

ANALYSIS AND DESIGN OF OFDM SYSTEMS: LOADING AND NONLINEAR  
DISTORTION MITIGATION

by

Mohammad Mohammadnia-Avval

M.Sc., Sharif University of Technology, Tehran, Iran, 2005

B.Sc., Amirkabir University of Technology, Tehran, Iran, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA  
(Vancouver)

September 2011

© Mohammad Mohammadnia-Avval, 2011

# Abstract

In wireless communication systems, multiples copies of the transmitted signal arrive at the receiver. This phenomenon causes channel variations in the frequency domain over the transmission bandwidth. Conventional systems use equalizers to tackle this problem. Another approach is to use multicarrier communication systems based on orthogonal frequency division multiplexing (OFDM). In this technique, the entire bandwidth is divided into several subchannels, and, as a result, each subchannel experiences almost a flat fading channel.

In this dissertation, we work on three different areas in OFDM systems: (1) We propose new analytical methods to evaluate the error rate of coded OFDM systems for a particular channel realization. (2) Assuming the instantaneous channel is known both at the transmitter and the receiver, we introduce adaptive transmission techniques to enhance the performance of these systems, and (3) we propose a new receiver-based technique to recover the distortion caused by the practical non-linear power amplifier.

To address the first subject, a novel analytical method for bit error rate evaluation of coded OFDM systems for a specific channel realization is proposed. As this method might be too complex for some applications, we also propose a simpler formula.

As for the second subject, we introduce new adaptive bit loading and interleaving techniques to minimize the bit error rate of the system. Also, we propose novel adaptive bit and power loading and code rate selection techniques to minimize bit error rate, to minimize the transmit power, or to maximize the throughput of the system.

To address the third subject, we propose a new method to estimate the original non-clipped signal. The estimation is done at the receiver and makes use of the newly proposed compressed sensing estimation technique.

# Preface

This thesis is based on the research work conducted at the Department of Electrical and Computer Engineering at UBC under the guidance and supervision of Professor Lutz Lampe.

Dr. Chris Snow, who was a PhD student in our group, helped me with the research related to Chapter 2 of this thesis. He provided me with codes for the simulations of this chapter. During the underlying research for Chapter 3, Mr. Alireza Kenarsari-Anhari helped me with deriving the analytical formulas. Finally, for the research conducted for Chapter 4, Dr. Abolfazl Ghassemi, who was a postdoctoral fellow in our group, helped me by having many fruitful discussions.

In the following, a list of articles and papers where the results of this thesis have been partially published/submitted is provided.

## **Journal articles**

1. (Part of Chapter 2) M. Mohammadnia-Avval, C. Snow and L. Lampe. Error Rate Analysis for Bit-Loaded Coded MIMO-OFDM. *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2340-2351, June 2010.
2. (Part of Chapter 3) M. Mohammadnia-Avval and A. Kenarsari-Anhari. Adaptive Bit-Interleaved Coded OFDM: Performance Analysis and Optimization. Submitted, May 2011.
3. (Part of Chapter 4) M. Mohammadnia-Avval, A. Ghassemi and L. Lampe. Nonlin-

ear Distortion Mitigation in OFDM Signals Using Compressed Sensing. Submitted, September 2011.

**Conference papers**

1. (Part of Chapter 2) M. Mohammadnia-Avval, C. Snow and L. Lampe. Error Rate Analysis for Bit-Loaded Coded OFDM. In *Proceedings of IEEE WCNC*, Las Vegas, NV, USA, April 2008.
2. (Part of Chapter 3) M. Mohammadnia-Avval, A. Kenarsari-Anhari and L. Lampe. Adaptive Bit-Interleaved Coded OFDM. In *Proceedings of IEEE ISPLC*, Rio de Janeiro, Brazil, March 2010.
3. (Part of Chapter 4) M. Mohammadnia-Avval, A. Ghassemi and L. Lampe. Compressed Sensing Based Recovery of Nonlinearly Distorted OFDM Signals. In *Proceedings of IEEE ICC*, Kyoto, Japan, June 2011.

# Table of Contents

<b>Abstract</b> . . . . .	ii
<b>Preface</b> . . . . .	iv
<b>Table of Contents</b> . . . . .	vi
<b>List of Tables</b> . . . . .	x
<b>List of Figures</b> . . . . .	xi
<b>List of Abbreviations</b> . . . . .	xv
<b>Notation</b> . . . . .	xvii
<b>Acknowledgments</b> . . . . .	xviii
<b>1 Introduction</b> . . . . .	1
1.1 History of Multicarrier Communications . . . . .	1
1.2 Use of OFDM in Standards . . . . .	2
1.3 Bit and Power Loading for OFDM . . . . .	2
1.4 Peak-to-Average Power Ratio Problem in OFDM . . . . .	3
1.5 BIC-OFDM System Model . . . . .	4
1.5.1 Channel Encoder and Puncturer . . . . .	4
1.5.2 Interleaver . . . . .	5

vi

*Table of Contents*

---

1.5.3	Modulation . . . . .	5
1.5.4	OFDM Symbol Framing . . . . .	5
1.5.5	Time Domain Processing . . . . .	5
1.5.6	Power Amplifier . . . . .	6
1.5.7	Channel . . . . .	6
1.5.8	Receiver . . . . .	6
1.6	Thesis Contributions and Organization . . . . .	6
<b>2</b>	<b>Performance Analysis for Bit-Loaded BIC-MIMO-OFDM . . . . .</b>	<b>10</b>
2.1	Introduction . . . . .	10
2.2	System Model . . . . .	13
2.2.1	Generic Model . . . . .	14
2.2.2	MB-OFDM System and Channel Model . . . . .	17
2.2.3	IEEE 802.11a/g System and Channel Model . . . . .	17
2.3	BER Analysis for Coded MIMO-OFDM with Bit-Loading . . . . .	18
2.3.1	Set of Error Vectors . . . . .	19
2.3.2	Pairwise Error Probability . . . . .	20
2.3.3	Performance Analysis . . . . .	21
2.3.4	Computational Complexity . . . . .	23
2.4	Bit-Loading, Adaptive Interleaving and Adaptive Coded Modulation for MIMO-OFDM . . . . .	23
2.4.1	Bit-Loading Algorithm for Coded MIMO-OFDM . . . . .	24
2.4.2	Adaptive Interleaving . . . . .	25
2.4.3	Adaptive Coded Modulation . . . . .	28
2.5	Results . . . . .	29
2.5.1	Accuracy of BER Approximation . . . . .	30

*Table of Contents*

---

2.5.2	Bit-Loading for Coded OFDM . . . . .	32
2.5.3	Adaptive Interleaving . . . . .	36
2.5.4	Adaptive Coded Modulation . . . . .	42
2.6	Summary . . . . .	44
<b>3</b>	<b>Adaptive BIC-OFDM: Performance Analysis and Optimization . . . . .</b>	<b>46</b>
3.1	Introduction . . . . .	46
3.2	Preliminaries . . . . .	47
3.2.1	System Model . . . . .	47
3.2.2	Analytical Approximation for the PDF of LLRs . . . . .	49
3.2.3	Optimization Problems For BIC-OFDM . . . . .	50
3.3	Expressions for the Performance of BIC-OFDM . . . . .	52
3.3.1	Known Interleaver . . . . .	53
3.3.2	Random Interleaver . . . . .	56
3.4	Solutions to the Optimization Problems . . . . .	59
3.4.1	Performance Evaluation for Optimized System . . . . .	59
3.4.2	Optimization Problems . . . . .	60
3.5	Numerical Results . . . . .	69
3.5.1	Variation of BER for the Optimized System for Different Interleavers . . . . .	70
3.5.2	BER Minimization Results . . . . .	70
3.5.3	Power Minimization Results . . . . .	72
3.5.4	Throughput Maximization . . . . .	74
3.6	Summary . . . . .	77
<b>4</b>	<b>Using Compressed Sensing to Mitigate Nonlinear Distortion in OFDM</b>	
	<b>Signals . . . . .</b>	<b>79</b>
4.1	Introduction . . . . .	79

*Table of Contents*

---

4.2	OFDM Signaling and Compressed Sensing . . . . .	82
4.2.1	OFDM System Model . . . . .	82
4.2.2	Background on CS . . . . .	86
4.3	CS-based Method for Nonlinear Distortion Mitigation . . . . .	88
4.3.1	Perfect Channel Estimation . . . . .	89
4.3.2	Imperfect Channel Estimation . . . . .	90
4.3.3	Modification of Other Iterative Techniques . . . . .	92
4.4	Numerical Results . . . . .	95
4.5	Summary . . . . .	99
<b>5</b>	<b>Conclusions and Future Work . . . . .</b>	<b>102</b>
5.1	Research Contributions . . . . .	102
5.2	Future Work . . . . .	104
	<b>Bibliography . . . . .</b>	<b>105</b>
 <b>Appendix</b>		
<b>A</b>	<b>Closed-form Expression for <math>\epsilon_1</math> and <math>\epsilon_2</math> for CS Reconstruction . . . . .</b>	<b>111</b>
A.1	Perfect Channel Estimation . . . . .	111
A.2	Imperfect Channel Estimation . . . . .	112

# List of Tables

2.1	Relative use of different loading schemes as best solution. MB-OFDM, UWB CM1 channel. . . . .	35
2.2	Relative use of different loading schemes as best solution. 802.11a/g, $R_c =$ $1/2$ , of (2.12), (2.13). . . . .	36
2.3	Different code rates and modulations used for the simulation results. . . . .	43
3.1	Values of the parameters used in the approximation for the PDF of LLRs. Only nonzero coefficients are shown here. . . . .	50
4.1	The proposed CS-based algorithm for estimation of nonlinear distortion as- suming imperfect channel knowledge. . . . .	92

# List of Figures

1.1	Different adaptation and PAPR reduction techniques for OFDM. Darker boxes are the categories of the proposed techniques. . . . .	3
1.2	Block diagram of the BIC-OFDM transmitter and receiver structure. . . . .	4
2.1	10% Outage BER versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ from analysis (lines) and simulations (markers) for various combinations of code rates and loading algorithms. MB-OFDM system, UWB CM1 channel. . . . .	30
2.2	10% Outage BER versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ from analysis (lines) and simulations (markers) for the MIMO-SVD system. . . . .	31
2.3	10% Outage BER versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ from analysis for various loading algorithms. MB-OFDM system with $R_c = 1/2$ , UWB CM1 channel. . . . .	32
2.4	10% outage BER versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ from analysis for various loading algorithms. 802.11a/g WLAN system with $R_c = 1/2$ , channel $T_{\text{rms}} = 250$ ns. . . . .	34
2.5	Cumulative distribution function of loading gain $G = \text{SNR}_{\text{NL}} - \text{SNR}_{\text{Loading}}$ required for BER = $10^{-5}$ for different loading schemes (NL: no loading). MB-OFDM system with $R_c = 1/2$ , UWB CM1 channel. . . . .	34
2.6	Comparing different interleavers for one channel realization. MB-OFDM system with $R_c = 1/2$ , UWB CM1 channel. . . . .	37
2.7	Average BER for different set-based adaptive interleaver and adaptive loading schemes. MB-OFDM system with $R_c = 1/2$ , UWB CM1 channel. . . . .	38

2.8	Cumulative distribution function of interleaving gain $G = \text{SNR}_{\text{standard}} - \text{SNR}_{\text{adaptive}}$ required for $\text{BER} = 10^{-5}$ for different set-based adaptive interleaver schemes. MB-OFDM system with $R_c = 1/2$ , UWB CM1 channel. . . . .	39
2.9	10% Outage BER for different adaptive interleaver schemes: symbol-based adaptive interleaver [1], bit-based adaptive interleaver [2], and our iterative adaptive interleaver algorithm. MB-OFDM system with $R_c = 1/2$ , 16-QAM modulation, UWB CM1 channel. . . . .	40
2.10	Cumulative distribution function of the interleaving gain $G = \text{SNR}_{\text{SF}} - \text{SNR}_{\text{our algorithm}}$ required for $\text{BER} = 10^{-5}$ for our iterative adaptive interleaver versus bit-based adaptive interleaving (SF) [2]. MB-OFDM system with $R_c = 1/2$ , 16-QAM modulation, UWB CM1 and CM3 channels. . . . .	41
2.11	Average BER for different adaptive interleaver schemes for MIMO-SVD. $R_c = 1/2$ , 4-QAM modulation. . . . .	42
2.12	Comparing the goodput for different adaptive coded modulation algorithms. IEEE 802.11a/g WLAN system, channel $T_{\text{rms}} = 250$ ns and $T_s = 50$ ns. . . . .	44
3.1	Different optimization problems in BIC-OFDM. . . . .	51
3.2	BER results for different number of subcarriers: (a) $N=64$ , (b) $N=256$ and (c) $N=1024$ . Square marker: highest BER for each SNR for 1000 random interleavers. Circle marker: lowest BER for each SNR for 1000 random interleavers. Solid line: union bound derived from (3.24) and (3.29). Dashed line: asymptotic BER derive from (3.24) and (3.37). . . . .	71
3.3	Average BER versus $\gamma_b$ for BER minimization algorithm. . . . .	71
3.4	10% outage BER versus $\gamma_b$ for BER minimization algorithm. . . . .	72

3.5	Cumulative distribution function (CDF) of power gain $G = \gamma_{b, \text{No Loading}} - \gamma_{b, \text{Power Minimization}}$ of power minimization algorithm compared with no loading system for $BER_t = 10^{-5}$ . . . . .	73
3.6	BER of the power minimization algorithm along with the asymptotic BER. . . . .	74
3.7	Comparing the average data rate for the proposed throughput maximization algorithm, AMC-BIC-OFDM and no loading system. . . . .	75
3.8	Average BER for the proposed throughput maximization algorithm along with the corresponding asymptotic and union bound BERs and also the BER of AMC-BIC-OFDM. . . . .	76
3.9	Percentage of code rate usage for the proposed rate maximization algorithm (line) and AMC-BIC-OFDM (dashed). . . . .	77
4.1	Compressed sensing example. . . . .	87
4.2	Distortion estimation example. . . . .	88
4.3	MSE versus clipping level ( $\alpha$ ) of SL power amplifier for systems employing different number of pilots ( $L_p$ ). Coded OFDM with 64-QAM constellation for all subcarriers. . . . .	96
4.4	The original distortion ( $\mathbf{d}$ ) and the estimated distortion ( $\hat{\mathbf{d}}$ ) for one specific OFDM symbol for SL power amplifier with different values of $\alpha$ , $10 \log_{10}(\bar{E}_b/\mathcal{N}_0) = 8\text{dB}$ . . . . .	96
4.5	BER versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ for perfect channel estimation and an SL power amplifier with $\alpha = 3\text{dB}$ . Coded and bit-loaded OFDM with an average of 4 bits per subcarrier. . . . .	97
4.6	BER versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ for perfect channel estimation and an SSPA with $\rho = 1$ , $p = 6$ and $\text{IBO} = 5\text{dB}$ . Coded and bit-loaded OFDM with an average of 4 bits per subcarrier. . . . .	98

4.7	MSE versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ for imperfect channel estimation when MMSE channel estimation, an SL power amplifier with $\alpha = 2.5\text{dB}$ , and coded and bit-loaded OFDM with an average of 4 bits per subcarrier are used. . . . .	99
4.8	BER versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ for imperfect channel estimation when MMSE channel estimation, an SL power amplifier with $\alpha = 2.5\text{dB}$ , and coded and bit-loaded OFDM with an average of 4 bits per subcarrier are used. . . . .	100
4.9	BER versus $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ for imperfect channel estimation when MMSE channel estimation, an SSPA with $\rho = 1$ , $p = 8$ and IBO = 6dB, and coded and bit-loaded OFDM with an average of 4 bits per subcarrier are used. . . . .	100

# List of Abbreviations

3GPP	3rd Generation Partnership Project
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BICM	Bit Interleaved Coded Modulation
BPSK	Binary Phase Shift Keying
CDF	Cumulative Distribution Function
CS	Compressed Sensing
CSI	Channel State Information
DFT	Discrete Fourier Transform
DSL	Digital Subscriber Line
ECMA	European Computer Manufacturers Association
FFT	Fast Fourier Transform
ICI	Inter-Channel Interference
IDFT	Inverse Discrete Fourier Transform
IEEE	Institute of Electrical and Electronic Engineers
IFFT	Inverse Fast Fourier Transform
i.i.d.	Independent, Identically Distributed
ISI	Inter-Symbol Interference
LLR	Log-Likelihood Ratio
LMMSE	Linear Minimum Mean-Squared Error

*List of Abbreviations*

---

LTE	Long Term Evolution
MB	Multi-Band
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
$M$ -PSK	$M$ -ary Phase Shift Keying
$M$ -QAM	$M$ -ary Quadrature Amplitude Modulation
MRC	Maximal Ratio Combining
OFDM	Orthogonal Frequency Division Multiplexing
PA	Power Amplifier
PAPR	Peak-to-Average Power Ratio
PDF	Probability Density Function
PEP	Pairwise Error Probability
PSK	Phase Shift Keying
QPSK	Quaternary Phase Shift Keying
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SOCP	Second Order Cone Program
SVD	Singular Value Decomposition
UWB	Ultra-Wideband
V-BLAST	Vertical-Bell Labs Layered Space-Time
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

# Notation

The notation and operators used in this thesis are listed below:

$(\cdot)^*$	Complex conjugation
$[\cdot]^T$	Transposition
$[\cdot]^H$	Hermitian transposition
$\det(\cdot)$	Matrix determinant
$\text{diag}(\mathbf{x})$	a matrix with the elements of $\mathbf{x}$ on the main diagonal
$ \cdot $	Absolute value of a complex number
$\ \cdot\ _1$	$L_1$ -norm of a vector
$\ \cdot\ _2$	$L_2$ -norm of a vector
$\mathbb{C}$	Field of complex numbers
$\mathbb{R}$	Field of real numbers
$\text{Re}\{\cdot\}$	Real part of a complex number
$\text{Im}\{\cdot\}$	Imaginary part of a complex number
$\mathbb{E}(\cdot)$	Statistical expectation
$\text{Pr}\{\cdot\}$	Probability of an event
$\oplus$	Element-wise XOR operator
$\otimes$	Convolution operator
$\mathbf{I}_X$	$X \times X$ identity matrix
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean $\mu$ and variance $\sigma^2$
$Q(\cdot)$	Gaussian Q-function, $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$

# Acknowledgments

I would like to thank my advisor, Professor Lutz Lampe, for all his support throughout the course of this work. Without his constant help and encouragement, this thesis would have not been possible.

I greatly thank the members of my examining committee for the time and effort they have put in evaluating my work and the feedback they have provided.

Also, I would like to take this opportunity to thank my colleagues in the Communication Theory Group for all the fruitful discussions and helpful comments. In particular, I thank Dr. Chris Snow, Dr. Abolfazl Ghassemi, and Alireza Kenarsari-Anhari for all the great helps they have provided.

Special thanks to my family for their love and inspiration. All my work is dedicated to my brother, Alireza, who passed away in Dec. 2009.

# Chapter 1

## Introduction

In single-carrier communication systems, the entire bandwidth is used by one symbol and successive symbols are transmitted in different time slots. When the channel varies over the transmission bandwidth, these systems require complicated equalizers at the receiver. Multicarrier systems have been proposed as a solution to this problem. In a multicarrier system, the available bandwidth is divided into several subchannels (or subcarriers). Each subchannel experiences flat fading and therefore simple equalization techniques are applicable. If the channel bandwidth and transmission parameters are chosen properly, each subcarrier will be orthogonal to other subcarriers. This can be implemented in different ways and the most popular form of implementation is called orthogonal frequency division multiplexing (OFDM).

### 1.1 History of Multicarrier Communications

The history of multicarrier communications dates back to the mid 60's, when Chang [3] proposed the concept of parallel data transmission. Weinstein and Ebert [4] proposed using discrete Fourier transform (DFT) to perform baseband modulation and demodulation, eliminating the need for banks of oscillators. To combat inter-channel interference (ICI) and inter-symbol interference (ISI), they proposed the use of a guard space between symbols. Next, Peled and Ruiz [5] introduced using cyclic prefix (CP) instead of empty guard space to assure the orthogonality of the subcarriers.

Although the idea of OFDM has been known for some time, it was only relatively recently that it has become the dominating type of signal modulation in wireless communication systems. This is because advances in very large scale integrated (VLSI) circuit technologies and digital signal processing (DSP) algorithms have eliminated the original obstacles of OFDM implementation. The interested reader can refer to [6] and [7] for a more complete history of OFDM.

## 1.2 Use of OFDM in Standards

Due to its many advantages, OFDM has gained considerable attention recently. One of the first implementations of OFDM was in the xDSL (digital subscriber lines) standards. OFDM is usually combined with binary coding and the combination is referred to as bit interleaved coded OFDM (BIC-OFDM). BIC-OFDM has gained interest from the communications community in recent years, as evidenced by standards such as IEEE 802.11a/g for Wireless Local Area Networks (WLANs) [8], IEEE 802.16 (broadband wireless access), ECMA Multiband OFDM (MB-OFDM) for high-rate Ultra Wideband (UWB) [9], and the 3GPP Long Term Evolution (LTE) wireless cellular systems. While OFDM accomplishes transmission in the temporal and spectral domains, multiple-input multiple-output (MIMO) technology exploits the spatial domain by using multiple antennas at the transmitter and receiver. MIMO-BIC-OFDM has been implemented in, for example, the IEEE 802.11n standard.

## 1.3 Bit and Power Loading for OFDM

Channel gain varies over different subcarriers of OFDM. If the channel changes slowly over time and is known at the transmitter, the number of bits and/or transmit power in each

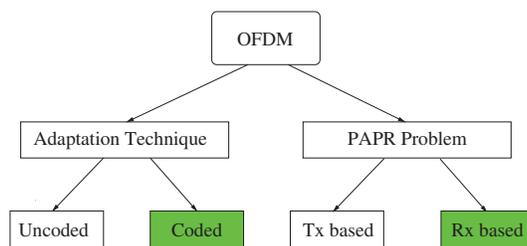


Figure 1.1: Different adaptation and PAPR reduction techniques for OFDM. Darker boxes are the categories of the proposed techniques.

subcarrier can be chosen to maximize the total bit rate, to minimize the total transmit power, or to minimize the bit error rate (BER) of the system. These schemes are known as bit loading and power loading. An OFDM system can apply bit loading, power loading or both.

The problem of maximizing the total rate through bit and power loading has long been known and the optimal solution to power loading is called *water-filling* [10]. This solution is not suitable for practical systems since the resulting bit rates are continuous. Therefore, there have been a plethora of research projects to find loading algorithms for different scenarios. However, most of these loading algorithms are developed assuming either uncoded transmission or channel coding along the bits being transmitted over individual subcarriers. A new generation of algorithms which consider the combination of OFDM with BICM has been developed in this thesis (see Section 1.6 and Figure 1.1).

## 1.4 Peak-to-Average Power Ratio Problem in OFDM

The transmitted signal of OFDM is the sum of many sinusoidal signals with different frequencies. Therefore, the resulting envelop of the OFDM signal is well approximated by a Rayleigh distribution [11]. This leads to a high peak-to-average power ratio (PAPR) of

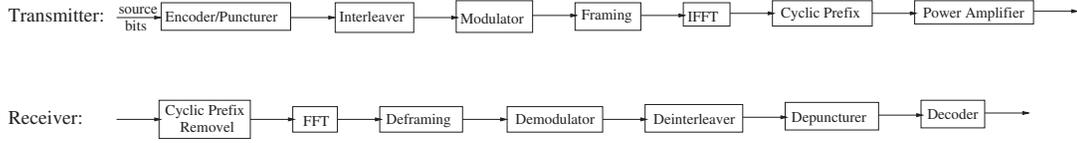


Figure 1.2: Block diagram of the BIC-OFDM transmitter and receiver structure.

the transmit signal. Since practical power amplifiers (PAs) have limited dynamic range and, the high PAPR causes signal distortion, resulting in degradation of the system BER and out-of-band radiation. To tackle this problem, a number of transmitter as well as receiver-based techniques have been proposed [12] (see Figure 1.1). Transmitter-based techniques decrease the PAPR of the transmit signal, while receiver-based techniques aim to reconstruct the original signal (the signal before power amplifier) at the receiver.

## 1.5 BIC-OFDM System Model

In this section, the BIC-OFDM transmission system considered in this thesis is introduced. The BIC-OFDM transceiver block diagram is shown in Figure 1.2. We consider single-input single-output (SISO) as well as MIMO systems.

### 1.5.1 Channel Encoder and Puncturer

For our channel encoding, we employ convolutional codes as they are the most common form of encoding in practical systems. For example, IEEE 802.11a/g, IEEE 802.16, and ECMA Multiband OFDM systems use convolutional encoding. The essence of the analysis presented in this thesis can be extended to other classes of codes such as low-density parity-check (LDPC) codes. The output of the encoder can be punctured to get higher rate codes.

### 1.5.2 Interleaver

Following encoding, the bits are interleaved. Interleaving is an important part of BICM, because it guarantees the nearby bits from convolutional encoder to be separated over different fading samples. Therefore, nearby bits experience different fading gains and thus the diversity present in the channel can be exploited during decoding at the receiver.

### 1.5.3 Modulation

Upon interleaving, the bits are modulated using a modulation scheme such as quadrature amplitude modulation (QAM) or phase shift keying (PSK). We consider two cases: 1) in the first case, fixed constellation sizes are employed where all subcarriers use the same constellation size, and 2) in the second case, a specific bit loading algorithm is applied and subcarriers use different constellation sizes accordingly.

### 1.5.4 OFDM Symbol Framing

The corresponding data symbols along with the pilots are framed into an OFDM symbol. For the MIMO-OFDM, we assume a different OFDM symbol for each antenna.

### 1.5.5 Time Domain Processing

Each OFDM symbol is transformed into time domain via IFFT. The cyclic prefix is added to the beginning of the symbol and is removed at the receiver. This causes each subcarrier to be affected by flat fading.

### 1.5.6 Power Amplifier

The time domain signal then goes through a power amplifier. The power amplifier causes signal distortion. There are different types of power amplifiers and each causes distortion differently. We talk about this in more details in Chapter 4.

### 1.5.7 Channel

We assume a fading channel model affected by AWGN. Different channel models used in this thesis are presented in Sections 2.2.2 and 2.2.3.

### 1.5.8 Receiver

At the receiver, the cyclic prefix is removed and the resulting signal passes through an FFT. Then, the symbols go through the demodulator, the deinterleaver, and finally the decoder.

## 1.6 Thesis Contributions and Organization

This thesis considers several issues regarding the analysis and design of OFDM systems. Our contributions are summarized as follows:

1. We develop a novel analysis technique for bit-loaded coded OFDM systems as well as bit-loaded coded MIMO-OFDM systems using singular value decomposition (MIMO-SVD-OFDM).
2. We use the above analysis and propose:
  - A new bit loading algorithm.
  - Three new adaptive interleaving algorithms.

- A new adaptive coded modulation algorithm.
3. We propose a simple formula for the performance of BIC-OFDM in the high SNR regime. We then use this formula to present algorithms which are capable of finding the optimum code rate, bit loading and power loading for the following optimization problems:
- Minimizing the BER subject to predefined total rate and transmit power.
  - Minimizing the total transmit power subject to predefined BER and total rate.
  - Maximizing the throughput subject to predefined BER and transmit power.
4. We develop a technique based on compressed sensing (CS) to estimate the original non-clipped signal at the receiver. The clippings are caused by the non-ideal power amplifier at the transmitter.

To present these contributions, the thesis is organized as follows. In Chapter 2, we develop a novel analytical method for the bit error rate (BER) and word error rate (WER) estimation of bit-loaded coded OFDM and MIMO-OFDM systems using singular value decomposition (SVD), operating over frequency-selective quasi-static channels with non-ideal interleaving. Then, we introduce three different applications of the proposed analysis. First, we compare the performance of several OFDM bit-loading schemes and propose a hybrid loading scheme which selects the best loading for each channel realization from a number of candidates. Second, we introduce three adaptive interleaving schemes: (i) selecting the best interleaver from a number of predefined interleavers, (ii) a novel adaptive bit interleaving algorithm based on pairwise error probability (PEP), and (iii) a spatial interleaving scheme for MIMO-SVD-OFDM systems with separate information sources. Third, we introduce an adaptive coded modulation algorithm by using our BER estimation technique.

