blind Channel Identification in OFDM Systems," *IEEE Transactions on Signal Processing*, vol. 47, no. 3, pp. 848–856, March 1999.
- [79] U. Tureli and H. Liu, "Blind Carrier Synchronisation and Channel Identification for OFDM Communications," in *Proc. of the ICASSP Conference*, Seattle, Washington, 1998, vol. 6, pp. 3509–3512.
- [80] Edfors, O. Sandell, M. van de Beek, J.-J. Wilson, S.K. Borjesson, P.O., "OFDM channel estimation by singular value decomposition," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 931–939, July 1998.
- [81] Z. Cheng and D. Dahlhaus, "Time versus Frequency Domain Channel Estimation for OFDM Systems with Antenna Arrays," in *Proc. of the 6th International Conference on Signal Processing (ICSP'02)*, Beijing, China, Aug. 2002, vol. 2, pp. 1340–1343.
- [82] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 4, pp. 1207–1215, July 2000.
- [83] M. Sandell and O. Edfors, "A Comparative Study of Pilot-Based Channel Estimators for Wireless OFDM," Tech. Rep., Lulea University of Technology, Lulea, Sweden, 1996.
- [84] C. S. Yeh and Y. Lin, "Channel estimation using pilot tones in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 45, no. 4, pp. 400–409, Dec 1999.
- [85] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3065–3073, Dec 2001.
- [86] B. Muquet, M. de Courville, P. Duhamel, and V. Buzenac, "A Subspace-Based Blind and Semiblind Channel Identification Method for OFDM Systems," in *Proc. of Signal Processing Advances in Wireless Communications (SPAWC 99)*, Annapolis, Maryland, May 1999, pp. 170–173.
- [87] S. Barbarossa, A. Scaglione and G. B. Giannakis, "Performance Analysis of a Deterministic Channel Estimator for Block Transmission Systems with Null Guard Intervals," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 684–695, March 2002.
- [88] O. Rousseaux, G. Leus, P. Stoica and M. Moonen, "A Stochastic Method for Training Based Channel Identification," in *Seventh International Symposium on Signal Processing and its Applications (ISSPA 2003)*, Paris, France, July 2003, pp. 657–660.

- [89] Y. Song, S. Roy and L. A. Akers, "Joint Blind Estimation of Channel and Data Symbols in OFDM," in *Proc. of the 51st Vehicular Technology Conference Proceedings*, Tokyo, Japan, 2000, vol. 1, pp. 46–50.
- [90] B. H. Juang, L. R. Rabiner, "The Segmental K-Means Algorithm for Estimating Parameters of Hidden Markov Models," *IEEE transactions on Acoustics, Speech and Signal Processing*, vol. 38(9), pp. 1639–1641, Sept. 1990.
- [91] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm," *Journal of Royal Statistical society B*, vol. 39, pp. 1–38, 1977.
- [92] N. Seshadri, "Joint Data and Channel Estimation Using Blind Trellis Search Techniques," *IEEE Transaction on Communications*, vol. 42, no. 2/3/4, pp. 1000–1011, April 1994.
- [93] S. Talwar and A. Paulraj, "Blind Estimation of Multiple Co-Channel Digital Signals Received at an Antenna Array," in *proc. of the fifth Annual IEEE Dual-Use Technologies and Applications Conference*, May 1995.
- [94] O. Rousseaux, G. Leus and M. Moonen, "A Blind Multi-User MIMO Transceiver Using Code Modulation in a Multipath Context," in *Proc. of the 14th Conference on Digital Signal Processing 2002 (DSP 2002)*, Santorini, Greece, July 2002, pp. 267–270.
- [95] O. Rousseaux, G. Leus and M. Moonen, "A Reduced Complexity Deterministic Blind Transceiver with Space- Only Block Coding in a Multi-User MIMO Context with Severe Multipath," in *Proc. of the Benelux Meeting on Systems and Controls*, Houffalize, Belgium, March 2001.
- [96] O. Rousseaux, G. Leus and M. Moonen, "A Blind Receiver for Block Transmission in a Multi-User MIMO Context with Multipath," in *Proc. of the Signal Processing Symposium (SPS2002)*, Leuven, Belgium., Mar 2002, pp. pp. 33–36.
- [97] John. G. Proakis, *Digital Communications*, Mc Graw-Hill International Edition, 4 edition, 2000.
- [98] X. Ma, G. B. Giannakis and S. Ohno, "Optimal Training for Block Transmission Over Doubly Selective Wireless Channels," *IEEE Transaction on Signal Processing*, vol. 51, no. 5, pp. 1351–1366, May 2003.
- [99] M. K. Tsatsanis and G. B. Giannakis, "Modeling and Equalization of Rapidly Fading Channels," *International Journal of Adaptive Control and Signal Processing*, vol. 10, pp. 159–176, 1996.