In Chapter 3, we look into the problem of adaptive bit-loading, power allocation, and code (rate) selection for BIC-OFDM. Three optimization problems are considered: BER minimization, power minimization, and throughput maximization. We develop a framework for performance evaluation of BIC-OFDM transmission over quasi-static fading channels. This analysis is based on the probability density function (PDF) of the reliability metrics developed in [13] and the resulting expression is simpler to evaluate compared to the expression proposed in Chapter 2. Then, the BER expression is derived based on the knowledge of interleaver and (error correction) code structure. Furthermore, it is shown that the optimization problem based on this approach leads to a high dimensional mixed-integer optimization problem which (to the best of our knowledge) cannot be solved in polynomial time. Hence, we assume that the structure of interleaver is not known, i.e. uniform random interleaving assumption, and derive the expected value of BER with this assumption. Then, this BER expression is further simplified assuming asymptotically-high SNR at the receiver. We show that the optimization problem with the expected value of BER as the objective function can be solved efficiently using greedy algorithms in polynomial time. Finally, we propose a simple formula which predicts the performance of the optimized system.

Adaptive techniques proposed in Chapters 2 and 3 as well as most of the techniques available in the literatures are designed for the case when there is perfect channel state information (CSI) at the transmitter. However, in practical systems, only imperfect CSI is available which is the result of noisy channel estimation and also the estimated channel might be outdated due to delay in the feedback channel. Therefore, a number of algorithms are proposed for imperfect CSI [14, 15, 16].

Chapter 4 considers the problem of high peak to average power ratio (PAPR) of the transmit signal in OFDM systems. In realistic systems, power amplifiers have a limited

dynamic range and, as a result, OFDM signals will be clipped at the transmitter. This will cause OFDM signal distortion. In this chapter, we employ pilot subcarriers to propose methods based on compressed sensing to recover nonlinearly distorted OFDM signal. First, we assume channel is perfectly known at the receiver and pilots are used to estimate distortion. Second, we consider practical channel estimation when the pilot tones are distorted by PA nonlinearity. In this case, we modify the CS-based method so that channel estimation is also improved by our technique.

Finally, Chapter 5 summarizes the contributions of this thesis and outlines future areas of research.

# Chapter 2

## Performance Analysis for Bit-Loaded BIC-MIMO-OFDM

### 2.1 Introduction

In this chapter, we develop a novel analysis technique for bit-loaded coded MIMO-OFDM systems, extending recent results for non-loaded coded SISO OFDM [17]. This method provides the system designers with a simple way to study the performance of different bit-loading and channel coding schemes without resorting to lengthy simulations. We propose novel adaptive bit loading, adaptive interleaving and adaptive coded modulation based on the derived analysis.

There are well-known techniques for bounding the performance of convolutionally-encoded transmission over many types of fading channels, e.g. [18, 19]. However, such classical techniques are not applicable to the OFDM systems mentioned above for several reasons. Firstly, the short-length channel-coded packet-based transmissions are non-ideally interleaved, which results in non-zero correlation between adjacent coded bits. Secondly, and more importantly, the quasi-static nature of the wireless channel limits the number of distinct channel gains to the (relatively small) number of OFDM subcarriers. This small number of distinct channel gains must not be approximated by the full fading distribution for a valid performance analysis, as would be the case in a fast-fading channel. Recent work has developed a pairwise error probability analysis for loaded coded OFDM [20, 21].

However, the analysis in [20] assumes a uniform distribution of label and OFDM subcarrier positions for coded bits and the resultant expression ([20, Ineq. (13)]) is not amenable to numerical evaluation. The analysis in [21] is based on a Gaussian approximation [22] and we found the resulting approximation of BER to be quite loose. As in [20], the effect of finite interleaving is not considered in the BER expression. In [23], the authors derived the capacity for OFDM system using bit and power loading under Rayleigh fading with Maximal Ratio Combining (MRC) at the receiver. They also proposed a bit and power loading for the same system when coding is not used. The authors in [24] derive the formula for the performance of the loading algorithm of [25]. Their result is accurate only if the number of subcarriers is large or the channel is highly frequency selective.

As stated in Chapter 1, adaptive transmission in the form of adaptive bit and power loading, is a powerful technique for performance enhancement in OFDM system [20, 25, 26, 27, 28, 29, 30, 31, 32, 33]. These algorithms aim to optimize the transmission process utilizing the priori knowledge about the channel gains over different subcarriers. Many of these algorithms are not directly applicable to BIC-OFDM transmission systems because they do not take into account the effect of coding across the bits from different subcarriers [25, 30, 31, 32, 33]. They are mainly developed assuming either uncoded transmission or channel coding along the bits being transmitted over individual subcarriers. Therefore, a new generation of algorithms which consider the combination of OFDM with BICM have been developed, e.g., [20, 26, 27, 28, 29]. As in almost all adaptive transmission systems, the proposed methods require the instantaneous channel knowledge at the transmitter or alternatively, the parameters of the optimized system. In quasi-static fading environments such as typical indoor WLAN channels, the slow time variation of the channel makes it feasible to implement a reverse link that sends back some channel information to the transmitter.

Adaptive modulation for single-carrier BICM was first introduced in [34] in which adaptation of constellation for a channel with correlated fading using outdated channel state information has been considered. Furthermore, [20] has considered combined adaptation of channel coding and bit/power allocation over different subchannels. As stated before, the derived BER expression is too complex to be used as an optimization criterion. Therefore, they used a heuristic criterion alike the one used in e.g. [25, 30, 35]. It has been found that the performance of the proposed adaptive system is not desirable and further optimization using auxiliary power allocation is needed in order to achieve considerable improvement. Unfortunately, a heuristic approach should be used to derive this auxiliary power allocation and its performance can only be verified through simulation. In [26] and [27] the problem of power allocation for BIC-OFDM transmission systems was studied. In particular, the authors of [26] developed an algorithm which optimizes the so called goodput of the system by using a BER approximation based on the saddlepoint approximation for the pairwise error probability (cf. [22]). Reference [27] considered BIC-OFDM transmission using binary phase-shift keying (BPSK) or quadrature phase-shift keying (QPSK) constellations and used the BER union bound in order to formulate the problem as a convex optimization problem. In [28], the author proposed power loading with equal constellation size for all subcarriers and also bit loading with equal transmit power for all subcarriers subject to a target BER and total rate. The channel model of [28] is frequency-domain block fading with some independent fading blocks in the transmission bandwidth. The authors of [29] proposed bit and power loading for BIC-OFDM systems to maximize the mutual information. To render the optimization problem convex, they used irregular modulation and power allocation which means it is possible to transmit only a fraction of a symbol over a subcarrier. This cannot be implemented in practical systems, so the authors suggested rounding, which results in some performance degradation.

Different from bit and power loading, interleaver affects the performance of the system only in the case of coded transmission. Since the analysis for coded OFDM has only been recently considered, adaptive interleaving has only been proposed in [1, 2]. These techniques will be explained in Section 2.4.2.

This chapter is organized as follows. In Section 2.2, the system model and notations are introduced. The analysis technique is presented in Section 2.3. We assume identical transmit power in all OFDM subcarriers, which is necessary for systems such as UWB MB-OFDM, where power spectral density is constrained. In Section 2.4 we present three applications for the proposed analysis. First, we propose and evaluate a hybrid loading scheme which selects the best loading for each channel realization from a number of candidates. Second, we introduce three different adaptive interleaving algorithms. The first algorithm is based on selecting the best interleaver among a set of interleavers. In the second algorithm, a bit interleaver is designed based on the PEPs derived using our error rate estimation technique. For MIMO-OFDM systems employing SVD, we propose an adaptive space-frequency interleaving. The third application is an adaptive coded modulation scheme using our BER estimation technique. Numerical results confirming the accuracy of our error rate estimation and illustrating the applications mentioned above are shown in Section 2.5 for MB-OFDM UWB and IEEE 802.11a/g systems, and also for the MIMO-SVD extension of IEEE 802.11a/g.

## 2.2 System Model

Before presenting the novel analysis and the improved designs, we first briefly establish the system model for coded bit-loaded MIMO-OFDM transmission. The block diagram with a brief explanation for each block is presented in Section 1.5.

### 2.2.1 Generic Model

Let us consider an  $L$ -subcarrier MIMO-OFDM system with  $N_T$  transmit and  $N_R$  receive antennas (we assume  $N_T \leq N_R$ ). It employs a general bit-loading scheme, which selects a  $2^{m_{i,j}}$ -ary quadrature amplitude modulation ( $2^{m_{i,j}}$ -QAM) for subcarrier  $1 \leq i \leq L$  over the antenna  $1 \leq j \leq N_T$  based on the channel conditions.<sup>1</sup> For non-square constellations, we use the constellations proposed in [36]. We denote the average number of coded bits per modulated symbol by  $\bar{m}$ . The particular loading algorithms applied will be discussed in Section 2.4.

The system employs a (possibly punctured) convolutional code of rate  $R_c$ . We assume that the transmitter selects a vector of  $B = R_c L \bar{m} N_T$  random message bits for transmission, denoted by

$$\underline{b} = [b_1 \ b_2 \ \dots \ b_B]^T . \quad (2.1)$$

The message bits are convolutionally encoded by the mapping

$$\mathcal{C} : \{0, 1\}^B \rightarrow \{0, 1\}^N \quad (2.2)$$

to produce the vector

$$\underline{c} = \mathcal{C}(\underline{b}) \quad (2.3)$$

of length  $N = L \bar{m} N_T$ . The vector  $\underline{c}$  is then interleaved by the mapping

$$\underline{\pi} : \{0, 1\}^N \rightarrow \{0, 1\}^N \quad (2.4)$$

resulting in the vector

$$\underline{c}^\pi = \underline{\pi}(\underline{c}) \quad (2.5)$$

---

<sup>1</sup>With a slight abuse of notation, we use the term 2-QAM to denote binary phase shift keying (BPSK).

of length  $N$ . The interleaved bits  $\underline{c}^\pi$  are finally modulated using  $2^{m_{i,j}}$ -QAM on subcarrier  $1 \leq i \leq L$  over antenna  $1 \leq j \leq N_T$ , where the modulation is represented by the mapping

$$\mathcal{M}_{\mathbf{h}} : \{0, 1\}^N \rightarrow \mathbb{C}^{NN_T} . \quad (2.6)$$

Hence, the vector of the  $NN_T$  modulated symbols is given by

$$\underline{x} = [x_{1,1} \ x_{1,2} \ \dots \ x_{1,N_T} \ x_{2,1} \ x_{2,2} \ \dots \ x_{2,N_T} \ \dots \ x_{N,1} \ x_{N,2} \ \dots \ x_{L,N_T}]^T = \mathcal{M}_{\mathbf{h}}(\underline{c}^\pi) , \quad (2.7)$$

where  $x_{i,j}$  is the symbol on subcarrier  $1 \leq i \leq L$  over antenna  $1 \leq j \leq N_T$ .

It is important to note the dependence of  $\mathcal{M}_{\mathbf{h}}$  on the frequency-domain channel gains as a result of the channel gain dependent loading algorithms. For a particular channel gain, the mapping  $\mathcal{M}_{\mathbf{h}}$  is obtained by running the chosen loading algorithm in order to select the modulation for each subcarrier. The channel gain for the  $i$ th subcarrier is given by the  $N_R \times N_T$  matrix

$$\mathbf{H}_i = \begin{bmatrix} h_{1,1,i} & h_{1,2,i} & \dots & h_{1,N_T,i} \\ h_{2,1,i} & h_{2,2,i} & \dots & h_{2,N_T,i} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_R,1,i} & h_{N_R,2,i} & \dots & h_{N_R,N_T,i} \end{bmatrix} , \quad (2.8)$$

where  $h_{k,j,i}$  is the frequency-domain channel gain from transmit antenna  $j$  to receive antenna  $k$  for the  $i$ th subcarrier. Also we define the matrix  $\mathbf{D}$  as the block diagonal matrix of size  $LN_R \times LN_T$  consisting of all  $\mathbf{H}_i$  matrices on the main diagonal.

We will assume that the MIMO-OFDM system is designed such that the cyclic prefix is longer than the channel impulse response and that timing and frequency synchronization have been established. Thus, we can equivalently consider the channel in the frequency

domain and express the received symbols as

$$\underline{r} = \sqrt{\bar{E}_s} \mathbf{D} \underline{x} + \underline{n}, \quad (2.9)$$

where  $\underline{n}$  is a vector of independent complex additive white Gaussian noise (AWGN) variables of length  $LN_R$  with variance  $\mathcal{N}_0$  and  $\bar{E}_s$  is the average received energy per modulated symbol assuming that  $\mathbb{E}\{\|\mathbf{D}\underline{x}\|^2\} = \mathbf{I}_{LN_R}$ . The average received energy per information bit is  $\bar{E}_b = \bar{E}_s / (R_c \bar{m})$ .

The receiver employs a soft-output detector followed by deinterleaving, depuncturing, and Viterbi decoding, resulting in an estimate

$$\hat{\underline{b}} = [\hat{b}_1 \ \hat{b}_2 \ \dots \ \hat{b}_B]^T \quad (2.10)$$

of the original transmitted information bits.

For MIMO processing, we assume singular value decomposition. Performing SVD on  $\mathbf{H}_i$  results in

$$\mathbf{H}_i = \mathbf{U}_i \mathbf{\Lambda}_i \mathbf{V}_i^H \quad (2.11)$$

where  $\mathbf{U}_i$  and  $\mathbf{V}_i$  are unitary matrices. The entries of the diagonal matrix  $\mathbf{\Lambda}_i$  are non-negative singular values of  $\mathbf{H}_i : \lambda_{i,1}, \lambda_{i,2} \dots \lambda_{i,N_T}$ . In the standard SVD operation we have  $\lambda_{i,1} \geq \lambda_{i,2} \geq \dots \geq \lambda_{i,N_T}$  and in MIMO-SVD transmission,  $\mathbf{V}_i$  is applied to the transmitted signal and  $\mathbf{U}_i^H$  is applied to the received signal. This will result in  $N_T$  parallel subcarriers with gains  $\lambda_{i,j}$  for  $1 \leq j \leq N_T$ . Then we put these gains in a vector  $\boldsymbol{\lambda}$  of length  $LN_T$  according to how encoded bits are assigned to different subcarriers. For convenience, we define the diagonal matrix  $\mathbf{\Lambda} = \text{diag}(\boldsymbol{\lambda})$ .

In the following, we describe two popular (practical) OFDM systems and channel models which we will consider in the performance evaluation in Section 2.5.

### 2.2.2 MB-OFDM System and Channel Model

As the first OFDM system example, we have chosen MB-OFDM for high data-rate UWB [9, 37]. MB-OFDM uses 128 subcarriers and operates by hopping over 3 sub-bands (one hop per OFDM symbol) in a predetermined pattern. We will assume that hopping pattern 1 of [9] is used (i.e., the sub-bands are hopped in order). As a result we can consider MB-OFDM as an equivalent 384 subcarrier OFDM system. After ignoring pilot, guard, and other reserved subcarriers, we have 300 data-carrying subcarriers.

Channel coding consists of classical BICM [19] with a punctured maximum free distance rate  $1/3$  constraint length 7 convolutional encoder and a multi-stage block-based interleaver (see [9] for details). In the standard, the interleaved coded bits are mapped to 4-QAM symbols using Gray labeling. To maintain the same data rates but decrease the error probability, we instead employ loading as described above with  $\bar{m} = 2$  bits per subcarrier.

For a meaningful performance analysis of the MB-OFDM proposal, we consider the channel model developed by IEEE 802.15 for UWB systems [38]. The channel impulse response is based on a modified Saleh-Valenzuela model [39]. As well, the entire impulse response undergoes an “outer” lognormal shadowing. The channel impulse response is assumed time invariant during the transmission period of (at least) one packet (see [38] for detailed description). We consider the UWB channel parameter sets CM1 and CM3 [38].

### 2.2.3 IEEE 802.11a/g System and Channel Model

The second example system we consider is IEEE 802.11a/g, which employs 64 subcarriers, of which 48 are used for data transmission [8]. Channel coding is again BICM, with a punctured maximum free distance rate  $1/2$  constraint length 7 convolutional encoder. We adopt the quasi-static exponentially-decaying multipath Rayleigh fading model used in [20, 21], where the gain over subcarrier  $i$  from transmit antenna  $k$  to receive antenna  $j$

is given by

$$h_{k,j,i} = \sum_{i=0}^{L_m-1} \bar{h}_{k,j}(i) \exp\left(-j\frac{2\pi il}{64}\right), \quad (2.12)$$

where  $L_m$  is the number of channel taps and  $\bar{h}_{k,j}(i)$  is the  $i$ th component of the channel impulse response, modeled as a complex Gaussian random variable [20, 21]

$$\bar{h}_{k,j}(i) \sim \mathcal{CN}(0, \sigma_0 \exp(-iT_s/T_{\text{rms}})), \quad (2.13)$$

where

$$\sigma_0 = 1 - \exp(-T_s/T_{\text{rms}}), \quad (2.14)$$

$T_s = 50$  ns is the receiver sampling rate, and  $T_{\text{rms}}$  is the RMS delay spread of the channel.

We also consider the extension to this system when  $N_T = 2$  antennas are used at the transmitter and  $N_R = 2$  antennas are used at the receiver. As stated before, we use SVD for MIMO processing.

## 2.3 BER Analysis for Coded MIMO-OFDM with Bit-Loading

We now present the method for approximating the performance of bit-loaded coded MIMO-OFDM systems operating over frequency-selective, quasi-static fading channels. This method is based on approximating the performance of the system over individual channel realizations. Following the approach for non-bit-loaded SISO-OFDM in [17], we start by considering the set of error vectors for bit-loaded OFDM.

### 2.3.1 Set of Error Vectors

Consider a convolutional encoder initialized to the all-zero state, where the reference (correct) codeword is the all-zero codeword. We construct all  $M$  input sequences which cause an immediate deviation from the all-zero state (i.e., those whose first input bit is 1) and subsequently return the encoder to the all-zero state with an output Hamming weight of at most  $w_{\max}$ . Let  $\mathcal{E}$  be the set of all vectors  $\mathbf{e}_\ell$  ( $1 \leq \ell \leq L$ ) representing the output sequences (after puncturing) associated with these input sequences, i.e.,  $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_L\}$ . Let  $l_\ell$  be the length of  $\mathbf{e}_\ell$  (the number of output bits after puncturing), and let  $a_\ell$  be the Hamming weight of the input associated with  $\mathbf{e}_\ell$ . Note that the choice of  $w_{\max}$  governs the value of  $M$  (i.e., once the maximum allowed Hamming weight is set, the number of error events  $M$  is known).

We term  $\mathbf{e}_\ell$  an “error vector” and  $\mathcal{E}$  the set of error vectors. The set  $\mathcal{E}$  contains all the low-weight error events, which are the most likely deviations in the trellis. As with standard union-bound techniques for convolutional codes [18], the low-weight terms will dominate the error probability. Hence, it is sufficient to choose a small  $w_{\max}$  — for example, the punctured MB-OFDM code of rate  $R_c = 1/2$  [9] has a free distance of 9, and choosing  $w_{\max} = 14$  (resulting in a set of  $M = 242$  error vectors of maximum length  $l = 60$ ) provides results which are not appreciably different from those obtained using larger  $w_{\max}$  values.

We note that in case of loading in OFDM, the combination of coding and modulation is not a linear operation, and thus the error-rate performance will depend on the transmitted codeword. Hence, it is not sufficient to consider only the all-zero word as reference codeword. Nevertheless, for tractability and simplicity of the analysis, we always choose one codeword (in which the message bits are randomly generated) as reference. Extensive simulations (see also Section 2.5) have confirmed that the choice of the reference codeword is not critical.

### 2.3.2 Pairwise Error Probability

Next, we determine the PEP by considering error events starting in a given position  $i$  of the chosen reference codeword. The set  $\underline{\zeta}$  of allowable starting positions  $i$  has size  $|\underline{\zeta}| = R_c N$ , and each element  $i$  of  $\underline{\zeta}$  is an index  $1 \leq i \leq N$ , which is code-dependent. For example, for a code of rate  $R_c = 1/2$  the allowable starting positions are  $\underline{\zeta} = \{1, 3, 5, \dots, N-1\}$ .

We consider each error vector  $\underline{e}_\ell$  for  $1 \leq \ell \leq M$ , and form the full error codeword

$$\underline{q}_{i,\ell} = \left[ \underbrace{0 \ 0 \ \dots \ 0}_{i-1} \ \underbrace{\underline{e}_\ell}_{l_\ell} \ \underbrace{0 \ 0 \ \dots \ 0}_{L_c - l_\ell - i + 1} \right]^T \quad (2.15)$$

of length  $N$  by padding  $\underline{e}_\ell$  with zeros on both sides as indicated above. Given the error codeword  $\underline{q}_{i,\ell}$  and given that codeword  $\underline{c}$  is transmitted, the competing codeword is given by

$$\underline{v}_{i,\ell} = \underline{c} \oplus \underline{q}_{i,\ell}. \quad (2.16)$$

Interleaving and modulation results in the vector of QAM symbols

$$\underline{z}_{i,\ell} = \mathcal{M}\mathbf{h}(\underline{v}_{i,\ell}^\pi), \quad (2.17)$$

where  $\underline{v}_{i,\ell}^\pi = \underline{\pi}(\underline{v}_{i,\ell})$  is the interleaved version of  $\underline{v}_{i,\ell}$ .

The PEP for the  $\ell$ th error vector starting in the  $i$ th position is then given by

$$\text{PEP}_{i,\ell}(\mathbf{\Lambda}) = Q \left( \sqrt{\frac{\bar{E}_s}{2\mathcal{N}_0} \|\mathbf{\Lambda}(\underline{x} - \underline{z}_{i,\ell})\|_2^2} \right). \quad (2.18)$$

Note that the PEP depends on the particular channel realization  $\mathbf{D}$  only through the matrix of singular values  $\mathbf{\Lambda}$ .

### 2.3.3 Performance Analysis

The PEP from (2.18) is now used to obtain an approximation of the BER for a particular matrix of channel gains  $\mathbf{\Lambda}$ , which we denote as  $P(\mathbf{\Lambda})$ . To this end, the bit error rate for the error event starting in a position  $i \in \underline{\zeta}$  ( $1 \leq i \leq N$ ) and error vector  $\underline{e}_\ell \in \mathcal{E}$  ( $1 \leq \ell \leq M$ ) is given by

$$P_{i,\ell}(\mathbf{\Lambda}) = a_\ell \cdot \text{PEP}_{i,\ell}(\mathbf{\Lambda}) , \quad (2.19)$$

where  $a_\ell$  is the number of information bit errors associated with  $\underline{e}_\ell$ . Summing over all  $M$  error vectors, we obtain an approximation of the BER for starting position  $i$  as

$$P_i(\mathbf{\Lambda}) = \sum_{\ell=1}^M a_\ell \cdot \text{PEP}_{i,\ell}(\mathbf{\Lambda}) . \quad (2.20)$$

Since all allowable starting positions are equally likely, the BER  $P(\mathbf{\Lambda})$  can be written as

$$P(\mathbf{\Lambda}) = \frac{1}{R_c N} \sum_{i \in \underline{\zeta}} \min \left[ \frac{1}{2}, \sum_{\ell=1}^M P_{i,\ell}(\mathbf{\Lambda}) \right] , \quad (2.21)$$

where we tightened the union bound (2.20) by the maximum value of 1/2 before averaging over starting positions.

Finally, the average BER for a given number  $N_c$  of channel realizations, where the  $i$ th channel realization is denoted by  $\mathbf{\Lambda}^{(i)}$  ( $1 \leq i \leq N_c$ ), is given by

$$\bar{P} = \frac{1}{N_c} \sum_{i=1}^{N_c} P(\mathbf{\Lambda}^{(i)}) . \quad (2.22)$$

Similarly, we can obtain the  $X\%$  outage BER performance as

$$P_{\text{out}} = \max_{\mathbf{\Lambda}^{(i)} \in \mathbf{\Lambda}_{\text{in}}} P(\mathbf{\Lambda}^{(i)}) . \quad (2.23)$$

where  $(100 - X)\%$  best channel realizations are contained in  $\mathbf{\Lambda}_{\text{in}}$ . This provides information about the minimum performance that can be expected of the system given the  $X\%$  outage rate.

In pseudocode, the algorithm to calculate  $P(\mathbf{\Lambda})$  according to (2.21) is

```

Run loading algorithm to obtain  $\mathcal{M}_{\mathbf{h}}$ 
 $P := 0$ 
for  $i \in \underline{\zeta}$  do
     $P_i := 0$ 
    for  $\ell := 1$  to  $M$ 
        form  $\underline{q}_{i,\ell}$  as per (2.15)
        calculate  $\underline{v}_{i,\ell} = \underline{c} \oplus \underline{q}_{i,\ell}$  as per (2.16)
        form  $\underline{v}_{i,\ell}^\pi = \underline{\pi}(\underline{v}_{i,\ell})$  using mapping (2.4)
        calculate  $\underline{z}_{i,\ell} = \mathcal{M}_{\mathbf{h}}(\underline{v}_{i,\ell}^\pi)$  as per (2.17)
        calculate  $P_{i,\ell}$  as per (2.19)
         $P_i := P_i + P_{i,\ell}$ 
    endfor
     $P := P + \min(\frac{1}{2}, P_i)$ 
endfor
 $P := P/(R_c N)$ 

```

The PEP formula of (2.18) can also be used to obtain the approximate word error rate (WER) for a specific channel realization  $\mathbf{\Lambda}$ , which we denote as  $WER(\mathbf{\Lambda})$ . First we derive the possibility of not having any error at position  $i \in \underline{\zeta}$  ( $1 \leq i \leq N$ ) as

$$C_i(\mathbf{\Lambda}) = \max \left[ 0, 1 - \sum_{\ell=1}^M \text{PEP}_{i,\ell}(\mathbf{\Lambda}) \right]. \quad (2.24)$$

The WER is the probability of having at least one error in the frame, which can be related to  $C_i(\mathbf{\Lambda})$  as follows

$$WER(\mathbf{\Lambda}) = 1 - \prod_{i \in \zeta} C_i(\mathbf{\Lambda}) . \quad (2.25)$$

### 2.3.4 Computational Complexity

We note that the computational complexity of the analysis scales linearly with  $M$  (the size of the set of error vectors) and with  $N$  (the codeword length). Especially the choice of  $M$  allows to trade off complexity and accuracy (see also Section 2.5.2). But even for large  $M$ , the analysis requires much fewer operations compared to the alternative of performing system simulations, especially for reasonably low error rates. To provide an appreciation of the run-time savings using the proposed method, we note that the analysis for  $M \approx 250$  and  $N = 600$  takes only a few seconds per channel realization on a modern PC, and the outage BER for a large number of channel realizations (e.g.,  $N_c = 500$ ) can be obtained in the time it would take to perform simulations for only one channel realization. Finally, we note that long packet lengths can be considered without any increase in complexity (if they are segmented into codewords of length  $N$ , as is usually the case in practical systems), since the error rate for each codeword will be identical as a result of the quasi-static channel conditions.

## 2.4 Bit-Loading, Adaptive Interleaving and Adaptive Coded Modulation for MIMO-OFDM

In this section we present several different algorithms to accomplish high-performance adaptive coded MIMO-OFDM based on the analytical error-rate expressions derived in

the previous section.