- [100] G. B. Giannakis and C. Tepedelenlioglu, "Basis Expansion Model and Diversity Techniques for Blind Identification and Equalization of Time Varying Channels," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct 1998.
- [101] A. M. Sayeed and B. Aazhang, "Joint Multipath-Doppler Diversity in Mobile Wireless Communications," *IEEE Trans. on Communications*, vol. 47, pp. 123–132, Jan 1999.
- [102] I. Barhumi, G. Leus and M. Moonen, "Time-Varying FIR Equalization of Doubly Selective Channels," *IEEE Transactions on Wireless Communications*, 2004, to appear.
- [103] G. Leus, I. Barhumi and M. Moonen, "MMSE Time-Varying FIR Equalization of Doubly-Selective Channels," in *Proc. of the International Conference on Acoustics, Speech, and Signal Processing (ICASP 2003)*, Hong Kong, April 2003, pp. IV-485–IV-488.
- [104] P. Schniter, "Low-Complexity Estimation of Doubly-Selective Channels," in *proc. of Signal Processing Advances in Wireless Communications*, Rome, Italy, June 2003, pp. 200–204.
- [105] W. C. Jakes, *Microwave Mobile Communications*, chapter 3, Wiley, 1974.
- [106] M. Niedzwiecki and T. Klaput, "Fast Recursive Basis Function Estimators For Identification of Time-Varying Processes," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1925–1934, Aug 2002.
- [107] M. Martone, "Multiresolution Sequence Detection in Rapidly Fading Channels Based on Focused Wavelet Decompositions," *IEEE Transactions on Communications*, vol. 49, no. 8, pp. 1388–1401, Aug 2001.
- [108] K. Higuchi, H. Andoh, K. Okawa, M. Sawahashi and F. Adachi, "Experimental Evaluation of Combined Effect of Coherent RAKE Combining and SIR-Based Fast Transmit Power Control for Reverse Link of DS-CDMA Mobile Radio," *IEEE Transactions on Selected Areas in Communications*, vol. 18, no. 8, pp. 1526–1535, Aug 2000.
- [109] M. Martone, "Wavelet-Based Separating Kernels for Array Processing of Cellular DS/CDMA Signals in Fast Fading," *IEEE Transactions on Communications*, vol. 48, no. 6, pp. 979–995, June 2000.
- [110] D. K. Borah and B. D. Hart, "Frequency-Selective Fading Channel Estimation with a Polynomial Time-Varying Channel Model," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 862–873, June 1999.
- [111] G. Yue, X. Zhou and X. Wang, "Performance Comparisons of Channel Estimation Techniques in Multipath Fading CDMA," *IEEE Transactions on Wireless Communications*, vol. 3, no. 3, pp. 716–724, 2004.

- [112] Y. V. Zakharov, T. C. Tozer and J. F. Adlard, "Polynomial Spline-Approximation of Clarke's Mode," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1198–1208, May 2004.
- [113] M. Hill and K. J. R. Liu, "Channel Estimation for multicarrier modulation systems using a time-frequency polynomial model," *IEEE Transactions on Communications*, vol. 50, no. 7, pp. 1045–1048, Jul 2002.
- [114] I. Barhumi, G. Leus and M. Moonen, "Time-Varying FIR Decision Feedback Equalization of Doubly Selective Channels," in *Proc. of the Global Telecommunications Conference (GLOBECOM 2003)*, San Francisco, CA, Dec 2003, pp. 2263–2268.
- [115] G. Leus, O. Rousseeaux and M. Moonen, "Direct Semi-Blind Design of Serial Linear Equalizers for Doubly-Selective Channels," in *Proc. of the International Conference on Communications (ICC 2004)*, Paris, France, June 2004, pp. 2626–2630.
- [116] X. Cai and G. B. Giannakis, "Bounding Performance and Suppressing Intercarrier Interference in Wireless Mobile OFDM," *IEEE Transactions on Communications*, vol. 51, no. 12, pp. 2047–2056, Dec 2003.
- [117] D. K. Borah and B. D. Hart, "A Robust Receiver Structure for Time-Varying, Frequency-Flat, Rayleigh Fading Channels," *IEEE Transactions on Communications*, vol. 47, no. 3, pp. 360–364, March 1999.
- [118] B. D. Hart, D. K. Borah and S. Pasupathy, "Autocovariance Preserving Estimator (APE) Interpretation of the MLSD Metric for Rayleigh Fading Channels," *IEEE Transactions on Communications*, vol. 48, no. 10, pp. 1614–1617, Oct 2000.
- [119] D. K. Borah and B. D. Hart, "Receiver Structures for Time-Varying Frequency-Selective Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1863–1875, Nov 1999.
- [120] B. D. Hart and S. Pasupathy, "Innovations-Based MAP Detection for Time-Varying Frequency-Selective Channels," *IEEE Transactions on Communications*, vol. 48, no. 9, pp. 1507–1519, Sept 2000.
- [121] S. Song, J. S. Lim, S. J. Baek and K. M. Sung, "Variable Forgetting Factor Linear Least Squares Algorithm for Frequency Selective Fading Channel Estimation," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 3, pp. 613–616, May 2002.
- [122] J.K. Tugnait and W. Luo, "Blind Identification of Time-Varying Channels Using Multistep Linear Predictors," *IEEE Transactions on Signal Processing*, vol. 52, no. 6, pp. 1739–1749, June 2004.

- [123] X. Ma and G. B. Giannakis, "Maximum-Diversity Transmissions over Doubly Selective Wireless Channels," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1832–1840, July 2003.
- [124] Shengli Zhou; Giannakis, G.B., "Single-carrier space-time block-coded transmissions over frequency-selective fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 164–179, Jan 2003.
- [125] Sampath, H.; Stoica, P.; Paulraj, A., "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2198–2206, Dec 2001.
- [126] Van Acker, K.; Leus, G.; Moonen, M.; van de Wiel, O.; Pollet, T., "Per tone equalization for DMT-based systems," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 109–119, Jan 2001.
- [127] Tuchler, M.; Koetter, R.; Singer, A.C., "Turbo equalization: principles and new results," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 754–767, May 2002.
- [128] Koetter, R.; Singer, A.C.; Tuchler, M., "Turbo equalization," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 860–864, Jan 2004.

List of Publications

Journal Papers

- O. Rousseaux, G. Leus, M. Moonen, “A suboptimal Iterative Method for Maximum Likelihood Sequence Estimation in a Multipath Context,” in *EURASIP Journal on Applied Signal Processing (JASP)*, vol. 2002, no. 12, pp. 1437-1447, Dec. 2002.
- O. Rousseaux, G. Leus, P. Stoica and M. Moonen, “Gaussian Maximum Likelihood Sequence Estimation with Short Training Sequences,” to appear in *IEEE Transactions on Wireless Communications*, 2004.
- O. Rousseaux, G. Leus and M. Moonen, “Estimation and Equalization of Doubly Selective Channels using Known Symbol Padding,” to appear in *IEEE Transactions on Signal Processing*, 2004.
- O. Rousseaux, G. Leus and M. Moonen, “Block Transmission and Shifted Known Symbol Padding for Efficient Data Communication in a WLAN Context”, in preparation, 2004.