### 2.4.1 Bit-Loading Algorithm for Coded MIMO-OFDM

A plethora of bit loading algorithms for OFDM systems have been proposed in the literature. Some examples of these are: the Hughes-Hartogs algorithm (HHA) [6, 40], the algorithm of Chow, Cioffi and Bingham (CCB) [35], the Piazzo algorithm [41], and the algorithm of Fischer and Huber [25]. The reader is referred to the respective papers for the details of each loading algorithm. Also, in [42], a new way of bit loading was proposed. In this method, the loading algorithm is not performed for every channel realization; it is done only for one channel realization and the resulting bit loading is just sorted for other channel realizations according to their subcarrier gains. Therefore, it is less complex than any other method. By using this method and using, for example, CCB as the primary loading algorithm, we have a bit loading algorithm which we will refer to as sorted CCB.

All the mentioned algorithms have the same shortcoming: they do not guarantee that the selected loading is appropriate for coded OFDM, i.e., BIC-OFDM. The simple reason for this is that coding and interleaving have a great impact on the error-rate performance, yet the above-mentioned algorithms simply do not consider them. Here, we propose a bit loading algorithm for coded OFDM system which also considers the effects of coding and interleaving. In our algorithm, for each specific channel realization, interleaver, average rate, and transmit power, we compute the approximate performance of the system for a set of bit loading algorithms and then the bit loading with minimum BER is selected. We call our algorithm “Best Solution Loading” (BSL). We note that BSL is a pragmatic approach to loading for coded OFDM, which explores a (hopefully promising) subset of all possibilities for loading. The identification of the optimum loading for coded OFDM requires a full search and thus is computationally prohibitive. It should be emphasized

that we only consider bit loading and the power is the same for all subcarriers.

It should be noted that the proposed methodology is general and it is straightforward to generalize it for the case of reducing transmit power subject to fixed BER and average rate. All that needs to be done is to run the corresponding bit loading algorithms (like Campello Margin Maximization algorithm [30]), then calculate the BER and finally select the one which satisfies the BER constraint and consumes less power. In this chapter, we do not consider this type of bit loading due to space limitation.

In terms of required signaling for loading, our algorithm does not increase the required signaling of the system. The reason is that if the algorithm is implemented at the receiver (transmitter), then receiver (transmitter) needs to send the bit loading to the transmitter (receiver) and the required amount of information which has to be sent as a result of our algorithm is equal to the case when only one of the bit loading algorithms is used. Also, our algorithm slightly increases the complexity of the system by requiring all the 5 bit loading algorithms to be run.

## 2.4.2 Adaptive Interleaving

Next, we consider adaptive interleaving. Different from bit-loading, the choice of interleaving has only an effect if coded transmission is considered, and thus interleaver optimization has received less attention than bit-loading in the literature. Recent references which address the problem of adaptive interleaving are [1, 2]. Here, we present three adaptive interleaving algorithms for BIC-MIMO-OFDM using the proposed error-rate analysis. We note that adaptive interleaving can be employed on its own or in combination with bit-loading. The former case may be preferable since no changes in signal constellations and thus modulation and demodulation are necessary.

### Best Interleaver of a Set

The first method is based on selecting the best interleaver among a set of interleavers, where the best interleaver is again determined using the general error-rate analysis derived earlier. When combining this adaptive interleaving method with bit-loading, the proposed scheme performs a search for both loading and interleaving for each channel realization. A pragmatic choice for this search is a two-step (greedy-type) approach, which in a first step selects the best interleaver for a non-loaded system, and in a second step selects the best loading for this given interleaver. There are a number of options for determining the set of candidate interleavers from which the “best” is chosen. In Section 2.5, we will present results for both random sets and sets optimized for particular channel models.

It should be noted that this algorithm and also finding the best bit loading are fast because for calculating the BER for each interleaver/bit loading, it is not necessary to redo all the required steps since only the channel gain  $\mathbf{\Lambda}$  is changing.

### Iterative Interleaver Improvement

The second adaptive interleaver algorithm starts from an initial interleaver and improves it by going through several iterations. Each iteration consists of the following steps:

- (1) Calculate  $P_{i,\ell}(\mathbf{\Lambda}) = a_\ell \cdot \text{PEP}_{i,\ell}(\mathbf{\Lambda})$ .
- (2) Find the error vector and starting position of the largest  $P_{i,\ell}(\mathbf{\Lambda})$ , which we denote as  $e_{max}$  and  $p_{max}$ , respectively.
- (3) Find the starting position of the smallest  $P_{i,e_{max}}(\mathbf{\Lambda})$ , which we denote as  $p_{min}$ .
- (4) Find the subcarriers that contain the error vector  $e_{max}$  at the starting position of  $p_{max}$  and store them at the vector  $\underline{g}_{max}$ .

- (5) Find the subcarriers that contain the error vector  $e_{max}$  at the starting position of  $p_{min}$  and store them at the vector  $\underline{g}_{min}$ .
- (6) Find the minimum of the vector  $\underline{g}_{max}$  and find the corresponding position of the interleaver which we denote as  $index_{min}$ .
- (7) Find the maximum of the vector  $\underline{g}_{min}$  and find the corresponding position of the interleaver which we denote as  $index_{max}$ .
- (8) Exchange the values of interleaver for  $index_{min}$  and  $index_{max}$ .
- (9) Calculate PEP for the new interleaver.

We observe that after several iterations, the algorithm may oscillate, i.e., it may find a repeated set of  $e_{max}$ ,  $p_{max}$  and  $p_{min}$ . In this situation, our algorithm is modified to find the next largest  $P_{i,\ell}(\mathbf{\Lambda})$  until it finds a set of  $e_{max}$ ,  $p_{max}$  and  $p_{min}$  that has not been used before. It is worthwhile to note that the PEP calculation in step 9 is fast, since the new interleaver is different from the old one in only two positions and, as a result, only a few PEPs need to be recalculated.

### Spatial Interleaving

A special case of MIMO-SVD transmission occurs if  $N_T$  separate information streams are transmitted.<sup>2</sup> In this case, the MIMO-OFDM system is considered as  $N_T$  parallel SISO-OFDM systems in which the data of each information stream is transmitted over its specific SVD subcarrier. We propose spatial interleaving for these systems. In spatial interleaving, we first sort the singular values and then circularly shift them by one for each subcarrier. For example, in a  $N_T = 2$ ,  $N_R = 2$  MIMO-SVD system, the SVD operation will result to two singular values:  $\lambda_{i,1}, \lambda_{i,2}$  for  $1 \leq i \leq L$ . In the standard system,  $\lambda_{i,1}$  is always greater than or equal to  $\lambda_{i,2}$  and therefore the resulting channel for the first source is better than

---

<sup>2</sup>Separate information streams per spatial layer are used in, e.g., 3GPP LTE systems.

the second source. When we want to deploy spatial interleaving in this case, we should keep  $\mathbf{\Lambda}_i$ ,  $\mathbf{U}_i$  and  $\mathbf{V}_i$  for the odd-numbered tones and switch the position of  $\lambda_{i,1}$  and  $\lambda_{i,2}$  in  $\mathbf{\Lambda}_i$  and change  $\mathbf{U}_i$  and  $\mathbf{V}_i$  correspondingly for even-numbered tones. This way we will have two links with almost similar quality. To evaluate the performance of these systems, each system should be analyzed separately, and the overall performance of the system is given by the average of these performances.

### Remarks

As mentioned above, the idea of adaptive interleaving has been considered in [1, 2]. In [1], the interleaver is symbol-based rather than bit-based as in our proposal, and the interleaver selected for each channel realization is derived according to the channel gains of each subcarrier such that error bursts are broken. Parallel to our work, [2] has devised an adaptive bit-interleaving scheme, where the interleaver is designed based on the bit level capacities of an equivalent binary channel model. However, using capacity as a metric may not be accurate in practical systems, as it does not take the particular coding scheme into account. Interestingly, the adaptive bit interleaver of [2] would result in the adaptive symbol interleaver of [1] for 4-QAM with Gray labeling. This is because bit level capacities for both bits in 4-QAM are equal and therefore, the corresponding metrics of the algorithm of [2] are the scaled versions of the metrics of [1]. Finally, we note that the BER approximations for coded OFDM derived in [20] and [21] are not applicable for adaptive interleaving, since in both cases ideal interleaving is assumed.

### 2.4.3 Adaptive Coded Modulation

Optimization of bit-loading and adaptive interleaving entails a computational complexity which may be too high for certain applications. An alternative to selecting the “optimal”

constellation for each subcarrier as well as the “optimal” interleaver for BIC-OFDM is to choose a coded modulation scheme (a code of a certain rate and a signal constellation of a certain size, combined via a pre-defined bit-interleaver) from a finite set of such schemes. Each of these coded modulation schemes provides different data rates, and using the devised BER and WER analysis, we select the coded modulation scheme that achieves the highest rate while not exceeding the desired target WER. If the WER constraint cannot be achieved by any coded modulation scheme, no data is transmitted. This adaptive coded modulation is, in terms of computational complexity, also particularly appealing, since the WER computation does not start from scratch but rather needs only to include the different channel realizations.

Adaptive coded modulation has also been considered in [21], where the authors select the best combination of code rate and constellation size based on a simplified BER expression. We found, however, that this expression does not yield tight approximations of the BER and thus is not well-suited for the purpose of adaptive coded modulation.

## 2.5 Results

In this section, we present numerical results to illustrate the usefulness of the analysis presented in Section 2.3 and the performance of the adaptive schemes proposed in Section 2.4. Throughout this section, we allow  $m_i \in \{0, \dots, 6\}$  bits per subcarrier, and employ loading schemes with an average  $\bar{m} = 2$  bits per subcarrier unless noted otherwise. We consider the two OFDM systems described in Section 2.2, namely MB-OFDM (Section 2.2.2) and 802.11a/g with its MIMO extension (Section 2.2.3).

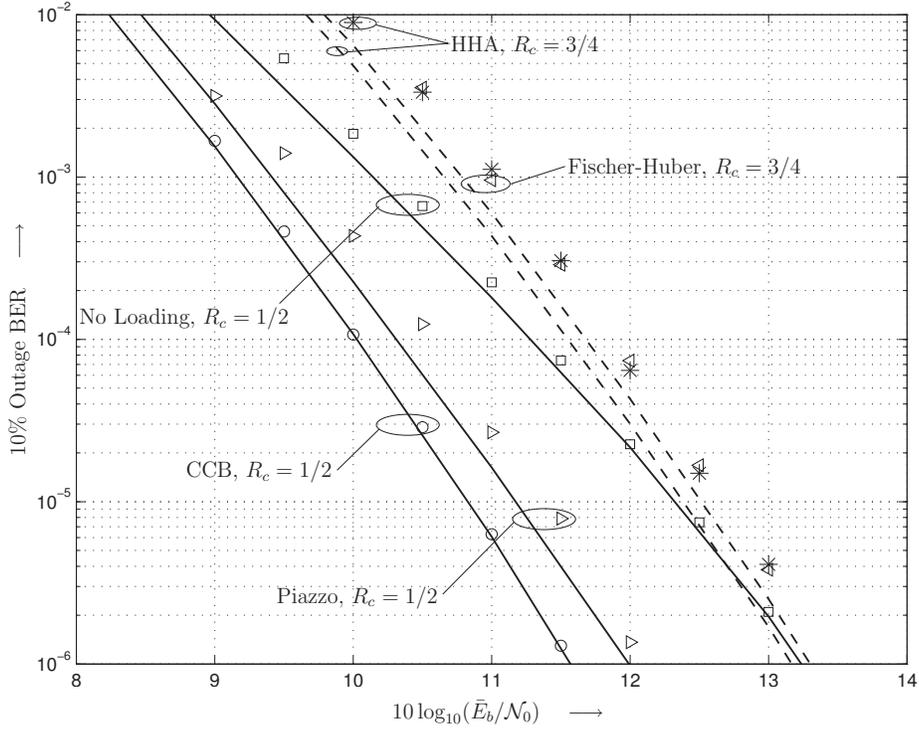


Figure 2.1: 10% Outage BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  from analysis (lines) and simulations (markers) for various combinations of code rates and loading algorithms. MB-OFDM system, UWB CM1 channel.

### 2.5.1 Accuracy of BER Approximation

First, we illustrate the accuracy of the proposed BER approximation for coded and loaded OFDM. To this end, in Figure 2.1, we plot the 10% outage BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  from analysis (lines) as well as the corresponding simulation results (markers) for various combinations of code rates and loading algorithms, for MB-OFDM over the UWB CM1 channel using a set of  $N_c = 100$  channel realizations. (We note that the 10% outage BER is a common performance measure in UWB systems, cf. e.g. [9, 37].) We can see that the simulation results confirm the analysis for all considered code rates and loading algorithms, with a maximum difference of 0.3 dB between simulation and analysis at low BERs.

In Figure 2.2, the 10% outage BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  from analysis (lines) as

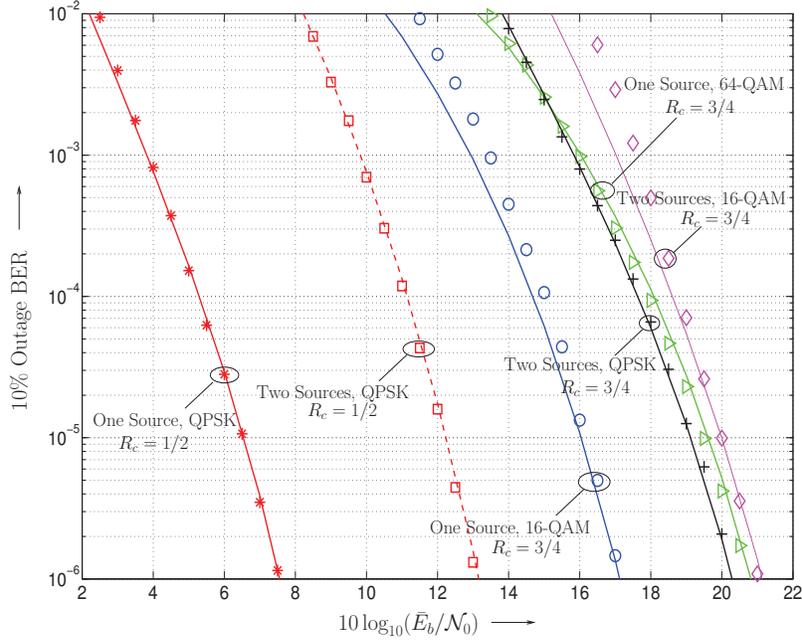


Figure 2.2: 10% Outage BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  from analysis (lines) and simulations (markers) for the MIMO-SVD system.

well as the corresponding simulation results (markers) for 802.11a/g WLAN are plotted for  $N_T = 2$ ,  $N_R = 2$  MIMO-SVD with different constellation sizes and code rates. We consider the case where there is the same information source for all antennas (“one source”) as well as the case where there is a separate information source for each SVD channel (“two sources”). It can be seen that again we have a very good match between simulation and analytical results which confirms the accuracy of our analytical expressions.

We note that obtaining the 10% outage BER via simulation is very time-consuming due to the need to simulate the system separately for each channel realization. On the other hand, the analysis can be performed quite quickly even for large sets of channel realizations.

In the following, we show BER results obtained from the analytical expressions unless

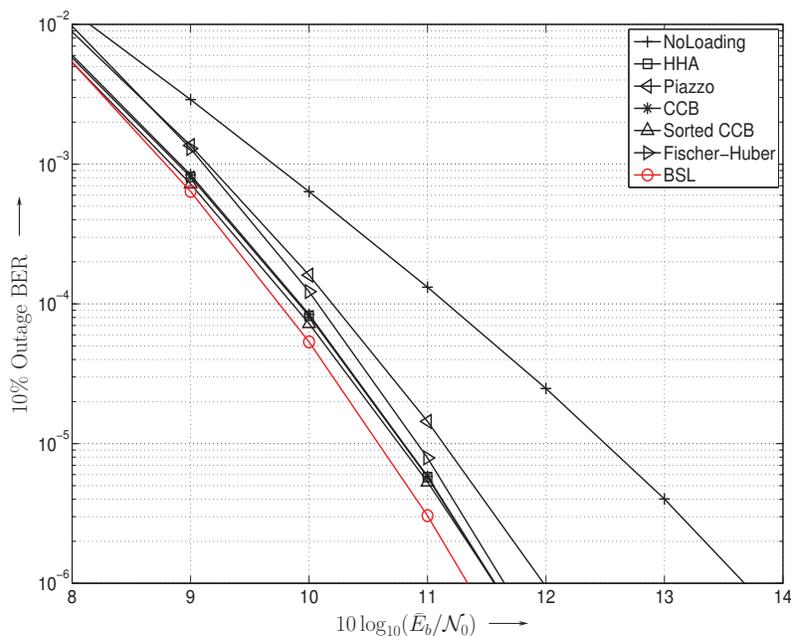


Figure 2.3: 10% Outage BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  from analysis for various loading algorithms. MB-OFDM system with  $R_c = 1/2$ , UWB CM1 channel.

noted otherwise.

## 2.5.2 Bit-Loading for Coded OFDM

Next, we show results for bit-loading for coded OFDM. Firstly, the MB-OFDM system with code rate  $R_c = 1/2$  and transmission for the UWB CM1 is considered. Figure 2.3 shows the 10% outage BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for a number of popular loading algorithms, and the proposed BSL algorithm. As a reference, the BER curve for coded OFDM without loading is also shown. We observe that for this system and channel model, decent gains of approximately 2 dB can be obtained by the application of loading. The Piazzo and Fischer-Huber algorithms provide slightly smaller gains than the HHA and CCB algorithms. The proposed BSL achieves the best performance, with modest additional gains

over the conventional loading algorithms. Of course, this comes at the cost of the increased complexity required to perform the BER analysis for all loadings. Interestingly, the performance of sorted CCB is almost exactly the same as those for HHA and CCB, which suggests use of sorted CCB also for coded OFDM due to its lower computational complexity [42]. The slightly better performance of sorted CCB compared to CCB stems from the fact that these bit loading algorithms are designed for uncoded systems and, as a result, there is no guarantee that they still works good for BICM systems. In order to reduce the computational complexity of BSL, we also performed the BSL using an error vector set of size  $M = 1$ , i.e., we only included the minimum distance error event in the analysis. The results obtained were identical to those in Figure 2.3. This suggests that reduced-complexity BER estimation with small error vector sets could be an attractive method for loading algorithms based on coded BER, such as BSL.

Secondly, we consider the 802.11a/g WLAN system with code rate  $R_c = 1/2$  and the channel model with RMS delay spread of  $T_{\text{rms}} = 250$  ns. In Figure 2.4, we again plot the 10% outage BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for the various loading algorithms. We observe relatively smaller gains due to loading than for the UWB scenario, and the Piazzo loading algorithm performs significantly worse than the other algorithms (see also Table 2.2 and the discussion below). Again, BSL achieves the best performance. Furthermore, the sorted CCB algorithm achieves a performance very similar to those of HHA and CCB.

In order to see more comprehensive statistics for the comparison of the different loading schemes, Figure 2.5 shows the corresponding cumulative distributions of the loading gains  $G = \text{SNR}_{\text{NL}} - \text{SNR}_{\text{Loading}}$  (NL: no loading) required to achieve a BER of  $10^{-5}$ , for the MB-OFDM system with  $R_c = 1/2$  over the UWB CM1 channel ( $\text{SNR} \triangleq 10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$ ). Interestingly, we note that there is a small probability that the Piazzo, Fischer-Huber and Sorted CCB loadings will result in a performance loss (negative gains), while the CCB,

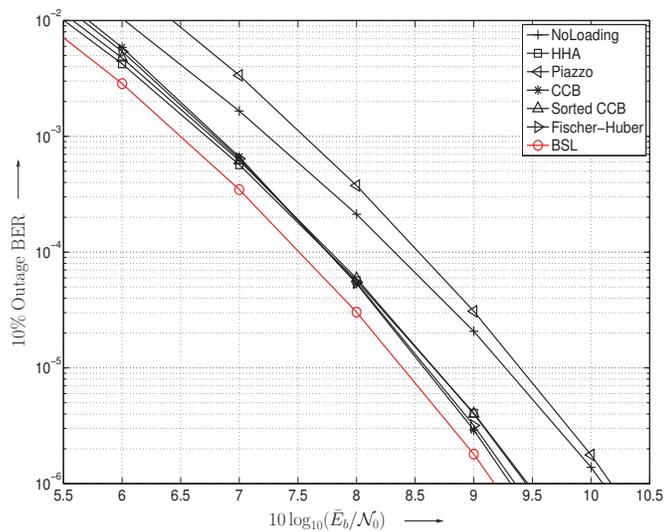


Figure 2.4: 10% outage BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  from analysis for various loading algorithms. 802.11a/g WLAN system with  $R_c = 1/2$ , channel  $T_{\text{rms}} = 250$  ns.

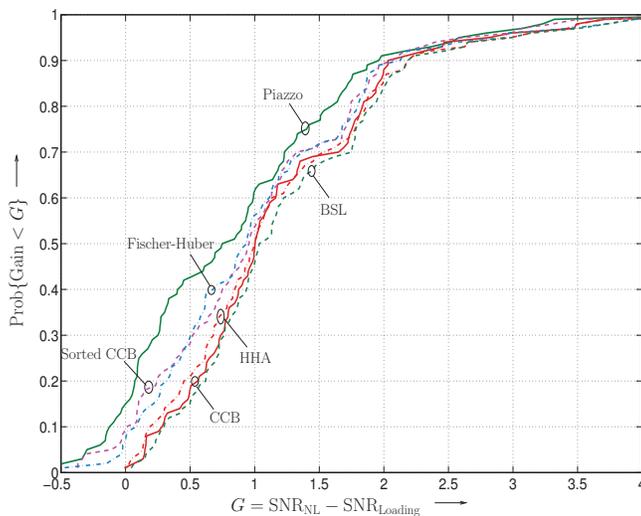


Figure 2.5: Cumulative distribution function of loading gain  $G = \text{SNR}_{\text{NL}} - \text{SNR}_{\text{Loading}}$  required for  $\text{BER} = 10^{-5}$  for different loading schemes (NL: no loading). MB-OFDM system with  $R_c = 1/2$ , UWB CM1 channel.

Table 2.1: Relative use of different loading schemes as best solution. MB-OFDM, UWB CM1 channel.

Loading	% Use ( $R_c = 1/2$ )		% Use ( $R_c = 3/4$ )	
	w/ HHA	w/o HHA	w/ HHA	w/o HHA
No Loading	0	0	0	0
HHA [6, 40]	42	—	86	—
CCB [35]	38	65	7	27
Sorted CCB [42]	7	16	3	10
Piazzo [41]	2	1	0	1
Fischer-Huber [25]	11	18	4	62

HHA, and BSL algorithms always provide a performance gain. We again note that HHA, CCB and Sorted CCB algorithms have similar performances. The BSL always results in the highest loading gain. Finally, we note that gains of up to 4 dB can be expected from loading, while gains of at least 1 dB can be expected for 50% of channel realizations.

Finally, we consider the relative use of different loading schemes for the BSL. Firstly, Table 2.1 lists the results for the case of MB-OFDM with  $R_c = 1/2$  and  $R_c = 3/4$  and the UWB CM1. Since calculating the HHA loading has a high computational complexity compared to the other algorithms, we list the relative use both including and excluding HHA (“w/ HHA” and “w/o HHA”, respectively). We note that for  $R_c = 1/2$ , CCB is the most-used algorithm, while for  $R_c = 3/4$  the Fischer-Huber algorithm is often the best. The Piazzo algorithm is rarely the best loading for either code rate. By comparing the two code rates, we can see that the best loading algorithm is rate dependent, indicating that when deploying coded loaded OFDM systems, some consideration should be given to the loading-coding combination during system design. Secondly, Table 2.2 lists the relative use of different loading schemes in BSL for the 802.11a/g system with  $R_c = 1/2$  and different channel RMS delay spreads  $T_{\text{rms}}$ . Interestingly, we note that for small  $T_{\text{rms}}$  the best loading is often the same modulation for all subcarriers (no loading). This is a result of the lack of

Table 2.2: Relative use of different loading schemes as best solution. 802.11a/g,  $R_c = 1/2$ , of (2.12), (2.13).

Loading	% Use ( $T_{\text{rms}} = 50$ ns)		% Use ( $T_{\text{rms}} = 100$ ns)		% Use ( $T_{\text{rms}} = 250$ ns)	
	w/ HHA	w/o HHA	w/ HHA	w/o HHA	w/ HHA	w/o HHA
No Loading	90	91	57	60	18	19
HHA [6, 40]	2	—	8	—	19	—
CCB [35]	1	2	6	7	14	20
Sorted CCB [42]	6	6	25	28	39	48
Piazzo [41]	0	0	1	1	2	2
Fischer-Huber [25]	1	1	3	4	8	11

variation in subcarrier channel gains due to the small delay spread. As  $T_{\text{rms}}$  increases, the channel gains have more variation and thus there is increased gain from loading. We also note that the best loading is more varied for the WLAN case, indicating again that the choice of a loading algorithm for coded OFDM systems is system-dependent and should be carefully considered during system design.

### 2.5.3 Adaptive Interleaving

In this section, we show results for the three proposed adaptive interleaving algorithms. We consider MB-OFDM with  $R_c = 1/2$  for UWB CM1 and also MIMO IEEE 802.11a/g with  $R_c = 1/2$ .

#### Impact of the Interleaver

In Figure 2.6, we plot the BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for one specific channel realization from analysis (lines) as well as the corresponding simulation results (markers) when 3 different interleavers are used. In particular, the best and worst interleavers are chosen among 1000 randomly generated interleavers, as well as the interleaver prescribed by the

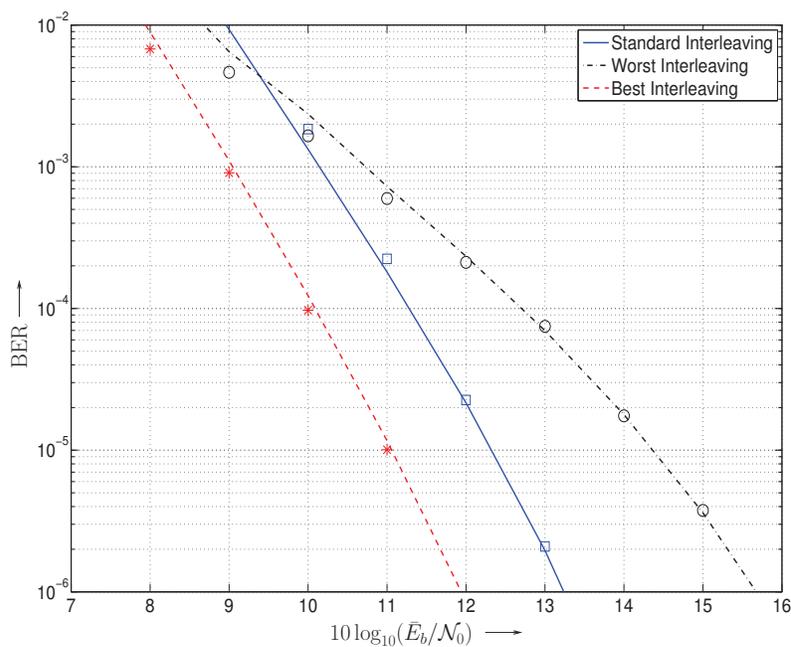


Figure 2.6: Comparing different interleavers for one channel realization. MB-OFDM system with  $R_c = 1/2$ , UWB CM1 channel.

MB-OFDM standard. We can see that interleaver has a great impact on the performance of the BICM systems. At the BER of  $10^{-5}$  there is a difference of about 3.5 dB between the performance for the best and worst interleavers. Therefore, we conclude that analyses such as [20, 21], which consider the ideal interleaver, are not accurate in practical systems.