Conference Papers

- O. Rousseaux, G. Leus, M. Moonen, “An Iterative Procedure for Semi-Blind Symbol Estimation in a Multipath SISO Channel Context Exploiting Finite Alphabet Properties,” in *Proc. of the International Zurich Seminar on Broadband Communications (IZS 2002)*, Feb 2002, Zurich, Switzerland, pp. 21.1-21.5.
- O. Rousseaux, G. Leus, P. Stoica and M. Moonen, “A Stochastic Method for Training Based Channel Identification,” in *Proc. of the Seventh International Symposium on Signal Processing and its Applications (ISSPA 2003)*, July 2003, Paris, France, pp. 657-660.
- O. Rousseaux, G. Leus, P. Stoica and M. Moonen, “Training Based Maximum Likelihood Channel Identification: Constant Training Sequences,”

in *Proc. of the IEEE Symposium on Signal Processing Advances in Wireless Communications (SPAWC 2003)*, June 2003, Rome, Italy, pp. 334-338.

- O. Rousseaux, G. Leus, P. Stoica and M. Moonen, “Generalized Training Based Channel Identification,” in *Proc. of the Global Conference on Communications (GLOBECOM 2003)*, December 2003, San Francisco, California, pp. 2432-2436.
- O. Rousseaux, G. Leus and M. Moonen, “An Iterative Method for Improved Training-Based Estimation of Doubly Selective Channels,” in *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2004)*, May 2004, Montreal, Canada, pp. iv.889-iv.892.
- O. Rousseaux, G. Leus and M. Moonen, “A Blind Receiver for Block Transmission in a Multi-User MIMO Context with Multipath,” in *Proc. of the Signal Processing Symposium (SPS2002)*, Mar. 2002, Leuven, Belgium, pp. 33-36.
- O. Rousseaux, G. Leus and M. Moonen, “A Reduced Complexity Deterministic Blind Transceiver with Space- Only Block Coding in a Multi-User MIMO Context with Severe Multipath,” in *Proc. of the Benelux Meeting on Systems and Controls*, March 2001, Houffalize, Belgium.
- O. Rousseaux, G. Leus and M. Moonen, “A Blind Multi-User MIMO Transceiver Using Code Modulation in a Multipath Context,” in *Proc. of the 14th Conference on Digital Signal Processing 2002 (DSP 2002)*, July 2002, Dantorini, Greece, pp. 267-270.
- G. Leus, I. Barhumi, O. Rousseaux and M. Moonen, “Direct Semi-Blind Design of Serial Linear Equalizers for Doubly-Selective Channels,” in *Proc. of the International Conference on Communications (ICC 2004)*, June 2004, Paris, France, pp. 2626-2630.
- R. Cendrillon, O. Rousseaux, M. Moonen, E. Van den Bogaert, J. Verlinden, “Simplified Power Allocation for the DSL Multi-Access Channel through Column-Wise Diagonal Dominance,” in *Proc. of 24th Symposium on Information Theory in the Benelux*, May 2003, Netherlands.

Biography

Olivier Rousseaux was born in Ottignies Louvain-la-Neuve on the 6th of October 1975.

He received his master of science in electrical and mechanical engineering from the Universite Catholique de Louvain-la Neuve (UCL), Belgium in 1998. His Master Thesis was in the field of Automatic Control Theory and was realized with the Automatic Control Department of the University of Linkoping, Sweden.

After working for Alcatel from Oct. 1998 to Aug. 1999, he joined the Electrical Engineering (ESAT) department of the Katholieke Universiteit Leuven (KUL) where he prepared this PhD under the supervision of Marc Moonen.

He joined the Interuniversity Micro-Electronics Center (IMEC) in November 2004 where he is since then working as a project officer in the Design Technology for Integrated Information and Communication Systems (DESICS) division.