### Best Interleaver of a Set

Next, we compare set-based adaptive interleaving using several different sets of interleavers. In particular, we compare

- (a) always using the MB-OFDM standard interleaver;
- (b) choosing the interleaver with the lowest BER from a set consisting of 1000 randomly generated bit-interleavers in addition to interleavers designed according to [1] with

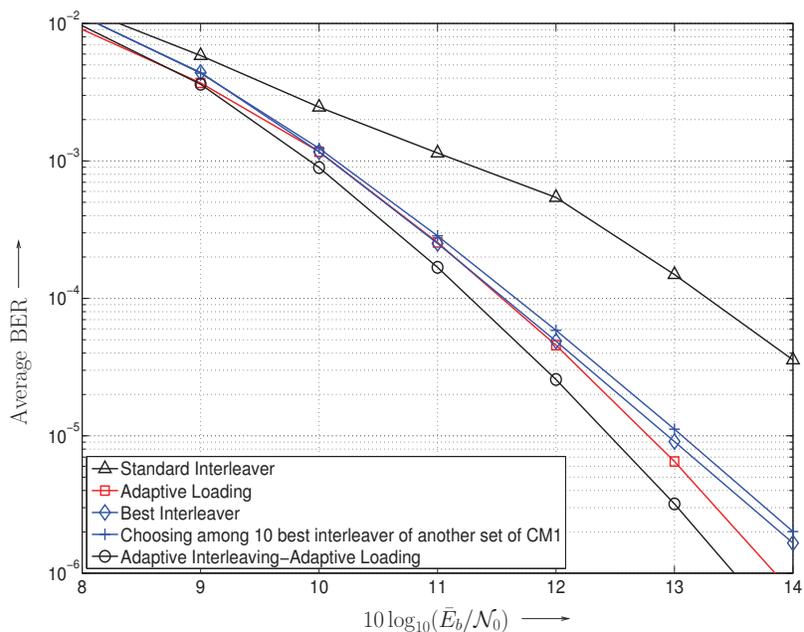


Figure 2.7: Average BER for different set-based adaptive interleaver and adaptive loading schemes. MB-OFDM system with  $R_c = 1/2$ , UWB CM1 channel.

- parameters  $D = 2$  and  $D = 4$  (we refer to this method as “best interleaver”); and
- (c) choosing the interleaver with the lowest BER from a set of 10 bit-interleavers (the 9 best interleavers among random interleavers tested for another set of UWB CM1 channels, and the standard interleaver).

We also consider combined bit-loading and adaptive interleaving, where the interleaver is selected using option (c) above.

In Figure 2.7 we plot the average BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for the different adaptive interleaver schemes. The BER is averaged over  $N_c = 2000$  channel realizations. As a reference, the BER for bit-loading with the fixed standard interleaver is also shown. We observe that adaptive interleaving without bit-loading results in a gain of about 1 dB for the considered scenario. If we restrict the search space to only 10 pre-selected interleavers,

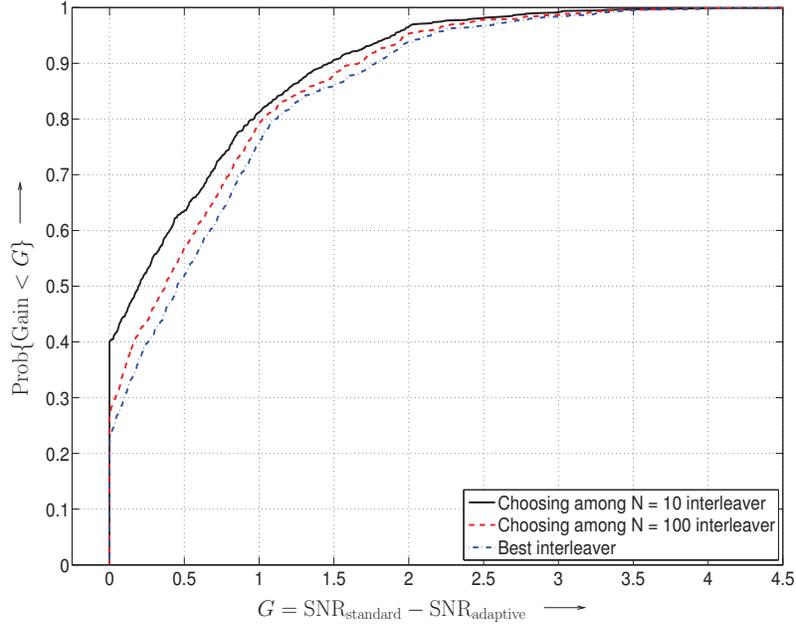


Figure 2.8: Cumulative distribution function of interleaving gain  $G = \text{SNR}_{\text{standard}} - \text{SNR}_{\text{adaptive}}$  required for  $\text{BER} = 10^{-5}$  for different set-based adaptive interleaver schemes. MB-OFDM system with  $R_c = 1/2$ , UWB CM1 channel.

the gain diminishes, especially for lower error rates. We note that we obtained practically the same results if the interleavers were pre-selected based on a different UWB CM, e.g., CM3. Hence, we conclude the diversity of interleavers available, not the pre-selection, is critical for the performance gains with adaptive interleaving. Furthermore, it can be seen that the combination of adaptive interleaving and bit-loading has a small gain. It has been said in [2] that combining bit-loading and adaptive interleaving will result in performance loss compared to plain adaptive interleaving. On the other hand, it can be seen that our scheme always results in performance gain, although the gain might be small.

Figure 2.8 provides further insight by showing the cumulative distribution of the adaptive interleaving gain  $G = \text{SNR}_{\text{standard}} - \text{SNR}_{\text{adaptive}}$ , where  $\text{SNR}_x$  is the SNR required to achieve a BER of  $10^{-5}$ , for different adaptive interleaving schemes for  $N_c = 1000$  channel

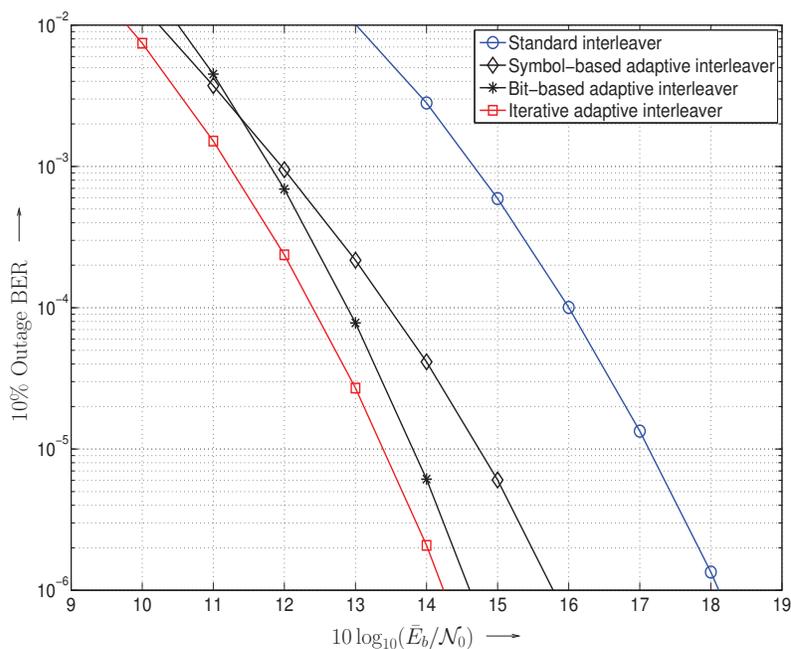


Figure 2.9: 10% Outage BER for different adaptive interleaver schemes: symbol-based adaptive interleaver [1], bit-based adaptive interleaver [2], and our iterative adaptive interleaver algorithm. MB-OFDM system with  $R_c = 1/2$ , 16-QAM modulation, UWB CM1 channel.

realizations. In particular, the effect of the size of the set from which interleavers are selected is highlighted. It can be seen that there is no performance loss using our method. We also observe that gains of up to 4.5 dB can be expected from our proposed adaptive interleaving scheme.

### Iterative Interleaver Improvement

The outage BER for our iterative interleaver improvement algorithm for 16-QAM is plotted in Figure 2.9, compared with the performance of the standard interleaver, symbol-based adaptive interleaving [1], and bit-based adaptive interleaving [2]. Our iterative algorithm starts with an initial interleaver from the bit-based adaptive algorithm. It can be seen

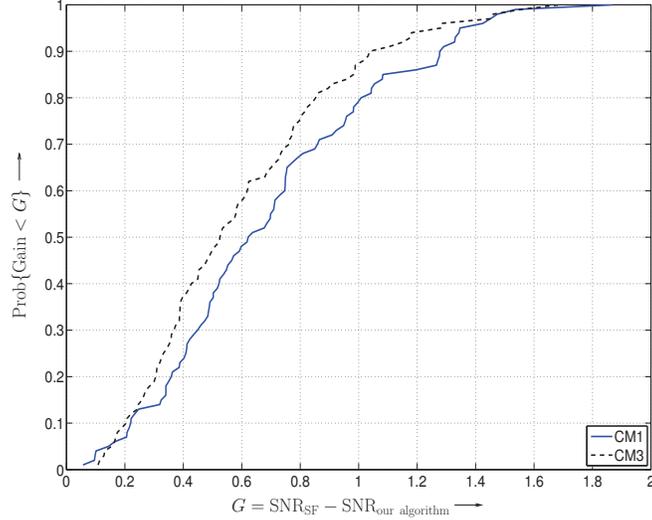


Figure 2.10: Cumulative distribution function of the interleaving gain  $G = \text{SNR}_{\text{SF}} - \text{SNR}_{\text{our algorithm}}$  required for  $\text{BER} = 10^{-5}$  for our iterative adaptive interleaver versus bit-based adaptive interleaving (SF) [2]. MB-OFDM system with  $R_c = 1/2$ , 16-QAM modulation, UWB CM1 and CM3 channels.

that our algorithm will result in 0.5 dB performance gain over the bit-based algorithm, and approximately 1.3 dB gain over the symbol-based algorithm.

In Figure 2.10 the cumulative distribution of the gain  $G = \text{SNR}_{\text{SF}} - \text{SNR}_{\text{our algorithm}}$  for our iterative interleaver improvement method is shown, where  $\text{SNR}_{\text{SF}}$  and  $\text{SNR}_{\text{our algorithm}}$  are the SNRs required to achieve a BER of  $10^{-5}$  for the bit-based adaptive algorithm of Stierstorfer and Fischer (SF) [2], and for our algorithm, respectively. It can be seen that gains of up to 2 dB can be expected from our method for both CM1 and CM3 channels.

### Spatial Interleaving

The average BER for different adaptive interleaving techniques for IEEE 802.11a/g with 2 transmit and 2 receive antennas and two separate information sources are plotted in Figure 2.11. We can see that our proposed spatial interleaving results in about 5 dB per-

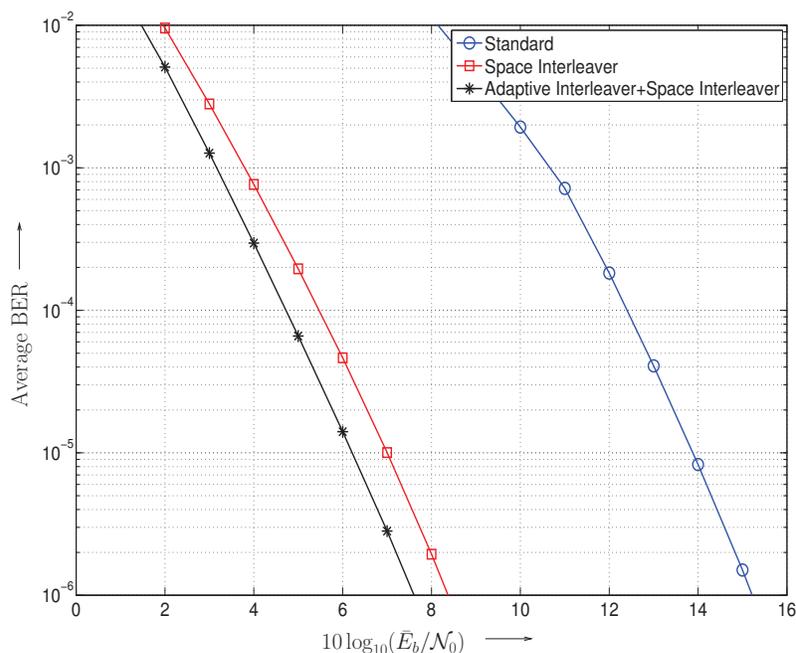


Figure 2.11: Average BER for different adaptive interleaver schemes for MIMO-SVD.  $R_c = 1/2$ , 4-QAM modulation.

formance gain. We can also combine spatial interleaving with other adaptive interleaving techniques. In this figure, we combine spatial interleaving with our set-based adaptive interleaving technique, where the best interleaver for each resulting channel is picked among the 10 best interleavers chosen for another set of channels. By doing so, we gain another 1.65 dB for the BER of  $10^{-5}$ .

## 2.5.4 Adaptive Coded Modulation

In this section, we show results for the proposed adaptive coded modulation scheme considering the 802.11a/g WLAN system with  $T_{\text{rms}} = 250$  ns and  $T_s = 50$  ns. We consider both the SISO system as well as  $2 \times 2$  MIMO-SVD system with only one information source. The set of available code rates and modulations are shown in Table 2.3, with total data

Table 2.3: Different code rates and modulations used for the simulation results.

$R_T$	$R_c$	Modulation
1	1/2	4-QAM
1.33	2/3	4-QAM
1.5	3/4	4-QAM
2	1/2	16-QAM
2.67	2/3	16-QAM
3	3/4	16-QAM
3	1/2	64-QAM
4	2/3	64-QAM
4.5	3/4	64-QAM

rates from 1 bit/symbol to 4.5 bit/symbol. We call each combination of modulation and code rate a “mode”. For the results, we generate a set of 1000 channel realizations and for each channel realization, we select the appropriate mode using our error rate approximation, send the OFDM symbol over the channel, and decode it at the receiver. This allows us to measure the goodput (the successfully transmitted data rate) which has also been considered in [21]. A target word error rate of  $10^{-2}$  is adopted for both schemes.

The goodput versus  $10 \log_{10}(\bar{E}/\mathcal{N}_0)$  per subcarrier is plotted in Figure 2.12. The effect of different numbers of modes is analyzed in this figure. The system with 9 modes denotes the system which uses all the available code rates and constellation sizes. For 6 modes,  $R_c = 2/3$  is not used and for 4 modes, 4-QAM with  $R_c = 3/4$  and 64-QAM with  $R_c = 1/2$  are also not used. It can be seen that using more modes will result in higher goodput because the system has more flexibility, however the resultant gain is dependent on the SNR. This is because in the low SNR region our method tends to select the lowest possible rate, thus having more modes does not help in this region. Simulation results for the 6 modes case are also plotted in this figure for both SISO and MIMO cases. It can be seen that analytical and simulation results perfectly match, again confirming the accuracy of

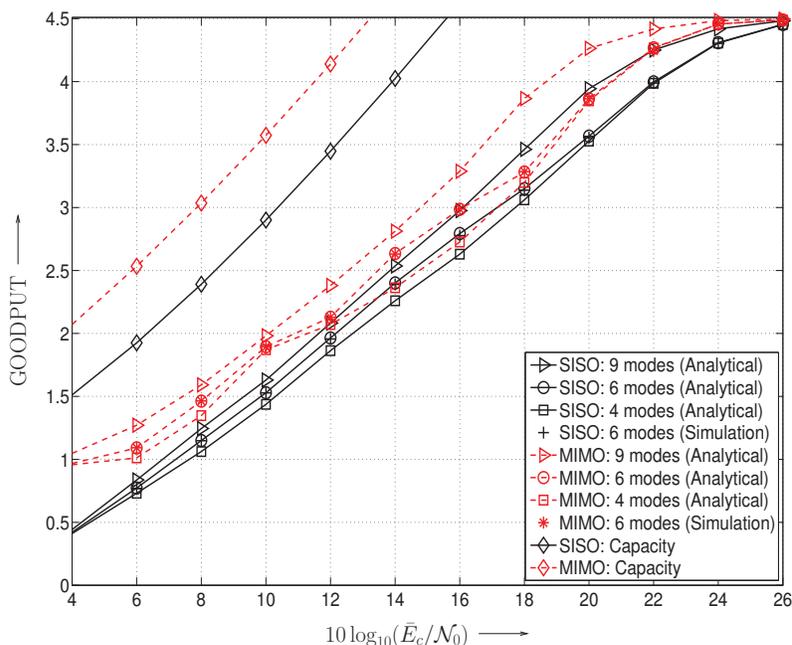


Figure 2.12: Comparing the goodput for different adaptive coded modulation algorithms. IEEE 802.11a/g WLAN system, channel  $T_{\text{rms}} = 250$  ns and  $T_s = 50$  ns.

the analytical results. Finally, we note that the gap to channel capacity for SISO and MIMO OFDM is consistently about 6 dB (using 9 modes), which is a promising result considering that convolutional codes are used.

## 2.6 Summary

In this chapter, we have developed a novel analytical method for BER and WER estimation of bit-loaded coded MIMO-OFDM systems operating over frequency-selective quasi-static channels with non-ideal interleaving. The presented numerical results illustrate that the proposed analysis technique provides an accurate estimation of the BER of loaded BIC-MIMO-OFDM systems. This allows for system performance analysis without resorting

to lengthy simulations. We have put the analysis to use in three different applications, namely bit-loading, adaptive interleaving, and adaptive coded modulation. In the case of bit loading, we have shown that the relative performance of bit-loading algorithms for coded OFDM is system-dependent, and thus some care should be given to the selection of loading algorithms for coded OFDM systems. The proposed “best solution” guarantees the best performance, at a cost of somewhat higher complexity when performing the loading. Adaptive interleaving has been confirmed to be an interesting alternative and addition to bit loading in coded OFDM. Finally, the application of the derived WER expressions to adaptive coded modulation algorithm leads to goodput close to the ultimate limit, while guaranteeing a certain target WER.

# Chapter 3

## Adaptive BIC-OFDM: Performance Analysis and Optimization

### 3.1 Introduction

While the analysis proposed in Chapter 2 enables us to derive the performance for bit-loaded BIC-OFDM, the resulting expression might be too complex in some applications. In this chapter, we present a general framework for performance evaluation of a BIC-OFDM transmission system given a channel realization and interleaver structure. The proposed analysis makes use of the approximations for the probability density function (PDF) of reliability metrics developed in [13]. Different from the results presented in Chapter 2 and [17], our BER expression is not based on analyzing the system for a fixed reference codeword. Based on the derived BER expression, the BER minimization optimization problem in the asymptotically-high signal-to-noise ratio (SNR) regime is formulated. Here, optimization refers to simultaneous power and bit loading and code rate selection. Unfortunately, this optimization problem cannot be solved optimally using any polynomial-time algorithm. For this reason, we derive the expression for the BER averaged over randomly selected interleavers. Using this newly derived BER expression, we are able to formulate BER minimization as a margin maximization problem (MMP) [30, 31], which can be solved using greedy type algorithms. Here, we also derive power minimization and throughput maximization algorithms under BER constraints. Finally, we derive an asymptotic BER

expression for the optimized system.

The rest of this chapter is organized as follows. Section 3.2 introduces the BIC-OFDM transmission system, the approximation for the PDF of reliability metrics, and important types of optimization problems arising in the design of such systems. In Section 3.3, we derive the BER expression for BIC-OFDM system for a specific channel and interleaver realization and simplify it for high SNR. Then, we calculate the BER expression averaged over interleaver realizations and further simplify it in the high SNR regime. Section 3.4 introduces computationally-efficient algorithms for solving the defined optimization problems and also the BER expression for the performance of the optimized system. In Section 3.5, we present a number of representative simulation results confirming the accuracy of the proposed analysis and the performance improvement made by optimizing the transceiver. Finally, we offer concluding remarks in Section 3.6.

## 3.2 Preliminaries

In this section, first the BIC-OFDM transmission system is introduced. More information about each block can be found in Section 1.5. Then, based on the results presented in [13], we introduce a closed-form approximation for the PDF of reliability metrics which will be used in Section 3.3. Finally, we briefly review different optimization problems arising in the design of BIC-OFDM systems.

### 3.2.1 System Model

At the transmitter, a block of  $B$  information bits  $\underline{b} = [b_1, b_2, \dots, b_B]$  is encoded into  $N$  coded bits  $\underline{c} = [c_1, c_2, \dots, c_N]$  using a binary encoder. In this chapter, we consider the set of available encoders  $\mathcal{C}$  as different puncturing patterns applied to a fixed encoder. Then, the output of the encoder is input to a bitwise interleaver  $\underline{\pi}$ . The result  $\underline{c}^\pi = [c_1^\pi, c_2^\pi, \dots, c_N^\pi]$  is

input to a mapper  $\mu : \{0, 1\}^N \rightarrow \underline{\mathcal{X}}$  where  $\underline{\mathcal{X}} = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_L$  and  $\mathcal{X}_i$  denotes the signal constellation used for the  $i$ th OFDM subcarrier and  $L$  is the number of subcarriers. We assume that  $\mathcal{X}_i$  is a quadrature amplitude modulation (QAM) constellation and mapping is done according to binary reflected Gray labeling. With the inverse fast Fourier transform at the transmitter, fast Fourier transform at the receiver, appropriate cyclic prefix processing, and assuming coherent reception, the equivalent baseband channel for the  $i$ th subcarrier is given by

$$y_i = \sqrt{p_i} h_i x_i + z_i, \quad i = 1, \dots, L, \quad (3.1)$$

where  $y_i \in \mathbb{C}$  denotes the received symbol,  $h_i \in \mathbb{R}^+$  is frequency-domain channel gain,  $p_i \in \mathbb{R}^+$  is the allocated power,  $x_i \in \mathcal{X}_i$  is the transmitted symbol, and  $z_i \in \mathbb{C}$  is the noise sample in the  $i$ th subcarrier. We assume that  $z_i$  is additive white Gaussian noise (AWGN) with  $\mathbb{E} \{ \|z_i\|^2 \} = 1$ . Furthermore, under the quasi-static fading assumption channel gains  $h_i$  remain unchanged during the transmission of a few consecutive symbols [17, 20]. Finally, the total power allocated to the  $L$  subcarriers is given by

$$\sum_{i=1}^L p_i = P. \quad (3.2)$$

At the receiver, the demapper outputs  $M_i = \log_2(|\mathcal{X}_i|)$  bitwise reliability metrics, in the form of log-likelihood ratios (LLRs), corresponding to the coded bits transmitted over the  $i$ th subcarrier. In practice, for ease of implementation, the LLR is approximated using the max-log simplification [19, 43]

$$\lambda_{i,j} = - \min_{a \in \mathcal{X}_{i,j}^1} \|y_i - \sqrt{p_i} h_i a\|^2 + \min_{a \in \mathcal{X}_{i,j}^0} \|y_i - \sqrt{p_i} h_i a\|^2, \quad (3.3)$$

where the second subscript of  $\lambda_{i,j}$  indicates the position of the bit in the binary label of the constellation under consideration and  $\mathcal{X}_{i,j}^b$  denotes the set of symbols in  $\mathcal{X}_i$  with the

$j$ th bit in the binary label fixed to  $b$ .

It is known that (3.3) provides practically maximum-likelihood (ML) decoding performance [44, 45]. Therefore, we adopt this simple metric expression. In slight abuse of terminology, we will also refer to  $\lambda_{i,j}$  in (3.3) as LLR. Finally, the LLRs are deinterleaved into  $\lambda_{i,j}^{\pi^{-1}}$  and input to the ML decoder of the binary code in order to retrieve the information bits.

### 3.2.2 Analytical Approximation for the PDF of LLRs

Recently, a novel approximation for PDF of LLRs given in (3.3) has been proposed in [13]. It is shown that this PDF approximation is accurate in the SNR range in which the BER union bound converges to the actual performance of the transmission system, and therefore it is useful for analyzing systems with moderate-complexity coding schemes, e.g., convolutional coded systems. In the following, based on the results in [13], we represent the PDF expressions for Gray-mapped square QAM constellations. We assume that the symmetry property holds and thus without loss of generality we consider the transmission of  $c^\pi = 1$  [19].

We define the LLR random variable  $\Lambda_{i,j}$  as [46]

$$\Lambda_{i,j} = \log \left( \frac{\Pr(\tilde{c}^\pi = c_{i,j}^\pi | \underline{\eta})}{\Pr(\tilde{c}^\pi = \bar{c}_{i,j}^\pi | \underline{\eta})} \right), \quad (3.4)$$

where  $\underline{\eta} = [y_i, h_i, p_i, \mathcal{X}_i]$  and

$$\Pr(\tilde{c}^\pi = c_{i,j}^\pi | \underline{\eta}) \propto \sum_{x \in \mathcal{X}_{i,j}^{c_{i,j}^\pi}} \exp(-|y_i - \sqrt{p_i} h_i x|^2), \quad (3.5)$$

and  $c_{i,j}^\pi$  is the interleaved coded bit transmitted in the  $j$ th position of the binary label of the symbol transmitted over the  $i$ th subcarrier and  $\bar{c}_{i,j}^\pi$  denotes its complement. Using the

Table 3.1: Values of the parameters used in the approximation for the PDF of LLRs. Only nonzero coefficients are shown here.

Constellation	$n$	$\beta$	$d_{\min}$
BPSK	1	$\beta_{1,1} = 1$	2
QPSK	1	$\beta_{1,1} = 1, \beta_{2,1} = 1$	$\sqrt{2}$
16-QAM	2	$\beta_{1,1} = 0.5, \beta_{1,2} = 0.5, \beta_{2,1} = 1, \beta_{3,1} = 0.5, \beta_{3,2} = 0.5,$ $\beta_{4,1} = 1$	$\frac{2}{\sqrt{10}}$
64-QAM	3	$\beta_{1,1} = 0.5, \beta_{1,2} = 0.25, \beta_{1,3} = 0.25, \beta_{2,1} = 0.75,$ $\beta_{2,2} = 0.25, \beta_{3,1} = 0.5, \beta_{3,2} = 0.25, \beta_{3,3} = 0.25$	$\frac{2}{\sqrt{42}}$

result from [13], it immediately follows that we can approximate the PDF of  $\Lambda_{i,j}$  by

$$f_{\Lambda_{i,j}}(\lambda) = \sum_{k=1}^{n(\mathcal{X}_i)} \beta_{j,k}^{\mathcal{X}_i} \mathcal{N}(\lambda; a_{i,k}, 2a_{i,k}) , \quad (3.6)$$

where  $a_{i,k} = (kd_{\min}^{\mathcal{X}_i})^2 p_i h_i^2$ ,  $d_{\min}^{\mathcal{X}_i}$  denotes the minimum Euclidean distance of constellation  $\mathcal{X}_i$ , and  $n(\mathcal{X}_i)$  and  $\beta_{j,k}^{\mathcal{X}_i}$  are parameters which depend on the constellation and bit position, which we summarized in Table 3.1 for some important constellations for convenience<sup>3</sup>, and

$$\mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) . \quad (3.7)$$

### 3.2.3 Optimization Problems For BIC-OFDM

There are three popular classes of optimization problems for loading and code-rate selection in OFDM as shown in Figure 3.1.

<sup>3</sup>Since Gray-mapped QAM constellations with size greater than 16 are not unique, we have used ([000, 001, 011, 111, 101, 100, 110, 010]) for 64-QAM.

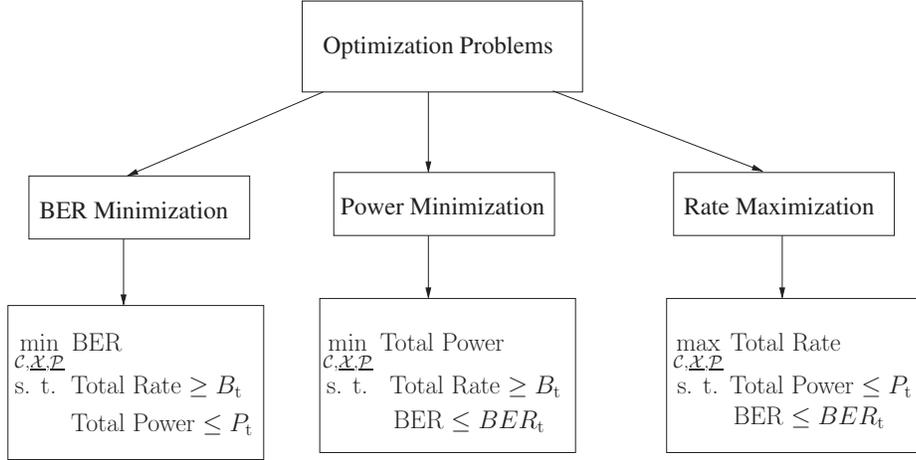


Figure 3.1: Different optimization problems in BIC-OFDM.

### BER Minimization

The BER is minimized given the power budget  $P_t$  and the target rate  $B_t$ . This BER minimization problem can be formulated as

$$\begin{aligned}
 \min_{\mathcal{C}, \underline{\mathcal{X}}, \underline{p}, \underline{\pi}} \quad & P_b(\underline{\theta}) \\
 \text{s.t.} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i \geq B_t, \\
 & \sum_{i=1}^L p_i \leq P_t,
 \end{aligned} \tag{3.8}$$

where  $P_b(\underline{\theta})$  denotes the BER given a code  $\mathcal{C}$ , an interleaver  $\underline{\pi}$ , a vector of constellations  $\underline{\mathcal{X}} = [\mathcal{X}_1 \dots \mathcal{X}_L]$ , a vector of powers  $\underline{p} = [p_1 \dots p_L]$ , and the channel gains  $\underline{h} = [h_1 \dots h_L]$ , which are collected in the parameter vector  $\underline{\theta} = [\mathcal{C}, \underline{\pi}, \underline{\mathcal{X}}, \underline{p}, \underline{h}]$ , and  $R(\mathcal{C})$  is the code rate. Note that the optimization (3.8) includes the selection of the code  $\mathcal{C}$ .

### Power Minimization

The aggregate transmitted power is to be minimized conditioned on the target rate  $B_t$  and the BER of  $BER_t$ . This power minimization problem can be formulated as

$$\begin{aligned}
 \min_{\mathcal{C}, \underline{\mathcal{X}}, \underline{p}} \quad & \sum_{i=1}^L p_i \\
 \text{s.t.} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i \geq B_t , \\
 & P_b(\underline{\theta}) \leq BER_t .
 \end{aligned} \tag{3.9}$$

### Throughput Maximization

Given a target BER and a power budget, it is desired to maximize the throughput of the system, i.e.

$$\begin{aligned}
 \max_{\mathcal{C}, \underline{\mathcal{X}}, \underline{p}} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i \\
 \text{s.t.} \quad & \sum_{i=1}^L p_i \leq P_t , \\
 & P_b(\underline{\theta}) \leq BER_t .
 \end{aligned} \tag{3.10}$$

The algorithm for solving each problem is presented in Section 3.4.2.

## 3.3 Expressions for the Performance of BIC-OFDM

In this section, first we derive an expression for the BER of BIC-OFDM. Then, based on this expression, the BER minimization problem (3.8) is formulated. Unfortunately, this optimization problem is not solvable in polynomial-time in general. Therefore, we next consider the performance of BIC-OFDM averaged over random interleavers. The resulting BER expression leads to formulations of the optimization problems from Section 3.2.3,

which can be solved computationally efficiently.

### 3.3.1 Known Interleaver

#### Performance Evaluation

Given a convolutional code  $\mathcal{C}$ , an interleaver  $\pi$ , and vectors of constellations  $\underline{\mathcal{X}}$ , powers  $\underline{p}$ , and channel gains  $\underline{h}$ , the BER union bound can be expressed as

$$P_b(\underline{\theta}) \leq \frac{1}{B2^B} \sum_{\underline{c}} \sum_{d_H, m, n} w(\mathcal{C}, \underline{\nu}) P_e(\underline{\theta}, \underline{c}, \underline{\nu}), \quad (3.11)$$

where the first sum runs over all codewords  $\underline{c}$ , and the second sum runs over Hamming distances  $d_H$  of error vectors, starting positions  $m$  of error events, and all error vectors for given  $(m, d_H)$ . Hence, the triple  $\underline{\nu} = [d_H, m, n]$  defines one specific error event. Furthermore,  $w(\mathcal{C}, \underline{\nu})$  denotes the input weight for the error event specified by  $\underline{\nu}$ , and  $P_e(\underline{\theta}, \underline{c}, \underline{\nu})$  is the probability of this event.

We now make the assumptions that (i) no two or more non-zero binary symbols belonging to an error vector are mapped to the same OFDM subcarrier and (ii) that the interleaved coded bits mapped to a signal point are uniformly i.i.d. These assumptions are necessary for the following analysis and are in the spirit of the usual “ideal” interleaving assumption. The first assumption is easily guaranteed with practical interleavers for short error events, which dominate the performance, and the second property is well approximated by interleaved convolutional codes. Then we can re-write (3.11) as

$$P_b(\underline{\theta}) \leq \frac{1}{B} \sum_{\underline{\nu}} \frac{w(\mathcal{C}, \underline{\nu})}{\prod_{k=1}^{d_H} M_{s_k(\underline{\nu})}} \sum_{k=1, \dots, d_H} P_e(\underline{\theta}, \underline{c}^{s_k(\underline{\nu})}, \underline{\nu}), \quad (3.12)$$

where  $s_k(\underline{\nu})$  is the subcarrier index for the  $k$ th non-zero element of the error vector specified

by  $\underline{\nu}$  and  $\underline{c}^{s_k(\underline{\nu})}$  is the binary  $M_{s_k(\underline{\nu})}$ -tuple of bits transmitted over subcarrier  $s_k(\underline{\nu})$ .

We define  $b_k(\underline{\nu})$  as the position of the  $k$ th non-zero bit of the error vector specified by  $\underline{\nu}$  in the binary label of  $\mathcal{X}_{s_k(\underline{\nu})}$  and also

$$\Delta(\underline{\theta}, \underline{\nu}) = \sum_{k=1}^{d_H} \Lambda_{s_k(\underline{\nu}), b_k(\underline{\nu})} . \quad (3.13)$$

It can be shown that

$$\begin{aligned} \bar{P}_e(\underline{\theta}, \underline{\nu}) &= \frac{1}{\prod_{k=1}^{d_H} M_{s_k(\underline{\nu})}} \sum_{\underline{c}^{s_k(\underline{\nu})}} P_e(\underline{\theta}, \underline{c}^{s_k(\underline{\nu})}, \underline{\nu}) \\ &= \Pr(\Delta(\underline{\theta}, \underline{\nu}) < 0) . \end{aligned} \quad (3.14)$$

Assuming ideal interleaving, random variables  $\Lambda_{s_k(\underline{\nu}), b_k(\underline{\nu})}$  for  $k = 1, \dots, d_H$  are independent. As a result, PDF of  $\Delta(\underline{\theta}, \underline{\nu})$  can be calculated as the convolution of the PDFs of  $\Lambda_{s_k(\underline{\nu}), b_k(\underline{\nu})}$  as

$$\begin{aligned} f_{\Delta(\underline{\theta}, \underline{\nu})}(\lambda) &= f_{\Lambda_{s_1(\underline{\nu}), b_1(\underline{\nu})}}(\lambda) \otimes \dots \otimes f_{\Lambda_{s_{d_H}(\underline{\nu}), b_{d_H}(\underline{\nu})}}(\lambda) \\ &= \sum_{l_1=1}^{n(\mathcal{X}_{s_1(\underline{\nu})})} \dots \sum_{l_{d_H}=1}^{n(\mathcal{X}_{s_{d_H}(\underline{\nu})})} \left[ \prod_{k=1}^{d_H} \beta_{b_k(\underline{\nu}), l_k}^{\mathcal{X}_{s_k(\underline{\nu})}} \right] \mathcal{N} \left( \lambda; \sum_{k=1}^{d_H} a_{s_k(\underline{\nu}), l_k}, 2 \sum_{k=1}^{d_H} a_{s_k(\underline{\nu}), l_k} \right) \end{aligned} \quad (3.15)$$

where  $\otimes$  denotes the convolution operator. By using (3.15), (3.14) can be calculated as

$$\begin{aligned} \bar{P}_e(\underline{\theta}, \underline{\nu}) &= \int_{-\infty}^0 f_{\Delta(\underline{\theta}, \underline{\nu})}(x) dx \\ &= \sum_{l_1=1}^{n(\mathcal{X}_{s_1(\underline{\nu})})} \dots \sum_{l_{d_H}=1}^{n(\mathcal{X}_{s_{d_H}(\underline{\nu})})} \left[ \prod_{k=1}^{d_H} \beta_{b_k(\underline{\nu}), l_k}^{\mathcal{X}_{s_k(\underline{\nu})}} \right] Q \left( \sqrt{\frac{1}{2} \sum_{k=1}^{d_H} a_{s_k(\underline{\nu}), l_k}} \right) . \end{aligned} \quad (3.16)$$

Using (3.16) in (3.12), the approximation

$$P_b(\underline{\theta}) \leq \frac{1}{B} \sum_{\underline{\nu}} w(\mathcal{C}, \underline{\nu}) \bar{P}_e(\underline{\theta}, \underline{\nu}) \quad (3.17)$$

for the BER union bound becomes closed form. We note that evaluation of (3.17) becomes computationally feasible when the sum is limited to the dominant error event with small  $d_H$ .

In the high SNR regime, (3.16) converges to

$$\bar{P}_e(\underline{\theta}, \underline{\nu}) \approx \left[ \prod_{k=1}^{d_H} \beta_{b_k(\underline{\nu}),1}^{\mathcal{X}_{s_k(\underline{\nu})}} \right] Q \left( \sqrt{\frac{1}{2} \sum_{k=1}^{d_H} a_{s_k(\underline{\nu}),1}} \right), \quad (3.18)$$

from which we define the following parameter

$$d_E(\underline{\theta}, \underline{\nu}) \triangleq \sum_{k=1}^{d_H} a_{s_k(\underline{\nu}),1} \quad (3.19)$$

Now we define  $d_E^*(\underline{\theta})$  as follows:

$$d_E^*(\underline{\theta}) = \min_{\underline{\nu}} d_E(\underline{\theta}, \underline{\nu}) \quad (3.20)$$

and the solution of (3.20) is denoted by  $\underline{\nu}^*$  which is the dominant error event. By only using the dominant term in (3.18),  $P_b(\underline{\theta})$  in high SNR regime can be estimated as

$$P_b(\underline{\theta}) \approx \frac{1}{B} w(\mathcal{C}, \underline{\nu}^*) \left[ \prod_{k=1}^{d_H} \beta_{b_k(\underline{\nu}^*),1}^{\mathcal{X}_{s_k(\underline{\nu}^*)}} \right] Q \left( \sqrt{\frac{1}{2} d_E^*(\underline{\theta})} \right) \quad (3.21)$$

### BER Minimization

Making use of the BER expression in (3.21), the optimization problem defined in (3.8) is equivalent to

$$\begin{aligned}
 \max_{\mathcal{C}, \mathcal{X}, \underline{p}, \pi} \quad & \min_{\underline{\nu}} d_E(\underline{\theta}, \underline{\nu}) \\
 \text{s.t.} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i \geq B_t, \\
 & \sum_{i=1}^L p_i \leq P_t,
 \end{aligned} \tag{3.22}$$

where  $d_E(\underline{\theta}, \underline{\nu})$  is defined in (3.19).

The optimization problem defined in (3.22), to the best of our knowledge, cannot be solved optimally using any polynomial-time algorithm. The only exception for which (3.22) can be formulated as a convex optimization problem (and only for QPSK and BPSK transmission) is the power allocation problem, cf. [27] which is not of much practical interest, and also the computational complexity of the solution renders it unsuitable for practical purposes [27].

### 3.3.2 Random Interleaver

While (3.16) and (3.17) enable a performance evaluation for a given parameter  $\underline{\theta}$ , it is not directly useful for adaptive BIC-OFDM as explained earlier. A pragmatic approach to overcome this difficulty is to consider the average performance for an ensemble of interleavers which is equivalent of assuming a random interleaver. In the following, we derive an expression for BER assuming a random interleaver.

We assume that interleavers are selected uniformly from the set of  $N!$  possible interleavers which was also considered in [20] and [26]. Hence, we are interested in an expression

for

$$P_b(\underline{\xi}) = \mathbb{E}_{\underline{\pi}}\{P_b(\underline{\theta})\} = \mathbb{E}_{\underline{\pi}}\{P_b([\underline{\xi}, \underline{\pi}])\}, \quad (3.23)$$

where  $\mathcal{C}$ ,  $\underline{\mathcal{X}}$ ,  $\underline{p}$ , and  $\underline{h}$  are collected in the parameter vector  $\underline{\xi} = [\mathcal{C}, \underline{\mathcal{X}}, \underline{p}, \underline{h}]$ . Using (3.17) we obtain

$$\begin{aligned} P_b(\underline{\xi}) &\leq \mathbb{E}_{\underline{\pi}} \left\{ \frac{1}{B} \sum_{\underline{\nu}} w(\mathcal{C}, \underline{\nu}) \bar{P}_e(\underline{\theta}, \underline{\nu}) \right\} \\ &\stackrel{(a)}{=} \sum_{d_H} \left( \frac{1}{B} \sum_m \sum_n w(\mathcal{C}, \underline{\nu}) \right) \mathbb{E}_{\underline{\pi}} \{ \bar{P}_e(\underline{\theta}, \underline{\nu}) \} \\ &= \sum_{d_H} w(\mathcal{C}, d_H) \bar{P}_e(d_H, \underline{\xi}), \end{aligned} \quad (3.24)$$

where we have defined

$$w(\mathcal{C}, d_H) = \frac{1}{B} \sum_m \sum_n w(\mathcal{C}, \underline{\nu}), \quad (3.25)$$

which is the distance spectrum of the code  $\mathcal{C}$ , and

$$\bar{P}_e(d_H, \underline{\xi}) = \mathbb{E}_{\underline{\pi}} \{ \bar{P}_e(\underline{\theta}, \underline{\nu}) \}. \quad (3.26)$$

To show that step (a) holds in (3.24), i.e.,  $\mathbb{E}_{\underline{\pi}} \{ \bar{P}_e(\underline{\theta}, \underline{\nu}) \}$  is independent of parameters  $m$  and  $n$ , we use the alternative expression of the Gaussian Q-function to write (3.16) as

$$\begin{aligned} \bar{P}_e(\underline{\theta}, \underline{\nu}) &= \frac{1}{\pi} \int_0^{\pi/2} \sum_{l_1=1}^{n(\mathcal{X}_{s_1}(\underline{\nu}))} \cdots \sum_{l_{d_H}=1}^{n(\mathcal{X}_{s_{d_H}}(\underline{\nu}))} \left[ \prod_{k=1}^{d_H} \beta_{b_k(\underline{\nu}), l_k}^{\mathcal{X}_{s_k}(\underline{\nu})} \exp \left( -\frac{a_{s_k(\underline{\nu}), l_k}}{4 \sin^2 \phi} \right) \right] d\phi \\ &= \frac{1}{\pi} \int_0^{\pi/2} \prod_{k=1}^{d_H} \sum_{l=1}^{n(\mathcal{X}_{s_k}(\underline{\nu}))} \beta_{b_k(\underline{\nu}), l}^{\mathcal{X}_{s_k}(\underline{\nu})} \exp \left( -\frac{a_{s_k(\underline{\nu}), l}}{4 \sin^2 \phi} \right) d\phi. \end{aligned} \quad (3.27)$$

Then, (3.26) can be developed as

$$\begin{aligned}
 \bar{P}_e(d_H, \underline{\xi}) &= \mathbb{E}_{\pi} \left\{ \bar{P}_e(\underline{\theta}, \underline{\nu}) \right\} \\
 &= \frac{1}{\pi} \int_0^{\pi/2} \mathbb{E}_{\pi} \left\{ \prod_{k=1}^{d_H} \sum_{l=1}^{n(\mathcal{X}_{s_k(\underline{\nu})})} \beta_{b_k(\underline{\nu}), l}^{\mathcal{X}_{s_k(\underline{\nu})}} \exp \left( -\frac{a_{s_k(\underline{\nu}), l}}{4 \sin^2 \phi} \right) \right\} d\phi \\
 &\stackrel{(a)}{\approx} \frac{1}{\pi} \int_0^{\pi/2} \prod_{k=1}^{d_H} \mathbb{E}_{s_k(\underline{\nu}), b_k(\underline{\nu})} \left\{ \sum_{l=1}^{n(\mathcal{X}_{s_k(\underline{\nu})})} \beta_{b_k(\underline{\nu}), l}^{\mathcal{X}_{s_k(\underline{\nu})}} \exp \left( -\frac{a_{s_k(\underline{\nu}), l}}{4 \sin^2 \phi} \right) \right\} d\phi \\
 &\stackrel{(b)}{=} \frac{1}{\pi} \int_0^{\pi/2} \prod_{k=1}^{d_H} \left[ \frac{1}{N} \sum_{i=1}^L \sum_{j=1}^{M_i} \sum_{l=1}^{n(\mathcal{X}_i)} \beta_{j, l}^{\mathcal{X}_i} \exp \left( -\frac{a_{i, l}}{4 \sin^2 \phi} \right) \right] d\phi \\
 &\stackrel{(c)}{=} \frac{1}{\pi} \int_0^{\pi/2} \left[ \frac{1}{N} \sum_{i=1}^L \sum_{j=1}^{M_i} \sum_{l=1}^{n(\mathcal{X}_i)} \beta_{j, l}^{\mathcal{X}_i} \exp \left( -\frac{a_{i, l}}{4 \sin^2 \phi} \right) \right]^{d_H} d\phi
 \end{aligned} \tag{3.28}$$

where (a) makes the approximation of independent assignments of interleaving positions for each pair  $(s_k(\underline{\nu}), b_k(\underline{\nu}))$ , whose effect is negligible for large enough interleavers, (b) is due to the assumption of uniform interleaving, and (c) follows from the independence of the arguments of the product.

Expanding (3.28) using the multinomial series representation [47, p. 823], the dominating terms of  $\bar{P}_e(d_H, \underline{\xi})$  can be expressed as a superposition of Gaussian Q-functions:

$$\bar{P}_e(d_H, \underline{\xi}) = \sum_{i=1}^L \left( \frac{1}{N} \sum_{j=1}^{M_i} \beta_{j, 1}^{\mathcal{X}_i} \right)^{d_H} Q \left( \sqrt{\frac{d_H a_{i, 1}}{2}} \right), \tag{3.29}$$

which can be used in (3.24) to calculate the BER union bound.

$P_b(\underline{\xi})$  can be further simplified by only considering the dominating term of the series, which is

$$\max_i \left\{ \left( \frac{1}{N} \sum_{j=1}^{M_i} \beta_{j, 1}^{\mathcal{X}_i} \right)^{d_H} Q \left( \sqrt{\frac{d_H a_{i, 1}}{2}} \right) \right\}, \tag{3.30}$$

and this maximization is equivalent to

$$\min_i \{a_{i,1}\}, \quad (3.31)$$

in high SNR and the solution is denoted by  $i^*$ . Consequently, we use the asymptotic BER approximation

$$P_b(\underline{\xi}) \approx w(\mathcal{C}, d_{\text{free}}(\mathcal{C})) \left( \frac{1}{N} \sum_{j=1}^{M_{i^*}} \beta_{j,1}^{X_{i^*}} \right)^{d_{\text{free}}(\mathcal{C})} Q \left( \sqrt{\frac{d_{\text{free}}(\mathcal{C}) a_{i^*,1}}{2}} \right), \quad (3.32)$$

where  $d_{\text{free}}(\mathcal{C})$  is the code's free distance.

## 3.4 Solutions to the Optimization Problems

Making use of the analysis results from the previous section, in this section we first derive a simple formula for the performance of the optimized systems. Then, we reformulate the optimization problems defined in Section 3.2.3 and present algorithms to find the optimal solutions for these revised problems.

### 3.4.1 Performance Evaluation for Optimized System

In this section, we develop a low-complexity method for performance analysis of the optimized BIC-OFDM system. We will show in the next section that in the optimized system,  $a_{i,1}$  is a constant which we denote by  $E$ , i.e.  $a_{i,1} = E$  for  $1 \leq i \leq L$ . Therefore, (3.28) becomes

$$\bar{P}_e(d_H, \underline{\xi}) = \frac{1}{\pi} \int_0^{\pi/2} \left[ \sum_{l=1}^{l_{\max}} \epsilon_l \exp \left( -\frac{l^2 E}{4 \sin^2 \phi} \right) \right]^{d_H} d\phi, \quad (3.33)$$

where we have defined the two parameters<sup>4</sup>

$$l_{\max} = \max_{1 \leq i \leq L} n(\mathcal{X}_i), \quad (3.34)$$

$$\epsilon_l = \frac{1}{N} \sum_{i=1}^L \sum_{j=1}^{M_i} \beta_{j,l}^{\mathcal{X}_i}. \quad (3.35)$$

Using the multinomial series expansion [47, p. 823], (3.33) can be written as

$$\begin{aligned} \bar{P}_e(d_H, \underline{\xi}) &= \frac{1}{\pi} \int_0^{\pi/2} \sum_{\substack{k_1, \dots, k_{l_{\max}} \\ k_1 + \dots + k_{l_{\max}} = d_H}} \left( \frac{d_H!}{\prod_{l=1}^{l_{\max}} k_l!} \right) \left( \prod_{l=1}^{l_{\max}} \epsilon_l^{k_l} \right) \exp \left( - \left[ \sum_{l=1}^{l_{\max}} k_l l^2 \right] \frac{E}{4 \sin^2 \phi} \right) d\phi \\ &= \sum_{\substack{k_1, \dots, k_{l_{\max}} \\ k_1 + \dots + k_{l_{\max}} = d_H}} \left( \frac{d_H!}{\prod_{l=1}^{l_{\max}} k_l!} \right) \left( \prod_{l=1}^{l_{\max}} \epsilon_l^{k_l} \right) Q \left( \sqrt{\left[ \sum_{l=1}^{l_{\max}} k_l l^2 \right] \frac{E}{2}} \right). \end{aligned} \quad (3.36)$$

Taking into account the dominant term in high SNR, we reach at the following asymptotic result

$$\bar{P}_e(d_H, \underline{\xi}) \stackrel{E \rightarrow \infty}{\approx} \epsilon_1^{d_H} Q \left( \sqrt{\frac{E d_H}{2}} \right). \quad (3.37)$$

Expression (3.36) can be used in (3.24) to obtain a closed-form union bound BER approximation for the optimized system. In the case of asymptotically high SNR, expression (3.37) is used in (3.24) considering only  $d_H = d_{\text{free}}$  meaning that we can derive the performance of the optimized system using a simple formula.

### 3.4.2 Optimization Problems

The algorithm for solving the optimization problems in Section 3.2.3 are presented here.

---

<sup>4</sup>Note that  $a_{i,l} = l^2 a_{i,1}$

### Margin Maximization Problem

We start by reviewing the so called margin maximization problem (MMP) which is proposed in [30, 31, 40]. The MMP can be formulated according to

$$\begin{aligned}
 \min_{\underline{x}, \underline{p}} \quad & \sum_{i=1}^L p_i \\
 \text{s.t.} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i \geq B_t, \\
 & E_i = E_t, \quad \forall i \in \{1, \dots, L\}.
 \end{aligned} \tag{3.38}$$

where  $E_i$  denotes a parameter which solely depends on the power allocated to the subcarrier  $p_i$  and the constellation  $b_i$ <sup>5</sup> which in our problem is

$$E_i = a_{i,1}. \tag{3.39}$$

Next, we define  $\Delta(\mathcal{P}_i(k))$  according to

$$\Delta(\mathcal{P}_i(k)) = \begin{cases} \mathcal{P}_i(k) - \mathcal{P}_i(k-1) & \text{if } k \geq 1 \\ 0 & \text{if } k = 0 \end{cases} \tag{3.40}$$

where  $\mathcal{P}_i(k)$  (or  $\mathcal{P}_i(k-1)$ ) is the required power such that the conditions  $E_i = E_t$  and  $M_i = k$  (or  $M_i = k-1$ ) hold.

The following gives a simple pseudo code for the algorithm which finds the optimal solution of (3.38) as it is proposed in [30, 31, 40]

---

<sup>5</sup>Originally,  $E_i$  was picked as the BER of the subcarrier when coding was along each subcarrier rather than across different subcarriers.

**set**  $\underline{M} = [0, \dots, 0]$   
**while**  $\sum_{i=1}^L M_i < \frac{B_t}{R(\mathcal{C})}$   
**find**  $k$  **s.t.**  $\Delta(\mathcal{P}_k(M_k + 1)) = \min_i \Delta(\mathcal{P}_i(M_i + 1))$   
 $M_k = M_k + 1$

**set**  $p_i = \mathcal{P}_i(M_i)$  for  $1 \leq i \leq L$

Note that the condition  $\Delta(\mathcal{P}_k(M_k + 1)) = \min_i \Delta(\mathcal{P}_i(M_i + 1))$  results in the extra bit being sent over the subcarrier with the least amount of extra power.

In each of the following subsections, we will show that all the proposed optimization problems in Section 3.2.3 can be solved by using MMP. Unlike Section 3.2.3, random interleaving is assumed here. Therefore, we do not optimize over  $\underline{\pi}$  in the following subsections.

### BER Minimization

Using the asymptotic BER expression given in (3.32) and noting that in the high SNR regime, the argument of the  $Q$ -function dominates over the multiplicative coefficient, we obtain that the BER minimization problem defined in (3.8) is equivalent to the following optimization problem (recall that  $a_{i,1} = (d_{\min}^{\mathcal{X}_i})^2 p_i h_i^2$ ):

$$\begin{aligned} \max_{\mathcal{C}, \underline{\mathcal{X}}, \underline{p}} \quad & d_{\text{free}}(\mathcal{C}) (d_{\min}^{\mathcal{X}_{i^*}})^2 p_{i^*} h_{i^*}^2 \\ \text{s.t.} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i \geq B_t, \\ & \sum_{i=1}^L p_i \leq P_t. \end{aligned} \tag{3.41}$$

The following lemma gives necessary conditions for the optimal solution.

**Lemma 3.4.1** *The optimal solution of (3.41)  $(\mathcal{C}^*, \underline{\mathcal{X}}^*, \underline{p}^*)$  satisfies*

$$\begin{aligned}
 (d_{\min}^{\mathcal{X}_i})^2 p_i h_i^2 &= (d_{\min}^{\mathcal{X}_{\hat{i}}})^2 p_i h_i^2, \quad i, \hat{i} \in \{1, \dots, L\}, \\
 R(\mathcal{C}) \sum_{i=1}^L M_i^* &= B_t, \\
 \sum_{i=1}^L p_i^* &= P_t.
 \end{aligned} \tag{3.42}$$

**Proof** The goal here is to maximize the minimum of  $(d_{\min}^{\mathcal{X}_i})^2 p_i h_i^2$  for  $i \in \{1, \dots, L\}$  subject to a target rate  $B_t$  and target power  $P_t$ . Assume the values of  $(d_{\min}^{\mathcal{X}_i})^2 p_i h_i^2$  for  $i \in \{1, \dots, L\}$  are not equal for the optimal solution. Therefore, power can be given to  $\min \left( (d_{\min}^{\mathcal{X}_i})^2 p_i h_i^2 \right)$  from other subcarriers which causes  $\min \left( (d_{\min}^{\mathcal{X}_i})^2 p_i h_i^2 \right)$  to increase and that's a contradiction to the optimal solution.

Due to Lemma 3.4.1, problem (3.41) can be decomposed into two consecutive optimization problems. First, we fix the coding scheme, i.e.,  $\mathcal{C}$ , and then we optimize the power and bit loading policy for the given code. The optimization with respect to the code  $\mathcal{C}$  can be done by full enumeration since usually there are only a few different coding schemes (e.g. puncturing patterns) available. Hence, (3.41) can be written as

$$\begin{aligned}
 \max_{\mathcal{C}} \quad & Ed_{\text{free}}(\mathcal{C}) \\
 \max_{\underline{\mathcal{X}}, \underline{p}} \quad & E \\
 \text{s.t. :} \quad & (d_{\min}^{\mathcal{X}_i})^2 p_i h_i^2 = E, \quad i = 1, \dots, L, \\
 & R(\mathcal{C}) \sum_{i=1}^L M_i = B_t, \\
 & \sum_{i=1}^L p_i = P_t.
 \end{aligned} \tag{3.43}$$

Letting

$$E_i = (d_{\min}^{x_i})^2 p_i h_i^2, \quad (3.44)$$

we can rewrite the inner optimization problem from (3.43) as

$$\begin{aligned} \max_{\underline{x}, \underline{p}} \quad & E \\ \text{s.t.} \quad & \sum_{i=1}^L p_i = P_t \\ & \sum_{i=1}^L M_i = B_t \\ & E_i = E, \quad \forall i \in \{1, \dots, L\}. \end{aligned} \quad (3.45)$$

The following theorem shows that we can use the greedy algorithms from [30, 31, 40] to solve the BER minimization problem.

**Theorem 3.4.2** *If  $\underline{x}^\dagger$  and  $\underline{p}^\dagger$  are the solution of (3.38), then solution  $\underline{x}^*$  and  $\underline{p}^*$  of (3.45) is given by*

$$\underline{x}^* = \underline{x}^\dagger, \quad (3.46)$$

$$p_i^* = \frac{P_t}{\sum_{i=1}^L p_i^\dagger} p_i^\dagger, \quad i = 1, \dots, L. \quad (3.47)$$

**Proof** Assume that the solution of (3.45) is  $\hat{\underline{x}}$  and  $\hat{\underline{p}}$  which is different from  $\underline{x}^*$  and  $\underline{p}^*$  given in (3.46) and (3.47). Accordingly, the corresponding values for  $E$  satisfy  $\hat{E} > E^*$ . Since,  $E_i$  is a linear function of  $p_i$  and

$$\sum_{i=1}^L \hat{M}_i = \sum_{i=1}^L M_i^* = B_t,$$

the vector  $\hat{\underline{p}}$  can be scaled down by the factor<sup>6</sup>  $E_t/\hat{E}$  which together with  $\hat{\underline{\mathcal{X}}}$  results in a better optimal solution for (3.38) than  $\underline{\mathcal{X}}^\dagger$  and  $\underline{p}^\dagger$  which is a contradiction.

### Power Minimization

By using relation (3.32), the power minimization problem defined in (3.9) can be reformulated as

$$\begin{aligned} \min_{\mathcal{C}, \underline{\mathcal{X}}, \underline{p}} \quad & \sum_{i=1}^L p_i \\ \text{s.t.} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i \geq B_t, \\ & P_b(a_{i^*,1}, \mathcal{C}, \underline{\mathcal{X}}) \leq BER_t, \end{aligned} \quad (3.48)$$

where  $P_b(a_{i^*,1}, \mathcal{C}, \underline{\mathcal{X}})$  is the BER expression defined in (3.32). From this equation, we have

$$P_b(a_{i^*,1} + \delta, \mathcal{C}, \underline{\mathcal{X}}_1) < P_b(a_{i^*,1}, \mathcal{C}, \underline{\mathcal{X}}_2), \quad \forall \underline{\mathcal{X}}_1, \underline{\mathcal{X}}_2 \in \mathbb{Z}^L. \quad (3.49)$$

Note that (3.49) is independent of bit loading  $\underline{\mathcal{X}}$ .

**Lemma 3.4.3** *The optimal solution of (3.48) satisfies*

$$a_{i,1} = a_{\hat{i},1}, \quad i, \hat{i} \in \{1, \dots, L\}. \quad (3.50)$$

**Proof** Assume that the optimal solution of (3.48) does not satisfy the condition in (3.50), then the power allocated to all subcarriers except the one with minimum  $a_{i,1}$  can be reduced to the point that they do not alter the minimum value of  $a_{i,1}$ , which is a contradiction.

---

<sup>6</sup>Note that  $\frac{P_t}{\sum_{i=1}^L p_i^\dagger} = \frac{E_t^*}{E_t}$

According to Lemma 3.4.3 and using the same argument as in Section 3.4.2, the problem of (3.48) can be decomposed into two problems as

$$\begin{aligned}
 & \min_{\mathcal{C}} P_t \\
 & \min_{\underline{\mathcal{X}}, \underline{p}} P_t \\
 & \text{s.t.} \quad \sum_{i=1}^L p_i = P_t, \\
 & \quad R(\mathcal{C}) \sum_{i=1}^L M_i = B_t, \\
 & \quad a_{i,1} = E, \quad i = 1, \dots, L, \\
 & \quad P_b(E, \mathcal{C}, \underline{\mathcal{X}}) \leq BER_t.
 \end{aligned} \tag{3.51}$$

Lemma 3.4.3 also allows us to use the approximation in (3.37) instead of the one in (3.32) for BER calculations.

The inner problem in (3.51) is an MMP but we cannot solve it directly since the value of  $E$  is unknown ( $E$  should be calculated from (3.37) but this expression depends on the bit loading  $\underline{\mathcal{X}}$  which is unknown.). The following theorem explains our proposed solution.

**Theorem 3.4.4** *If  $\underline{\mathcal{X}}^\dagger$  and  $\underline{p}^\dagger$  are the solution of (3.38), then solution  $\underline{\mathcal{X}}^*$  and  $\underline{p}^*$  of the inner problem of (3.51) is given by*

$$\underline{\mathcal{X}}^* = \underline{\mathcal{X}}^\dagger, \tag{3.52}$$

$$p_i^* = \frac{E^*}{E_t} p_i^\dagger, \quad i = 1, \dots, L, \tag{3.53}$$

where  $E^*$  is calculated by (3.37) and using  $\underline{\mathcal{X}}^\dagger$  for bit loading.

**Proof** Assume that the optimal solutions are given by  $\hat{\underline{\mathcal{X}}}$ ,  $\hat{\underline{p}}$  and  $\hat{E}$  rather than  $\underline{\mathcal{X}}^*$ ,  $\underline{p}^*$ , and  $E^*$ , where  $\sum_{i=1}^L \hat{p}_i < \sum_{i=1}^L p_i^*$ . The resulting BER according to  $\hat{\underline{\mathcal{X}}}$ ,  $\hat{\underline{p}}$  and  $\hat{E}$  should be less than or equal to  $BER_t$ , therefore according to (3.49) we have  $\hat{E} \geq E^*$ . Also, we

can find the solution for the original MMP by multiplying  $\hat{p}$  by  $\frac{E_t}{E}$ . We have  $\frac{E_t}{E} \sum_{i=1}^L \hat{p}_i \leq \frac{E_t}{E^*} \sum_{i=1}^L \hat{p}_i \leq \frac{E_t}{E^*} \sum_{i=1}^L p_i^* = P^\dagger$  which means we found a solution for the MMP problem which is better than (or equivalent to) the optimal solution. The case of better than optimal results is a contradiction and the case of equivalent to optimal results is trivial.

### Throughput Maximization

The throughput maximization problem defined in (3.10) can be reformulated as

$$\begin{aligned} \max_{\mathcal{C}, \underline{\mathcal{X}}, \underline{p}} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i \\ \text{s.t.} \quad & \sum_{i=1}^L p_i \leq P_t, \\ & P_b(a_{i^*,1}, \mathcal{C}, \underline{\mathcal{X}}) \leq BER_t, \end{aligned} \tag{3.54}$$

where  $P_b(a_{i^*,1}, \mathcal{C}, \underline{\mathcal{X}})$  is defined in (3.32).

**Lemma 3.4.5** *The optimal solution of (3.54) either satisfies*

$$a_{i,1} = a_{\hat{i},1}, \quad i, \hat{i} \in \{1, \dots, L\}, \tag{3.55}$$

*or can be manipulated such that it satisfies this condition*

**Proof** Assume that the optimal solution of (3.54) does not satisfy the condition in (3.55). Then, the power allocated to all subcarriers except the one with minimum  $a_{i,1}$  can be reduced to the point that they do not alter the minimum value of  $a_{i,1}$  ( $a_{i^*,1}$ ). This power may be allocated to one subcarrier in order to carry one more bit and if the extra power is not sufficient for such an action, it always can be distributed to subcarriers such that they preserve the condition in (3.55).

Therefore, this problem can be reformulated as

$$\begin{aligned}
 & \max_{\mathcal{C}} B_t \\
 & \max_{\underline{\mathcal{X}}, \underline{p}} B_t \\
 \text{s.t.} \quad & R(\mathcal{C}) \sum_{i=1}^L M_i = B_t \\
 & \sum_{i=1}^L p_i = P_t, \\
 & a_{i,1} = E, \quad i = 1, \dots, L, \\
 & P_b(E, \mathcal{C}, \underline{\mathcal{X}}) \leq BER_t.
 \end{aligned} \tag{3.56}$$

This means we can solve the inner problem of (3.56) for different values of  $\mathcal{C}$  and choose the one with largest  $B_t$ .

The following theorem gives the relation between the solution of MMP and rate maximization problem.

**Theorem 3.4.6** *Assume that the solutions of (3.38) for total rate of  $B$  to be  $\tilde{\underline{\mathcal{X}}}_B, \tilde{\underline{p}}_B$  and for total rate of  $B + 1$  to be  $\tilde{\underline{\mathcal{X}}}_{B+1}, \tilde{\underline{p}}_{B+1}$ . We also define*

$$p_{B+1}^*(i) = \frac{P_t}{\sum_{i=1}^L \tilde{p}_B(i)} \tilde{p}_B(i), \quad i \in \{1, \dots, L\}, \tag{3.57}$$

and

$$p_{B+1}^*(i) = \frac{P_t}{\sum_{i=1}^L \tilde{p}_{B+1}(i)} \tilde{p}_{B+1}(i), \quad i \in \{1, \dots, L\}, \tag{3.58}$$

and calculate  $E_{t1}$  (and  $E_{t2}$ ) by inserting  $\underline{p}_B^*$  and  $\tilde{\underline{\mathcal{X}}}_B$  ( $\underline{p}_{B+1}^*$  and  $\tilde{\underline{\mathcal{X}}}_{B+1}$ ) in (3.44). Then  $\underline{\mathcal{X}}^* = \tilde{\underline{\mathcal{X}}}_B$  and  $\underline{p}_B^*$  are the optimal solution of the inner problem of (3.56) if  $P_b(E_{t1}, \mathcal{C}, \tilde{\underline{\mathcal{X}}}_B) \leq BER_t$  and  $P_b(E_{t2}, \mathcal{C}, \tilde{\underline{\mathcal{X}}}_{B+1}) > BER_t$ .

**Proof** Assume that the optimal solution of the inner problem (3.56) when  $a_{i,1} = \hat{E}$  to be  $\hat{\underline{\mathcal{X}}}$  and  $\hat{\underline{p}}$  where  $R(\mathcal{C}) \sum_{i=1}^L \hat{M}_i \geq B + 1$ . According to (3.49),  $\hat{E} \geq E_{t2}$  which is equivalent to  $\frac{E_t}{\hat{E}} \leq \frac{E_t}{E_{t2}}$ . Therefore, if we scale  $\hat{\underline{p}}$  by  $\frac{E_t}{\hat{E}}$  we get a better or equivalent solution for (3.38) than when we solve it for  $B + 1$  bits since  $\frac{E_t}{\hat{E}} \sum_{i=1}^L \hat{p}(i) \leq \frac{E_t}{E_{t2}} \sum_{i=1}^L p_{B+1}^*(i)$ <sup>7</sup>. The former is obviously a contradiction and the latter results in contradiction due to the uniqueness of the optimal solution of (3.38).

According to Theorem 3.4.6, we solve the inner problem of (3.56) iteratively. We start by  $B = 1$  in (3.38) and calculate  $P_b$  by using (3.32). If the resulting  $P_b$  is less than or equivalent to  $BER_t$ , we increase  $B$  by 1 and do the whole process again. This process is continued until  $P_b$  condition is violated.

Since for relatively high values of  $P_b$ , i.e. values above  $10^{-6}$ , the asymptotic expression of (3.32) is not tight, we propose to use a backtrack algorithm using the BER union bound (3.29). That is, after that the algorithm stopped, we compute the BER union bound at each step and remove the bits added in a step if the BER is larger than  $BER_t$ . We continue this process until the BER condition is satisfied.

## 3.5 Numerical Results

In this section we present selected simulative results confirming the usefulness of the proposed algorithms and the accuracy of the performance approximation. We use WLAN IEEE 802.11a/g OFDM system with  $L = 48$  active subcarriers [8]. The quasi-static exponentially-decaying multipath Rayleigh fading channel model which is explained in Section 2.2.3 with  $T_s = 50$  ns and  $T_{\text{rms}} = 150$  ns is used. The ensemble of available encoders are obtained using the quasi-standard rate-1/2 memory-6 convolutional code (generator polynomials are  $(171, 133)_8$ ) as the mother code and applying different puncturing patterns. In

---

<sup>7</sup>Note that  $\tilde{\underline{\mathcal{X}}}_{B+1}$  and  $\underline{p}_{B+1}^*$  are the solution of (3.38) for  $E_t = E_{t2}$  and  $\sum_{i=1}^L \hat{p}(i) = \sum_{i=1}^L p_{B+1}^*(i) = P_t$ .

particular, the rates 1/2, 2/3, 3/4, and 1 (i.e. uncoded transmission) are used. The set of available modulations are 4-QAM, 16-QAM, and 64-QAM. In the following, first we show the accuracy of the proposed union bound and asymptotic BER expressions. Then, the results for BER minimization, power minimization and throughput maximization algorithms are shown.

### 3.5.1 Variation of BER for the Optimized System for Different Interleavers

Figure 3.2 shows the variation of BER versus  $\gamma_b$  for the system optimized according to (3.43) when one information bit per subcarrier is transmitted, and  $\gamma_b = P_t \sum_{i=1}^L \mathbb{E}\{h_i^2\}/B$  (recall that  $\mathbb{E}\{\|z_i\|^2\} = 1$ ) is a bit-wise SNR measure. A single channel realization and  $N = 64$ ,  $N = 256$  and  $N = 1024$  are considered, to highlight the influence of the number of subcarriers. Each sub-figure depicts (i) the BER union bound which is computed using (3.24) and (3.29) (ii) the asymptotic BER which is achieved using (3.24), (3.37), and considering only the first PEP and (iii) the BER simulation results for 1000 randomly generated interleavers. Instead of showing all BER results, at each SNR we have shown the lowest and largest BERs. We observe that for (properly) loaded BIC-OFDM the effect of interleaving is marginal. This corroborates the simplifications made in the derivation of the loading algorithm, which required averaging over interleaving. It can also be seen that the variation of BER, i.e. the difference between the lowest and largest BER at a particular SNR, is converging to zero as the number of subcarriers grows larger.

### 3.5.2 BER Minimization Results

Next, BER minimization results are shown. Figure 3.3 shows the average BER over 1000 randomly generated channels, while Figure 3.4 shows the 10% outage BER for the same

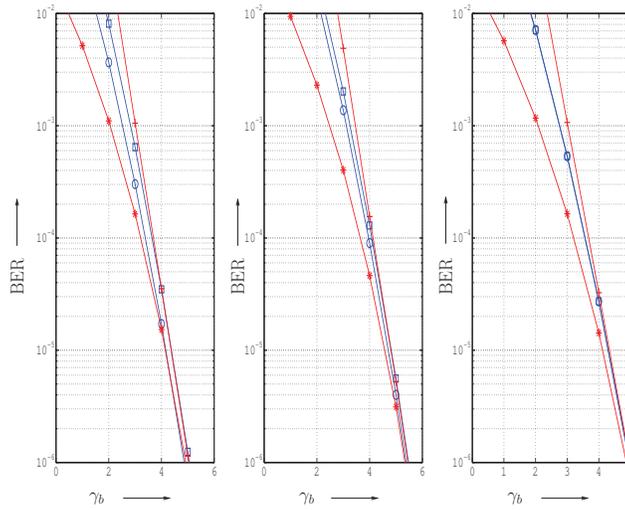


Figure 3.2: BER results for different number of subcarriers: (a)  $N=64$ , (b)  $N=256$  and (c)  $N=1024$ . Square marker: highest BER for each SNR for 1000 random interleavers. Circle marker: lowest BER for each SNR for 1000 random interleavers. Solid line: union bound derived from (3.24) and (3.29). Dashed line: asymptotic BER derive from (3.24) and (3.37).

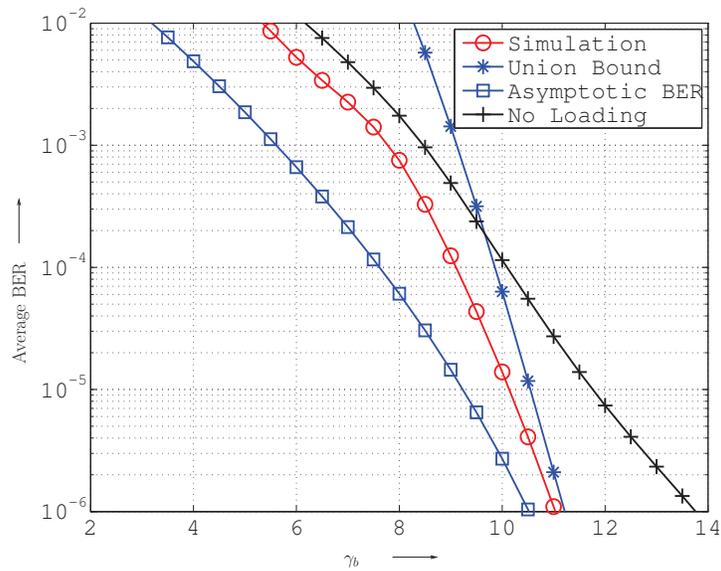


Figure 3.3: Average BER versus  $\gamma_b$  for BER minimization algorithm.

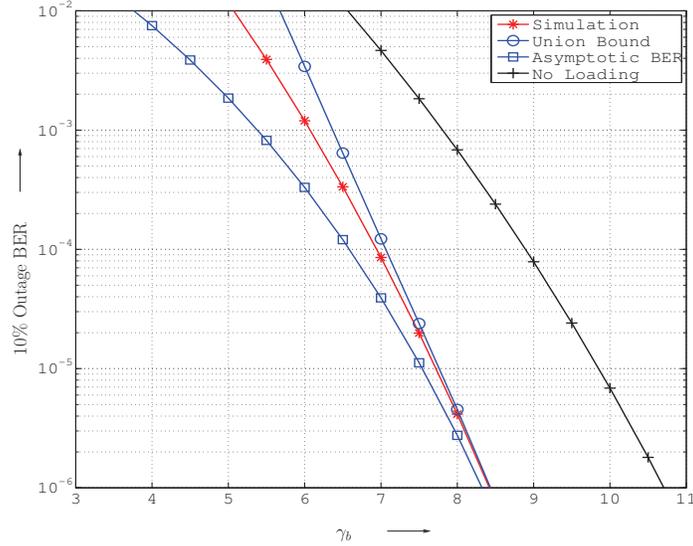


Figure 3.4: 10% outage BER versus  $\gamma_b$  for BER minimization algorithm.

1000 channels. Each figure contains (i) the BER performance of the non-adaptive system using 4-QAM modulation with code rate 1/2, (ii) the BER performance of the optimized system, (iii) the BER union bound, and (iv) the asymptotic BER approximation. We observe a substantial gain of the adaptive system over the non-adaptive system. Furthermore, it can be seen that the BER union bound approximation predicts the actual performance very accurately for target error rates of about  $10^{-4}$  and below. Finally, the asymptotic BER approximation is shown to be quite useful for a quick performance estimation. We note that these findings are true for both average (Figure 3.3) and outage (Figure 3.4) BER results.

### 3.5.3 Power Minimization Results

We now consider the power minimization algorithm. In particular, we consider the power minimization given the constraint that the BER of the system is required to be less than  $10^{-5}$  (i.e.  $BER_t = 10^{-5}$ ). Again we consider the case where 1 information bit is to be

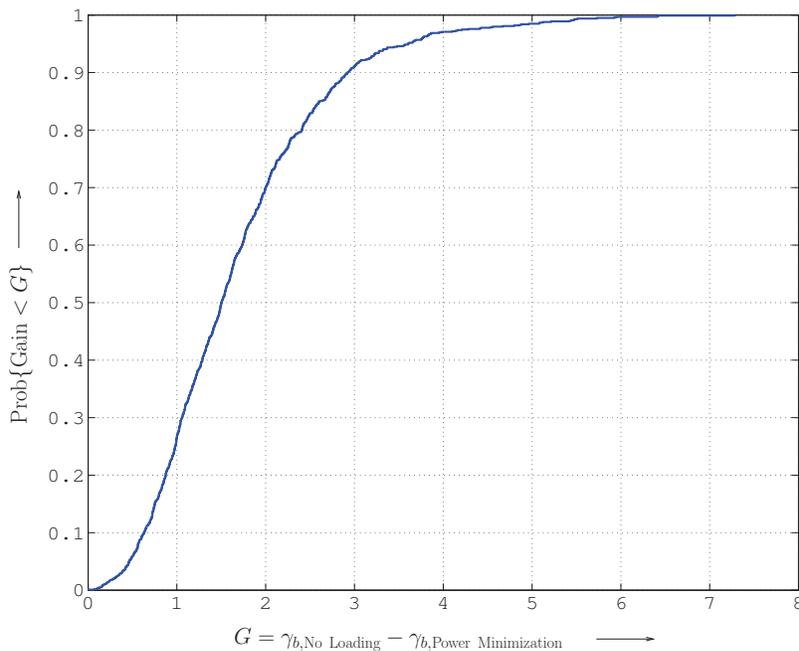


Figure 3.5: Cumulative distribution function (CDF) of power gain  $G = \gamma_{b, \text{No Loading}} - \gamma_{b, \text{Power Minimization}}$  of power minimization algorithm compared with no loading system for  $BER_t = 10^{-5}$ .

transmitted over each subcarrier. Figure 3.5 shows the empirical cumulative distribution function (CDF) of the power gain of the proposed algorithm compared to the non-adaptive system using 4-QAM modulation  $G = \gamma_{b, \text{No Loading}} - \gamma_{b, \text{Power Minimization}}$  for 1000 randomly generated channels. We can see a substantial gain can be achieved by using our proposed method; for instance we can expect gains of more than 1.5 dB for 50% of channels. Also, Figure 3.6 shows the simulated BER performance of the system along with the asymptotic BER approximation for each channel realization. In fact, Figure 3.6 shows that the achieved gains do not compromise the required performance of the system. Again, we can see that the proposed asymptotic approximation is an accurate estimate for the true performance of the system.

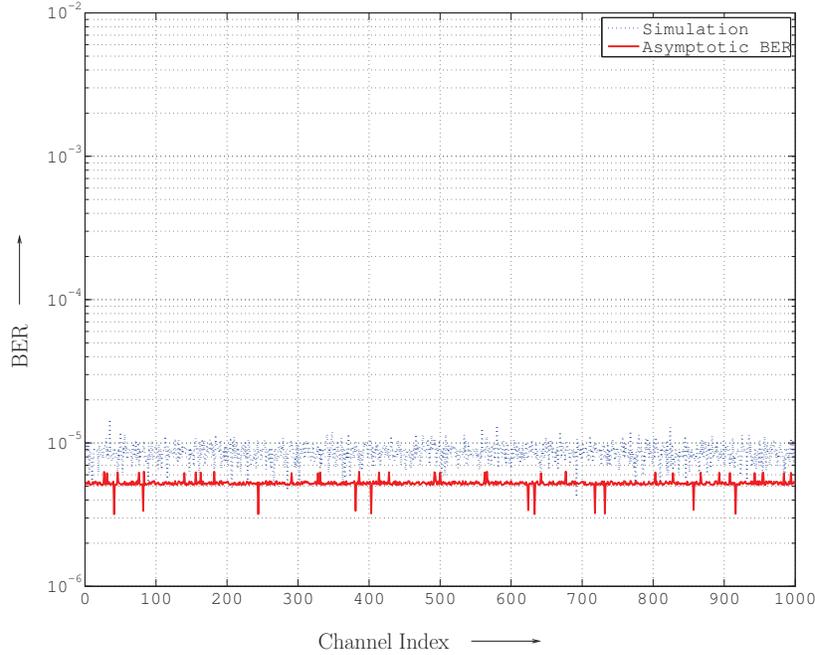


Figure 3.6: BER of the power minimization algorithm along with the asymptotic BER.

### 3.5.4 Throughput Maximization

In this section, we consider the problem of rate maximization for BIC-OFDM. We compare our result with the method presented in [20] and we will call their algorithm AMC-BIC-OFDM. AMC-BIC-OFDM relies on simulative optimization of three parameters, i.e.  $\Gamma$  and two auxiliary power assignments  $p_{16}$  and  $p_{64}$  for 16-QAM and 64-QAM constellations.  $\Gamma$  is a parameter from which the error-rate can be computed disregarding the type of coding and modulation. This parameter is assumed to be constant for the given target error-rate. Furthermore, since the error rate union bound derived in [20] is not accurate when higher order modulations like 16-QAM are used in the OFDM symbol, the authors have proposed to use two heuristic parameters i.e. auxiliary power assignments to improve the performance of the system.

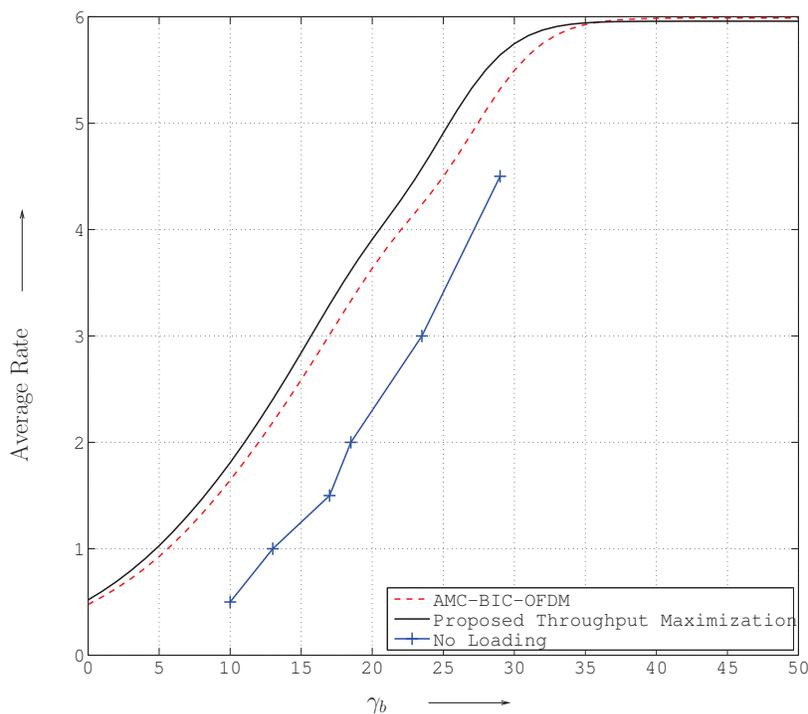


Figure 3.7: Comparing the average data rate for the proposed throughput maximization algorithm, AMC-BIC-OFDM and no loading system.

According to our correspondence to the authors of [20] and considering the simulation results section of the same paper, they did not derive these parameters analytically. In fact, they claimed that due to the complicated nature of error-rate prediction, it is practically impossible to come up with exact  $\Gamma$  for a coded system. Furthermore, they mentioned that  $\Gamma$  is a heuristic parameter and it may not be corresponding to target error-rate. Therefore, the process of determining the value of these parameters for a given target error rate is not straightforward and requires time consuming simulative optimization [48].

Figure 3.7 compares the average rate per subcarrier of our proposed method, AMC-BIC-OFDM and no loading system (IEEE 802.11a/g standard) for target BER of  $10^{-5}$ . For AMC-BIC-OFDM, we found  $\Gamma = 9.7$  dB,  $p_{16} = 0.5$  dB and  $p_{64} = 0.9$  dB by exhaustive

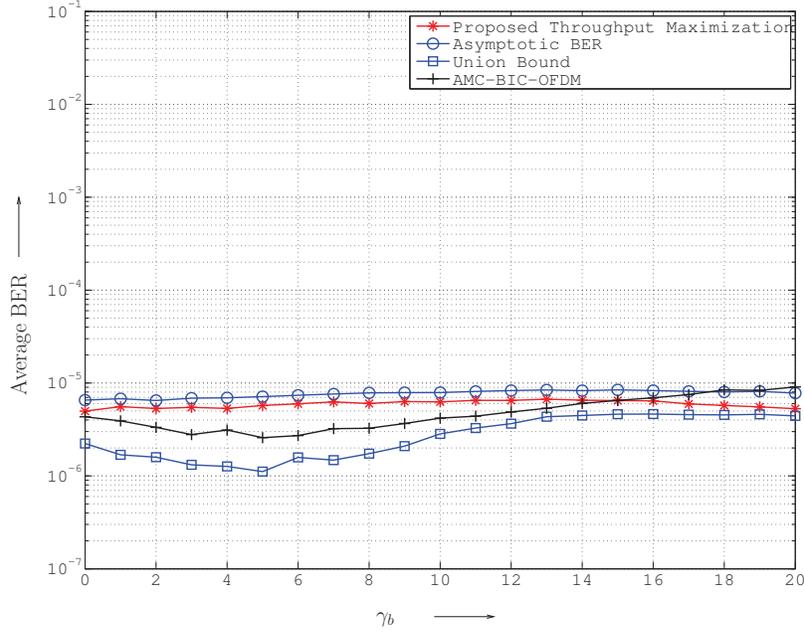


Figure 3.8: Average BER for the proposed throughput maximization algorithm along with the corresponding asymptotic and union bound BERs and also the BER of AMC-BIC-OFDM.

search which results in a BER less than the target BER. It can be seen that our method outperforms AMC-BIC-OFDM by almost 2 dB while its complexity is the same. Also, it does not rely on complex simulative optimization method. Furthermore, Figure 3.8 shows the average BER of our proposed algorithm in comparison with AMC-BIC-OFDM. Besides, the BER union bound and asymptotic BER are shown. We can see that the BER for AMC-BIC-OFDM is about the target BER of  $10^{-5}$  which shows our proposed values for  $\Gamma$ ,  $p_{16}$  and  $p_{64}$  are acceptable. Again, the accuracy of the proposed asymptotic BER expression can be observed. Finally, Figure 3.9 shows the average usage of each code as a function of  $\gamma_b$  for the proposed method and AMC-BIC-OFDM. We can see that our algorithm tends to select the code rate of  $2/3$  more often compared to AMC-BIC-OFDM. We also compared our algorithm for target WER of  $10^{-2}$  and AMC-BIC-OFDM with  $\Gamma$ ,

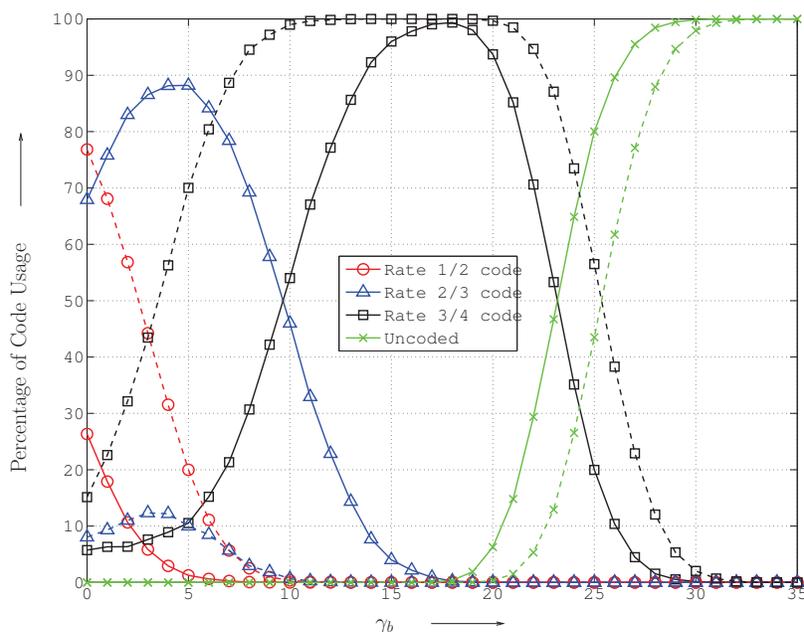


Figure 3.9: Percentage of code rate usage for the proposed rate maximization algorithm (line) and AMC-BIC-OFDM (dashed).

$p_{16}$  and  $p_{64}$  as in [20] (These parameters are set for target WER of  $10^{-2}$ ). The results were similar to those we presented in Figures 3.7, 3.8 and 3.9 but are not shown here.

### 3.6 Summary

In this chapter, we have considered adaptive BIC-OFDM. In particular, we have focused on BER minimization, power minimization and throughput maximization using bit and power allocation and code rate selection. We have derived simplified BER expressions considering the effect of interleaver and also assuming random interleaver using the PDF of reliability matrices. BER expression for random interleaving enabled us to formulate the optimization problems. Based on the derived expression, we introduced algorithms to solve those optimization problems using known loading algorithm for uncoded OFDM

transmission known as margin maximization problem. We have also proposed closed-form expression for the performance of the optimized system. Selected numerical results have confirmed significant performance improvement using the proposed methods and a good accuracy of the performance approximation.

# Chapter 4

## Using Compressed Sensing to Mitigate Nonlinear Distortion in OFDM Signals

### 4.1 Introduction

Although OFDM systems have many advantages over single carrier systems, they also have their own drawbacks. One of the major problems of OFDM is the relatively high peak-to-average power ratio (PAPR) of the time domain transmit signal, which is the result of adding many signals with different frequencies. This causes nonlinear signal distortion when the signal passes through the high-power amplifier (PA). To address this problem, numerous techniques have appeared in the literature based on signal and/or data modification at the transmitter [12] or mitigation of nonlinear distortion at the receiver [49, 50, 51, 52, 53]. When there are computational complexity, power consumption, and implementation cost limitations for the OFDM transmitter, receiver-based techniques to reduce nonlinear distortion are highly desirable.

Assuming that the PA clipping level is known at the receiver, receiver-based techniques [49, 50, 51, 52, 53] basically reconstruct the clipping distortion caused by the PA. These techniques can be mainly divided into two classes: 1) the first class of techniques assumes

that nonlinear distortion does not have an impact on channel estimation [49, 50, 51, 52], and 2) in the second class of techniques the effect of nonlinear distortion on pilot symbols that degrades channel estimation performance is taken into account [53]. Within the first class, the iterative methods proposed in [49] and [50] reconstruct the clipped signal to its original form and estimate the clipping noise and cancel it from the received signal, respectively. The technique proposed in [52] reconstructs the clipped samples by using other samples in an oversampled system. The problem for this method is that it requires a bandwidth expansion between 25% to 100%. Belonging to the second class, the technique presented in [53] employs adaptive clipping and noise filtering to avoid distorting pilots and improve the performance of channel estimation. This technique increases PAPR and complexity at the transmitter due to performing adaptive clipping and noise filtering.

All the above-mentioned techniques can effectively mitigate the effect of nonlinear distortion on the transmitted signal if the clipping level of the nonlinear PAs is set relatively large [49, 50, 51, 52, 53], i.e., the probability of clipping is relatively small. In this case, the effect of clipping can be modeled as additive impulsive noise that is *sparse* in time domain [54]. This noise can significantly degrade the performance of OFDM transmission as even a single impulse in an OFDM block can notably corrupt all symbols of the block due to its large energy.

According to recent results in sparse signal processing, a.k.a. compressed sensing (CS) (cf. e.g. [55]), a sparse signal can be recovered using a small number of linear projections over a random basis that is incoherent with respect to the basis in which the signal is sparse. CS can be employed to estimate impulsive noise at the receiver due to the inherent time-domain sparsity of impulsive noise. This forms the main idea of the CS-based method for impulsive noise cancellation proposed in [56]. In this method, non-modulated (null, pilot, etc.) subcarriers are used to estimate impulsive noise. In [56], it is assumed that impulsive

noise, e.g. large nonstationary electromagnetic disturbances as they are common in digital subscriber line and power line communication systems, occurs *at the receiver*. However, the PA nonlinearity distorts the OFDM signal *at the transmitter* and the CS-based approach for impulsive noise estimation and cancellation uses the received signal. Moreover, the effect of impulsive noise on the estimated channel transfer function, that is needed for the CS-based impulsive noise estimation and cancellation, has not been addressed in [56].

In this chapter, we propose a CS-based approach for detection of nonlinearly distorted OFDM signals at the receiver. The proposed technique exploits pilot tones inserted in the OFDM signal. We consider two different scenarios for the receiver-side detection, namely perfect and imperfect channel estimation. For the case of imperfect channel estimation, our proposed scheme first obtains an initial estimate of the channel in frequency domain which is used to estimate nonlinear distortion by performing CS in time domain. Then, the estimated nonlinear distortion is used to improve the channel estimation and it is also removed from the received signal. We also modify the iterative techniques from [49] and [50] to use the estimated distortion to refine channel estimation, which allows us to do a direct performance comparisons. Simulation results demonstrate that the proposed CS-based technique can accurately estimate clipping noise caused by PA nonlinear distortion. Furthermore, the proposed scheme can notably improve the BER performance over the benchmark techniques from [49] and [50] for the case of imperfect channel estimation.

The remainder of this chapter is organized as follows. In Section 4.2, the OFDM system model including nonlinear power amplifier is introduced and some basics of CS theory are provided. The proposed CS-based technique is described in Section 4.3, which also provides the extension of the iterative techniques [49] and [50] to the case of imperfect channel estimation. Numerical results are presented and discussed in Section 4.4, and finally concluding remarks follow in Section 4.5.

## 4.2 OFDM Signaling and Compressed Sensing

In this section, we first present the formulation for OFDM transmission. In doing so, we also describe the nonlinear PA model used in this chapter. In the second part, we review essential facts from CS theory.

### 4.2.1 OFDM System Model

#### OFDM transmitter

We assume the frequency-domain OFDM symbol  $\mathbf{X} = [X_0, \dots, X_{L-1}]^T$  is transmitted on  $L$  orthogonal subcarriers. The symbol  $X_k$  sent over the  $k$ th subcarrier is selected from a phase-shift keying (PSK) or quadrature amplitude modulation (QAM) constellation. The complex envelope of the baseband OFDM signal, defined over the time interval  $t \in [0, T_s]$ , where  $T_s$  is the OFDM symbol duration, can be expressed as

$$x(t) = \frac{1}{L} \sum_{k=0}^{L-1} X_k e^{j2\pi kt/T_s} . \quad (4.1)$$

In practice,  $JL$  discrete-time samples of  $x(t)$  are efficiently computed by an inverse discrete fourier transform (IDFT) [51]

$$x_n = \frac{1}{L} \sum_{k=0}^{L-1} X_k e^{j2\pi kn/JL}, \quad n = 0, \dots, JL - 1 \quad (4.2)$$

where  $J$  is the oversampling factor. A cyclic prefix of length  $L_g$  is appended to  $x_n$  in order to prevent intersymbol interference among blocks and enable simple single-tap equalization. Using digital-to-analog conversion,  $x(t)$  is obtained from the cyclically prefixed  $x_n$ . Defining

the time-domain vector  $\mathbf{x} = [x_0, \dots, x_{JL-1}]^T$ , (4.2) can be represented in matrix form as

$$\mathbf{x} = \mathbf{Q}\mathbf{X} , \quad (4.3)$$

where  $\mathbf{Q}$  denotes the  $JL \times L$  IDFT matrix whose elements are  $Q_{n,k} = \frac{1}{L}e^{j\frac{2\pi(k-1)(n-1)}{JL}}$ ,  $1 \leq k \leq L, 1 \leq n \leq JL$ .

### PA nonlinearity

In order to describe the nonlinear character of the PA, we consider the oversampled discrete-time signal. This is correct if the dynamic range of the time-domain signal is limited in the digital domain, or a good approximation if the nonlinearity acts on the continuous-time signal  $x(t)$ [51]. From an extension of Bussgang's theory [57] it can be shown that the nonlinearly distorted signal can be expressed as

$$x'_n = \beta x_n + d_n , \quad (4.4)$$

where  $\beta$  is chosen such that signal sequence  $x_n$  and distortion sequence  $d_n$  are uncorrelated. Since the distortion is introduced to reduce the dynamic range of the transmit signal, the distortion vector  $\mathbf{d} = [d_1, \dots, d_{JL}]^T$  is essentially a vector of impulses, resulting from clipping events. In other words,  $\mathbf{d}$  is sparse. This point of view will be exploited for data detection in Section 4.3.

When presenting numerical results in Section 3.5 we consider two power amplifier models:

- Soft limiter (SL) power amplifier:

The SL power amplifier model is defined according to

$$x'_n = \begin{cases} x_n, & |x_n| \leq A \\ A \frac{x_n}{|x_n|}, & |x_n| > A \end{cases}, \quad (4.5)$$

where  $A$  is a predefined threshold. To measure the amount of distortion due to the SL nonlinearity, the clipping level  $\alpha$  is defined as

$$\alpha = \frac{A}{\sqrt{\mathbb{E}(|x_n|^2)}}. \quad (4.6)$$

- Solid State Power Amplifier (SSPA):

According to [58],  $x'_n$  can be written as

$$x'_n = \frac{\rho x_n}{\left(1 + \left|\frac{\rho x_n}{B}\right|^{2p}\right)^{\frac{1}{2p}}}, \quad (4.7)$$

where  $B$  is the output saturation level,  $p$  controls the smoothness of the SSPA, and  $\rho$  is the amplifier small signal gain. The amplifier saturation power is defined as  $P_{sat} = B^2$ . In order to reduce the nonlinear distortion due to signal peaks, the amplifier is driven with an input back-off (IBO)

$$\text{IBO}_{\text{dB}} = 10 \log_{10} \left( \frac{P_{sat}}{\mathbb{E}(|x_n|^2)} \right). \quad (4.8)$$

It should be noted that for large values of  $p$  (typically  $p > 10$ ) and  $\rho = 1$ , the SSPA model from (4.7) approaches the SL model from (4.5).

For  $\alpha \gg 1$  or  $\text{IBO} \gg 1$  the factor  $\beta$  in (4.4) approaches one [51, 59]. For simplicity, we assume  $\beta = 1$  in the following.

### OFDM receiver

At the receiver and after removing the cyclic prefix, the oversampled received OFDM signal vector can be written as

$$\mathbf{y} = \mathbf{H}(\mathbf{x} + \mathbf{d}) + \mathbf{z} , \quad (4.9)$$

where the vector  $\mathbf{z} = [z_0, \dots, z_{JL-1}]^T$  denotes additive white Gaussian noise (AWGN) with variance  $\sigma_n^2$ ,  $\mathbf{H}$  is an  $(JL \times JL)$  circulant channel matrix, whose first column is  $[h_0, \dots, h_{L_m}, \mathbf{0}_{(JL-L_m-1) \times 1}]^T$ , and  $h_\ell$ ,  $0 \leq \ell \leq L_m$ , is the channel impulse response. It is assumed that  $L_g \geq L_m$  so that successive OFDM symbols do not interfere. The discrete frequency-domain received signal obtained after the DFT is given by

$$\mathbf{Y} = \mathbf{\Lambda} \mathbf{X} + \mathbf{Q}^H \mathbf{H} \mathbf{d} + \mathbf{Z} , \quad (4.10)$$

where

$$\mathbf{\Lambda} = \mathbf{Q}^H \mathbf{H} \mathbf{Q} = \text{diag}([\Lambda_0, \dots, \Lambda_{L-1}]) \quad (4.11)$$

is the  $(L \times L)$  diagonal matrix whose elements are the channel frequency response coefficients and

$$\mathbf{Z} = \mathbf{Q}^H \mathbf{z} \quad (4.12)$$

is the frequency-domain noise. The maximum likelihood (ML) receiver seeks an estimate  $\hat{\mathbf{X}}$  of  $\mathbf{X}$  according to

$$\hat{\mathbf{X}} = \underset{\tilde{\mathbf{X}}}{\text{argmin}} \|\hat{\mathbf{\Lambda}} \tilde{\mathbf{X}} + \mathbf{Q}^H \hat{\mathbf{H}} \hat{\mathbf{d}} - \mathbf{Y}\|_2 , \quad (4.13)$$

where  $\hat{\mathbf{\Lambda}}$ ,  $\hat{\mathbf{H}}$  and  $\hat{\mathbf{d}}$  are estimates of the channel representations  $\mathbf{\Lambda}$  and  $\mathbf{H}$  and the clipping noise  $\mathbf{d}$ . The estimation of these quantities from the received signal  $\mathbf{Y}$  based on CS

reconstruction is explained in Section 4.3.

## 4.2.2 Background on CS

In a typical compressed sensing problem, the goal is to recover the signal  $\mathbf{u} \in \mathbb{R}^{\bar{N}}$  which is a  $\xi$ -sparse vector in an original basis (such as time domain) based on observation vector  $\mathbf{v} \in \mathbb{R}^{\bar{M}}$  in an observation domain (such as frequency domain). By  $\xi$ -sparse we mean that at most  $\xi$  of its elements are nonzero.  $\mathbf{u}$  and  $\mathbf{v}$  are linearly related to each other based on the following formula

$$\mathbf{v} = \Phi \mathbf{u} + \mathbf{n} , \quad (4.14)$$

where  $\Phi \in \mathbb{R}^{\bar{M} \times \bar{N}}$  is called measurement matrix and  $\mathbf{n} \in \mathbb{R}^{\bar{M}}$  is the observation noise with bounded energy  $\|\mathbf{n}\|_2^2 \leq \epsilon^2$ .

The number of observations is far less than the dimension of the original signal, i.e.  $\bar{M} \ll \bar{N}$ . The authors in [60] showed that if there are enough observations and matrix  $\Phi$  obeys the *Uniform Uncertainty Principle (UUP)*, then vector  $\mathbf{u}$  can be reconstructed by solving an optimization problem as follows

$$\begin{aligned} \min \quad & \|\mathbf{u}\|_1 \\ \text{s.t.} \quad & \|\Phi \mathbf{u} - \mathbf{v}\|_2 \leq \epsilon \end{aligned} \quad (4.15)$$

The special case of Fourier measurements has been addressed in [56, 60, 61, 62]. That is,  $\Phi$  is a pruned Fourier matrix obtained by selecting  $\bar{M}$  rows from the Fourier matrix. While random selection of the rows as considered in e.g. [60, 61] may not be always practical in the considered scenario, there is recent work on deterministic selection of the rows of the Fourier matrix [56, 62].

*Definition:* Assume  $a$ ,  $b$  and  $\lambda$  be positive integers. An  $(a, b, \lambda)$  cyclic difference set

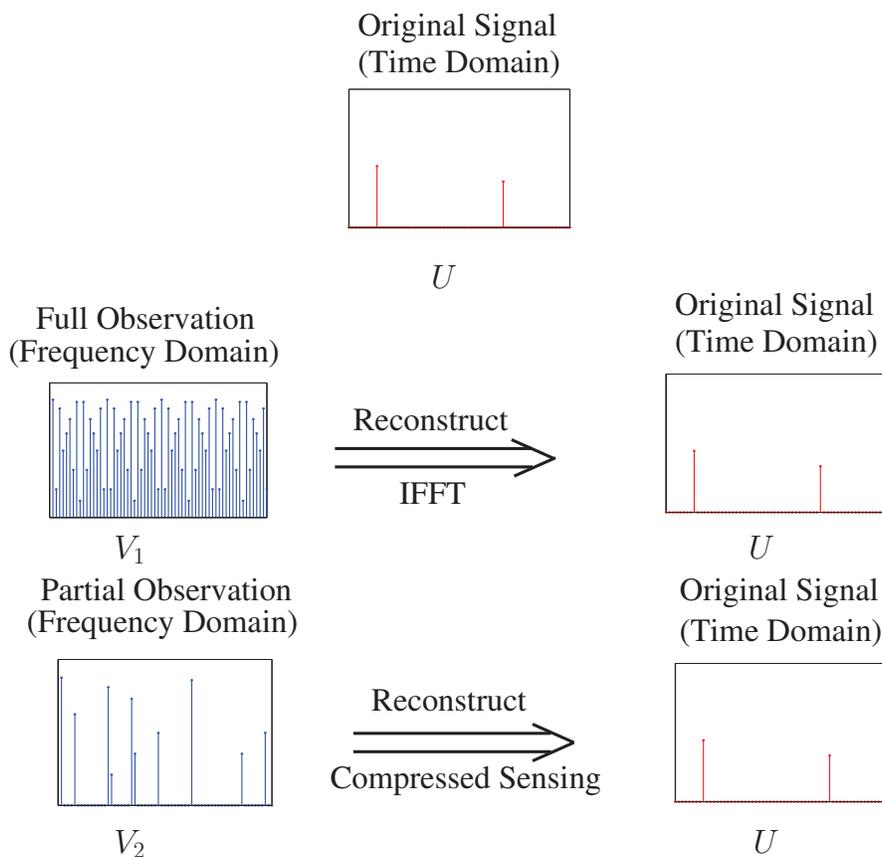


Figure 4.1: Compressed sensing example.

(CDS) is defined as a set of  $b$  distinct elements of  $\psi = \{w_1, w_2, \dots, w_b\}$  where a difference  $\eta \equiv w_i - w_j \pmod{a}$  with  $i \neq j$  takes values in  $1 \leq \eta \leq a - 1$  exactly  $\lambda$  times.

It has been shown in [56] and [62] that if the  $\bar{M}$  rows of the  $\bar{N} \times \bar{N}$  Fourier matrix are selected from an  $(\bar{N}, \bar{M}, \lambda)$  CDS, then the resulting partial Fourier matrix satisfies the UUP condition. Unfortunately, this result is not practical since there is no CDS for practical values of  $\bar{N}$  like  $2^6$ ,  $2^7$ , or  $2^8$ . For this reason, random selection has been used for numerical results of [56], and we also adopt this strategy for the numerical results presented in Section 4.4.

As an example, consider the case presented in Figure 4.1. The goal here is to reconstruct

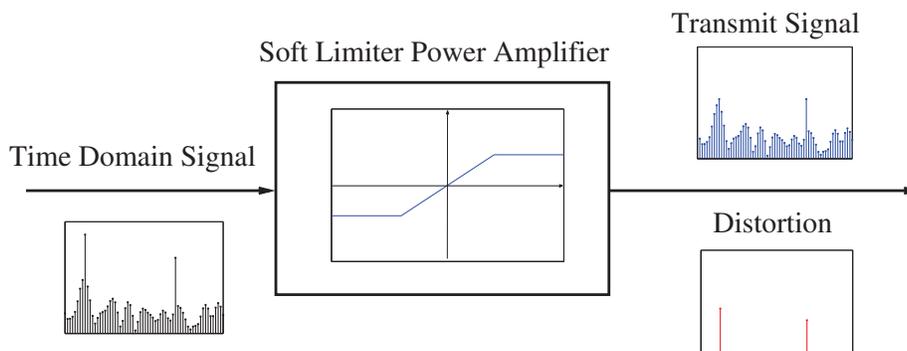


Figure 4.2: Distortion estimation example.

the time domain signal  $U$ . Signal  $U$  is of length 64 but has only 2 non-zero elements. Observation is in frequency domain and is denoted by  $V_1$  and  $V_2$ .  $V_1$  is full observation but  $V_2$  has only few Fourier coefficients (partial observation).  $U$  can be derived from  $V_1$  by taking IFFT. Interestingly,  $U$  can also be estimated from  $V_2$  by using CS technique.

### 4.3 CS-based Method for Nonlinear Distortion Mitigation

A detailed description of the proposed technique for estimation and cancellation of nonlinear distortion is now presented. We distinguish between the cases of perfect and imperfect channel estimation at the receiver. This section also presents an extension of the iterative techniques from [49] and [50] to account for imperfect channel estimation.

As an example, consider the case presented in Figure 4.2. In this figure, the time-domain signal is being clipped by the power amplifier resulting to transmit signal and distortion. The goal here is to reconstruct the distortion at the receiver by using compressed sensing.

### 4.3.1 Perfect Channel Estimation

The case of “perfect channel estimation” makes the idealized assumption that the channel impulse response is perfectly known at the receiver due to channel estimation based on previously received symbols. This means that the pilot (and possibly null) subcarriers in the current OFDM symbol can exclusively be used for mitigation of the nonlinear distortion. Let  $L_p$  denote the number of pilots and  $\mathcal{I}_p = \{i_1, \dots, i_{L_p}\}$  the index set of pilot subcarriers, respectively. Accordingly, we define  $\mathbf{Y}_p = [Y_{i_1}, \dots, Y_{i_{L_p}}]^T$ ,  $\mathbf{X}_p = [X_{i_1}, \dots, X_{i_{L_p}}]^T$ ,  $\mathbf{Z}_p = [Z_{i_1}, \dots, Z_{i_{L_p}}]^T$ ,  $\mathbf{\Lambda}_p = \mathbf{Q}_p^H \mathbf{H} \mathbf{Q}_p$ , where  $\mathbf{Q}_p$  is the  $JL \times L_p$  matrix with columns  $i_1, \dots, i_{L_p}$  of  $\mathbf{Q}$ . Using these definitions, the transmission equation (4.10) reads for the pilot subcarriers

$$\mathbf{Y}_p = \mathbf{\Lambda}_p \mathbf{X}_p + \mathbf{Q}_p^H \mathbf{H} \mathbf{d} + \mathbf{Z}_p. \quad (4.16)$$

Let us define the observation vector  $\mathbf{V}_p$  and measurement matrix  $\mathbf{\Phi}_p$  as

$$\begin{aligned} \mathbf{V}_p &= \mathbf{Y}_p - \mathbf{\Lambda}_p \mathbf{X}_p \\ \mathbf{\Phi}_p &= \mathbf{Q}_p^H \mathbf{H}. \end{aligned} \quad (4.17)$$

Then from (4.16) it follows that

$$\mathbf{V}_p = \mathbf{\Phi}_p \mathbf{d} + \mathbf{Z}_p. \quad (4.18)$$

Using (4.14), (4.15), and (4.18)  $\mathbf{d}$  can be estimated from pilot subcarriers by

$$\begin{aligned} \min \quad & \|\mathbf{d}\|_1 \\ \text{s.t.} \quad & \|\mathbf{\Phi}_p \mathbf{d} - \mathbf{V}_p\|_2 \leq \epsilon_1 \end{aligned} \quad (4.19)$$

The choice of  $\epsilon_1$  is discussed in Appendix A.

The optimization problem in (4.19) can be recast as the second order cone program (SOCP)

$$\min \sum_{i=1}^{JL} t_i \quad (4.20a)$$

$$\text{s.t. } \tilde{s}_i^2 + \tilde{s}_{i+JL}^2 \leq t_i^2, \quad 1 \leq i \leq JL \quad (4.20b)$$

$$\left\| \begin{bmatrix} \text{Re}(\Phi_p) & -\text{Im}(\Phi_p) \\ \text{Im}(\Phi_p) & \text{Re}(\Phi_p) \end{bmatrix} \tilde{\mathbf{s}} - \begin{bmatrix} \text{Re}(\mathbf{V}_p) \\ \text{Im}(\mathbf{V}_p) \end{bmatrix} \right\|_2 \leq \epsilon_1, \quad (4.20c)$$

where  $\tilde{\mathbf{s}} = [\tilde{s}_1, \dots, \tilde{s}_{2JL}]^T \in \mathbb{R}^{2JL}$ . From the solution  $\hat{\mathbf{s}}$  of (4.20) we obtain the estimated receiver side distortion vector

$$\hat{\mathbf{d}} = [\hat{s}_1, \dots, \hat{s}_{JL}]^T + j[\hat{s}_{JL+1}, \dots, \hat{s}_{2JL}]^T. \quad (4.21)$$

### 4.3.2 Imperfect Channel Estimation

We now consider the case where the channel is estimated based on the pilot symbols in the presently processed OFDM symbols. To this end, we consider minimum mean-square error (MMSE) channel estimation, and we denote the available channel estimate as  $\hat{\mathbf{H}}$ .

Since  $\hat{\mathbf{H}}$  is the only information we have about  $\mathbf{H}$ , we compute  $\mathbf{V}_p$  as

$$\begin{aligned} \mathbf{V}_p &= \mathbf{Y}_p - \hat{\Lambda}_p \mathbf{X}_p \\ &= \mathbf{Q}_p^H \mathbf{H} \mathbf{d} + (\Lambda_p - \hat{\Lambda}_p) \mathbf{X}_p + \mathbf{Z}_p, \end{aligned} \quad (4.22)$$

where  $\hat{\Lambda}_p = \mathbf{Q}_p^H \hat{\mathbf{H}} \mathbf{Q}_p$ . We note that while  $(\Lambda_p - \hat{\Lambda}_p) \mathbf{X}_p$  acts as additional noise for the estimation of  $\mathbf{d}$  from  $\mathbf{V}_p$  independent of  $\mathbf{d}$ , writing (4.22) in terms of  $\mathbf{Q}_p^H \hat{\mathbf{H}}$  would add a “data” dependent disturbance  $\mathbf{Q}_p^H (\mathbf{H} - \hat{\mathbf{H}}) \mathbf{d}$ , which is not amenable for an  $l_2$ -norm

constraint. We thus consider direct estimation of the receiver-side distortion

$$\mathbf{d}_r = \mathbf{H}\mathbf{d}, \quad (4.23)$$

using the measurement matrix  $\Phi'_p = \mathbf{Q}_p^H$ , such that (4.22) reads

$$\mathbf{V}_p = \Phi'_p \mathbf{d}_r + (\Lambda_p - \hat{\Lambda}_p) \mathbf{X}_p + \mathbf{Z}_p. \quad (4.24)$$

If  $\mathbf{d}$  is a  $\bar{\xi}$ -sparse signal, then the number of nonzero elements of  $\mathbf{d}_r$  is in the range of  $\bar{\xi} + L_m$  to  $\bar{\xi}(L_m + 1)$ <sup>8</sup> which is still sparse, especially when  $L \gg L_m$ .

Using (4.14), (4.15), and (4.24),  $\mathbf{d}_r$  can be estimated from pilot subcarriers by

$$\begin{aligned} \min \quad & \|\mathbf{d}_r\|_1 \\ \text{s.t.} \quad & \|\Phi'_p \mathbf{d}_r - \mathbf{V}_p\|_2 \leq \epsilon_2 \end{aligned} \quad (4.25)$$

where the value of  $\epsilon_2$  is calculated in Appendix A.

The estimated  $\hat{\mathbf{d}}_r$  from (4.25) is then used to refine the channel estimate  $\hat{\mathbf{H}}$ . This is done by subtracting the distortion from pilots according to

$$\hat{\mathbf{Y}}_p = \mathbf{Y}_p - \Phi'_p \hat{\mathbf{d}}_r, \quad (4.26)$$

and using  $\hat{\mathbf{Y}}_p$  for channel estimation.

To obtain better results, we iterate between channel and distortion estimation, i.e., the result of (4.25) can be used to refine the estimated channel, which then is used again in (4.22) and (4.25). The resulting iterative algorithm is summarized in Table 4.1.

---

<sup>8</sup> $(\bar{\xi} + L_m)$ -sparsity happens when all the nonzero elements of  $\mathbf{d}$  are successive.  $\bar{\xi}(L_m + 1)$ -sparsity happens when the gap between all non-zero elements of  $\mathbf{d}$  is greater than the channel length  $(L_m + 1)$ .

Table 4.1: The proposed CS-based algorithm for estimation of nonlinear distortion assuming imperfect channel knowledge.

---

- (a) Obtain initial channel estimate of the channel  $\hat{\mathbf{H}}$ . Set  $i = 1$ .
  - (b) Compute  $\mathbf{Y}_p$  from (4.22).
  - (c) Estimate  $\hat{\mathbf{d}}_r$  by solving the optimization problem (4.25).
  - (d) Obtain refined channel estimate  $\hat{\mathbf{H}}$  based on  $\hat{\mathbf{Y}}_p$  from (4.26).
  - (e) Increment the counter  $i = i + 1$  and go to Step (b) if  $i$  is less than the predefined value..
  - (f) Perform decoding according to  $\hat{\mathbf{X}} = \underset{\tilde{\mathbf{x}}}{\operatorname{argmin}} \|\hat{\mathbf{A}}\tilde{\mathbf{x}} + \mathbf{Q}^H \hat{\mathbf{d}}_r - \mathbf{Y}\|_2$ .
- 

### 4.3.3 Modification of Other Iterative Techniques

In this section, we consider the techniques proposed in [49] and [50], which we refer to as decision-aided reconstruction (DAR) and iterative estimation and cancellation (IEC), respectively. DAR and IEC were originally proposed assuming perfect knowledge of the channel impulse response. In the following, we modify these techniques to mitigate channel estimation errors due to clipping.

- Modified DAR:

DAR [49] attempts to reconstruct the time domain nonclipped signal  $\mathbf{x}$ . It can be extended to the imperfect channel estimation scenario as follows:

- (a) Set  $i = 1$  and obtain an initial estimates of the channel  $\hat{\mathbf{A}}^i$  and  $\hat{\mathbf{H}}^i$ . Set  $\hat{\mathbf{Y}}^i = \mathbf{Y}$ ,  $\hat{\mathbf{Y}}_p^i = \mathbf{Y}_p$ .
- (b) Obtain an unconstrained estimate of the transmitted signal as  $\hat{X}_{s,k}^i = \hat{Y}_k^i / \hat{\Lambda}_k^i$  for  $k = 0, \dots, L - 1$ .

- (c) Convert the estimates from Step (b) to the time domain using an IDFT, resulting in  $\hat{x}_{s,k}^i$  for  $k = 0, \dots, JL - 1$ .
- (d) Obtain decisions  $\hat{X}_k^i$  for the transmitted signal according to

$$\hat{X}_k^i = \underset{X}{\operatorname{argmin}} |\hat{Y}_k^i - \hat{\Lambda}_k^i X|, \quad k = 0, \dots, L - 1. \quad (4.27)$$

- (e) Convert the decisions from Step (d) to the time domain using an IDFT, resulting in  $\hat{x}_k^i$  for  $k = 0, \dots, JL - 1$ .
- (f) Detect the clipped samples and obtain a new signal  $\bar{x}_{s,k}^i$  according to

$$\bar{x}_{s,k}^i = \begin{cases} \hat{x}_{s,k}^i & |\hat{x}_k^i| \leq A \\ \hat{x}_k^i & |\hat{x}_k^i| > A \end{cases} \quad (4.28)$$

for  $k = 0, \dots, JL - 1$ .

- (g) Estimate the distortion vector  $\hat{\mathbf{d}}^i$  as  $\hat{d}_k^i = \hat{x}_{s,k}^i - \bar{x}_{s,k}^i$  for  $k = 0, \dots, JL - 1$ .
- (h) Update the received signal vector according to

$$\hat{\mathbf{Y}}_p^{i+1} = \hat{\mathbf{Y}}_p^i - \mathbf{Q}_p^H \hat{\mathbf{H}}^i \hat{\mathbf{d}}^i. \quad (4.29)$$

- (i) Use  $\hat{\mathbf{Y}}_p^{i+1}$  to obtain the new channel estimation which is denoted by  $\hat{\mathbf{\Lambda}}^{i+1}$  and  $\hat{\mathbf{H}}^{i+1}$ .
- (j) Convert the sequence  $\bar{x}_{s,k}^i$  to the frequency domain signal  $\bar{\mathbf{X}}_s^i$  using DFT.
- (k) Update  $\hat{\mathbf{Y}}^i$  as  $\hat{\mathbf{Y}}^{i+1} = \hat{\mathbf{\Lambda}}^{i+1} \bar{\mathbf{X}}_s^i$ .
- (l) Increment the counter  $i = i + 1$  and go to Step (b).

- Modified IEC:

IEC [50] estimates the clipping noise  $\mathbf{d}$  and cancels it from the received signal assuming perfect channel estimation. Similarly, assuming imperfect channel knowledge, modified IEC can be summarized as:

- (a) Set  $i = 1$  and obtain an initial estimates of the channel  $\hat{\mathbf{\Lambda}}^i$  and  $\hat{\mathbf{H}}^i$ . Set  $\hat{\mathbf{Y}}^i = \mathbf{Y}$ ,  $\hat{\mathbf{Y}}_p^i = \mathbf{Y}_p$ .

- (b) Obtain symbol decisions  $\hat{X}_k^i$  according to

$$\hat{X}_k^i = \underset{X}{\operatorname{argmin}} |\hat{Y}_k^i - \hat{\Lambda}_k^i X|, \quad k = 0, \dots, L - 1. \quad (4.30)$$

- (c) Convert the decisions from Step (b) to the time domain using an IDFT, resulting in  $\hat{x}_k^i$  for  $k = 0, \dots, JL - 1$ .

- (d) Generate the clipped time-domain signal  $\hat{x}_{s,k}^i$  by inserting  $\hat{x}_k^i$  in (4.5) or (4.7).

- (e) Convert  $\hat{\mathbf{x}}_s^i$  to the frequency-domain signal  $\hat{\mathbf{X}}_s^i$  using DFT.

- (f) Estimate the frequency-domain clipping noise as  $\hat{\mathbf{D}}^i = \hat{\mathbf{X}}_s^i - \hat{\mathbf{X}}^i$ .

- (g)  $\hat{\mathbf{D}}^i$  is converted to the time-domain using an IDFT, resulting in  $\hat{\mathbf{d}}^i$ .

- (h) Update  $\hat{\mathbf{Y}}_p^i$  according to

$$\hat{\mathbf{Y}}_p^{i+1} = \hat{\mathbf{Y}}_p^i - \mathbf{Q}_p^H \hat{\mathbf{H}}^i \hat{\mathbf{d}}^i. \quad (4.31)$$

- (i) Use  $\hat{\mathbf{Y}}_p^{i+1}$  to obtain the new channel estimation  $\hat{\mathbf{\Lambda}}^{i+1}$  and  $\hat{\mathbf{H}}^{i+1}$ .

- (j) Update  $\hat{\mathbf{Y}}^i$  as  $\hat{\mathbf{Y}}^{i+1} = \hat{\mathbf{Y}}^i - \hat{\mathbf{\Lambda}}^{i+1} \hat{\mathbf{D}}^i$ .

- (k) Increment the counter  $i = i + 1$  and go to Step (b).

## 4.4 Numerical Results

In this section, we evaluate the performance of the proposed CS-based approach to mitigate nonlinear distortion in OFDM transmission. The OFDM signal is oversampled by a factor of  $J = 4$ , and QAM subcarrier modulation is employed. We consider both fixed constellation sizes and also systems employing the bit loading algorithm proposed in [35], which improves the BER subject to a given average data rate and transmit power. The number of bits per subcarrier are chosen from the range of 0 to 6, i.e., 64QAM is the largest available constellation. In our simulations, the number of subcarriers is set to  $L = 64$  and the rate- $\frac{1}{2}$  convolutional encoder defined by the generator polynomials  $(133, 171)_8$  and Viterbi decoding are used.  $\bar{E}_b$  is the average received power per bit and  $\mathcal{N}_0$  is the one-sided noise power spectral density. As suggested in [49] and [50], for the simulation results, the original and modified versions of the DAR and IEC techniques have been run 3 and 2 times, respectively.

For the frequency-selective fading channel, we adopt the quasi-static exponentially-decaying multipath Rayleigh fading model which is explained in Section 2.2.3. Throughout the following, we assume  $T_r = 50$  ns,  $T_{\text{rms}} = 150$  ns and  $L_m = 4$ .

Figure 4.3 plots the mean square error (MSE) of the proposed technique for the perfect channel estimation scenario versus the clipping level of the SL power amplifier for different numbers of pilot tones  $L_p = 8, \dots, 12$ . For this experiment, coded OFDM with the 64-QAM constellation for all subcarriers is used. The simulation results suggest that increasing  $L_p$  beyond 10 does not improve MSE in this case. A choice of  $L_p = 10$  would seem to be a practical solution; for example IEEE 802.11a/g, which has  $L = 64$  subcarriers, uses 16 pilots [8]. We adopt  $L_p = 10$  for the rest of the simulation results.

Figure 4.4 compares the estimated nonlinear distortion ( $\hat{\mathbf{d}}$ ) obtained by CS with the original distortion ( $\mathbf{d}$ ) for  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0) = 8$  dB. An SL power amplifier model and dif-

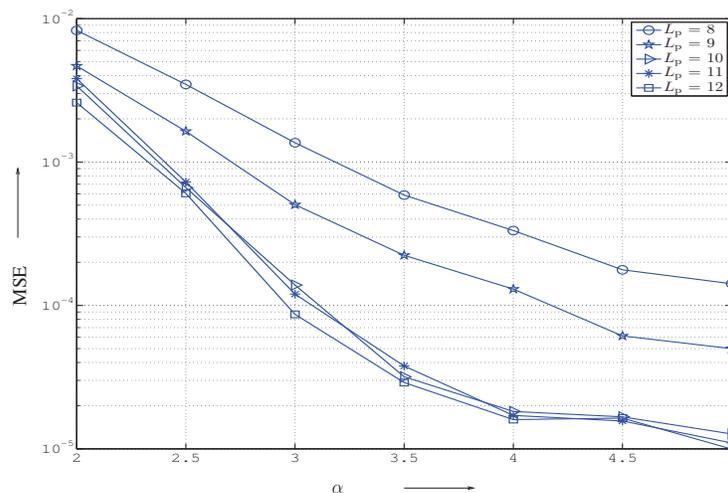


Figure 4.3: MSE versus clipping level ( $\alpha$ ) of SL power amplifier for systems employing different number of pilots ( $L_p$ ). Coded OFDM with 64-QAM constellation for all subcarriers.

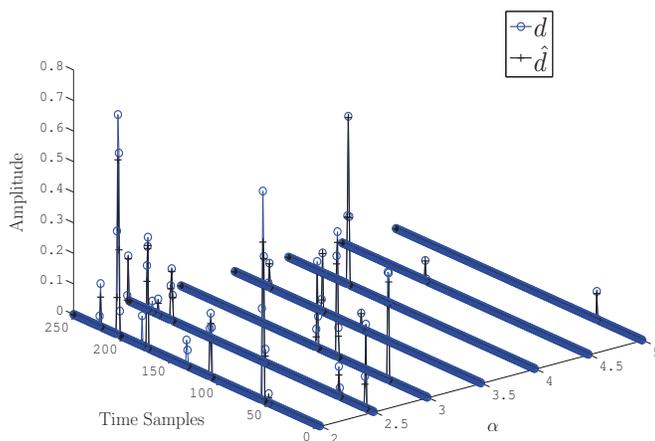


Figure 4.4: The original distortion ( $\mathbf{d}$ ) and the estimated distortion ( $\hat{\mathbf{d}}$ ) for one specific OFDM symbol for SL power amplifier with different values of  $\alpha$ ,  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0) = 8\text{dB}$ .

ferent values of  $\alpha$  at the transmitter and perfect channel estimation at the receiver are assumed. We observe that the discrepancy between the true and estimated distortion becomes negligible for  $\alpha \geq 2.5\text{dB}$ , which confirms the accuracy of the CS based distortion

estimation. The lower estimation error for increasing  $\alpha$ , as also suggested by Figure 4.3, is due to the smaller number of clipping events and thus sparser distortion vector in this case.

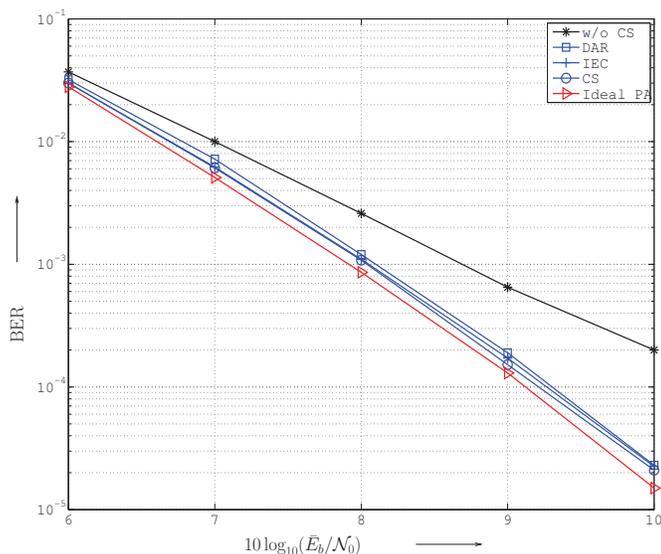


Figure 4.5: BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for perfect channel estimation and an SL power amplifier with  $\alpha = 3\text{dB}$ . Coded and bit-loaded OFDM with an average of 4 bits per subcarrier.

In Figure 4.5, the average BER achieved with the proposed CS-based and the benchmark DAR and IEC techniques are plotted versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for the case of perfect channel estimation and the SL power amplifier model with  $\alpha = 3\text{dB}$ . For this figure and Figures 4.6, 4.7, 4.8, and 4.9, coded and bit-loaded OFDM transmission with an average of 4 bits per subcarrier is used. “Ideal PA” refers to the case where the power amplifier does not clip the transmitted signal. It can be seen that the proposed method provides a small advantage over the benchmark methods in this case. For the same setup but the SSPA model with  $\rho = 1$ ,  $p = 6$  and  $\text{IBO} = 5\text{dB}$ , the BER curves are shown in Figure 4.6. Again, all three methods accomplish effective cancellation of the nonlinear distortion, with

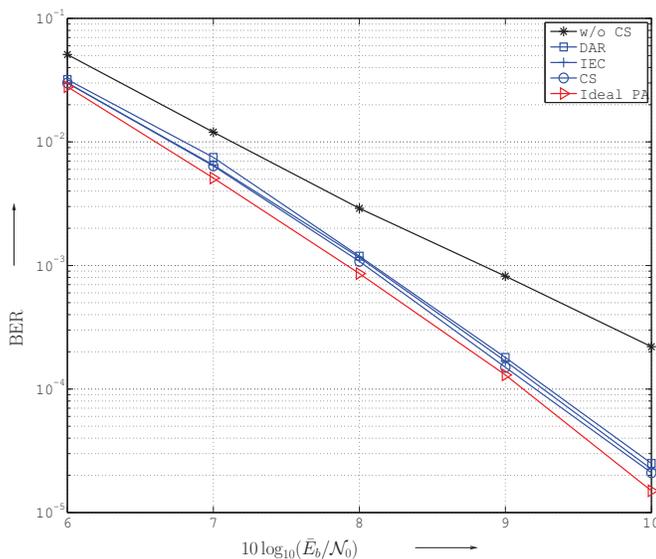


Figure 4.6: BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for perfect channel estimation and an SSPA with  $\rho = 1$ ,  $p = 6$  and  $\text{IBO} = 5\text{dB}$ . Coded and bit-loaded OFDM with an average of 4 bits per subcarrier.

a slight advantage for the CS-based method.

We now turn to the case where the channel is estimated based on the pilot subcarriers in the currently detected OFDM symbol. Figure 4.7 shows the MSE for channel estimation versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for the modified DAR and IEC techniques and the proposed CS method as well as for the case when an ideal power amplifier is used. The SL power amplifier model with  $\alpha = 2.5\text{dB}$  is applied. It can be seen that the proposed CS-based method significantly improves the MSE compared to conventional channel estimation, and it also proves to be superior compared to modified DAR and IEC. Figure 4.8 shows the average BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for the three distortion mitigation methods in this case. We observe that the CS-based method achieves notable performance gains over the benchmark methods. Especially the modified DAR technique does not provide much gain over detection with unprocessed distortion and falls behind compared to modified IEC

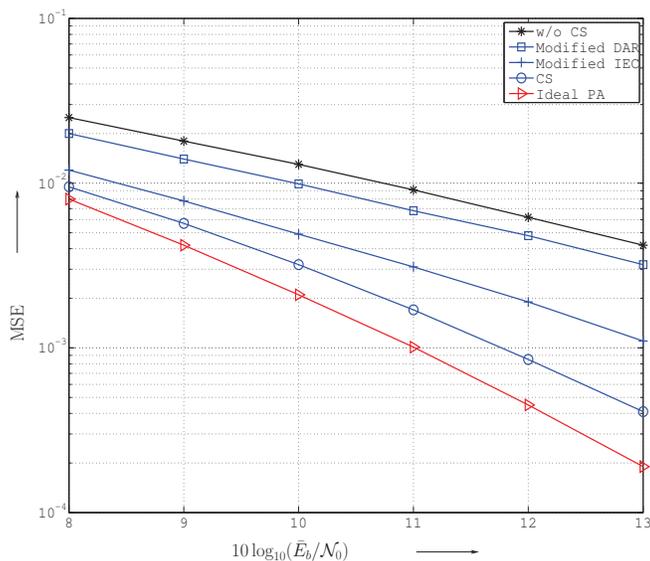


Figure 4.7: MSE versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for imperfect channel estimation when MMSE channel estimation, an SL power amplifier with  $\alpha = 2.5\text{dB}$ , and coded and bit-loaded OFDM with an average of 4 bits per subcarrier are used.

and CS. The CS reconstruction has been run 3 times here, but interestingly the results almost do not change by the second and third iteration compared to the first iteration. In Figure 4.9, the BER curves versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for the same system setup but the SSPA model with  $\rho = 1$ ,  $p = 8$  and IBO = 6dB are plotted. Again, CS outperforms the other two benchmark techniques, especially DAR.

## 4.5 Summary

In this chapter, we have considered the problem of mitigating the distortion experienced at an OFDM receiver due to nonlinear power amplification at the transmitter. We have developed a compressed sensing technique making use of pilot subcarriers to estimate the distortion. Two cases of OFDM transmission have been distinguished. The first case assumes

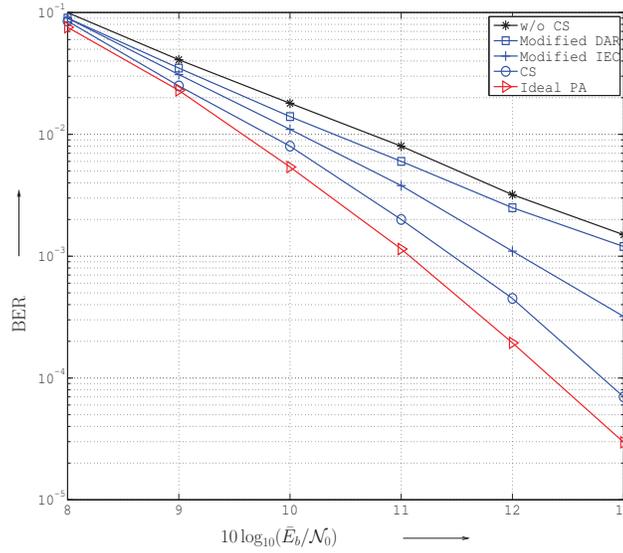


Figure 4.8: BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for imperfect channel estimation when MMSE channel estimation, an SL power amplifier with  $\alpha = 2.5\text{dB}$ , and coded and bit-loaded OFDM with an average of 4 bits per subcarrier are used.

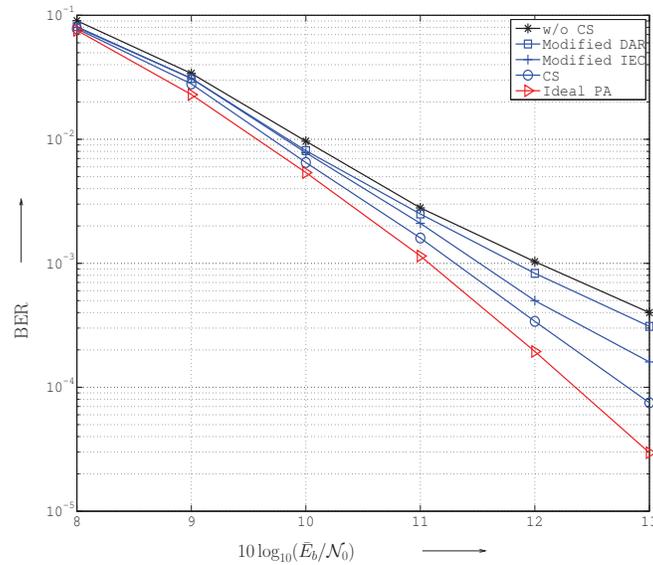


Figure 4.9: BER versus  $10 \log_{10}(\bar{E}_b/\mathcal{N}_0)$  for imperfect channel estimation when MMSE channel estimation, an SSPA with  $\rho = 1$ ,  $p = 8$  and IBO = 6dB, and coded and bit-loaded OFDM with an average of 4 bits per subcarrier are used.

the receiver knows the channel perfectly. The second case considers practical channel estimation scenarios when pilot tones are also distorted. The two previously proposed methods from [49] and [50] have been modified to be applicable to the case of imperfect channel estimation such that a fair performance comparison could be made. The presented numerical results have demonstrated (i) that the proposed compressed sensing based method can accurately estimate the clipping events for different values of clipping levels, (ii) that our method works as good as previous techniques for perfect channel estimation, and (iii) that our detection method notably outperforms the modified versions of [49] and [50] for imperfect channel estimation.

# Chapter 5

## Conclusions and Future Work

In this final chapter, we summarize our results and highlight the contributions of this dissertation. We also suggest topics and open problems for further research.

### 5.1 Research Contributions

This thesis focused on the analysis and design of OFDM systems, namely (1) performance analysis of BIC-OFDM systems; (2) adaptive techniques for performance enhancement of BIC-OFDM systems; and (3) a new technique based on compressed sensing to mitigate the nonlinear distortion caused by the power amplifier.

In Chapter 2, we derived a novel analytical method for the performance of bit-loaded BIC-MIMO-OFDM systems for a specific known interleaver operating over a quasi-static fading channel. We then used this formula to design new adaptive bit-loading, adaptive interleaving, and adaptive coded modulation techniques. In the case of bit-loading, we used the proposed formula and calculated BER for different bit-loading techniques, which were originally proposed for uncoded OFDM systems, and selected the one with the lowest BER. We noticed that the best loading depends on the specific channel and system parameters. For adaptive interleaving, we proposed three adaptive interleaving techniques: (1) In the first technique, we select the best interleaver among a set of predefined interleavers; (2) the second technique, uses the PEPs derived using our formula, and designs the interleaver accordingly; and (3) in the third technique, we propose an adaptive interleaver for MIMO-

SVD-OFDM systems. Finally, the proposed adaptive coded and modulation technique uses our BER formula and selects the best combination of code rate and constellation size. Numerical results confirm the accuracy of the proposed performance evaluation formula and the usefulness of the proposed adaptive technique.

The BER formula proposed in Chapter 2 might be too complex to evaluate in some applications. Therefore, we proposed a BER formula for BIC-OFDM operating over a quasi-static fading channel based on the PDF of the reliability metrics in Chapter 3. Then, we simplified this BER assuming random interleaving. This formula was then used to formulate three main optimization problems: BER minimization, transmit power minimization, and throughput maximization. We found optimal bit loading, power loading, and code rate selection schemes for each case. Finally, we proposed a simple formula for the performance of the optimized system. Numerical results showed the gains by using our adaptive techniques.

In Chapter 4, we considered the problem of mitigating the nonlinear distortion at the receiver caused by a nonideal power amplifier at the transmitter. We used the recently proposed compressed sensing technique to estimate the clippings. Two scenarios were considered. In the first scenario, it's assumed that the channel is perfectly known at the receiver. In the second scenario, we considered the realistic case where the channel is estimated at the receiver using pilots and pilots are also affected by the power amplifier nonlinearity. Numerical results showed that our proposed technique works as good as existing techniques for the case of perfect channel estimation but it significantly outperforms those techniques for the imperfect channel case.

## 5.2 Future Work

The work presented in this thesis can be extended in several ways. In the following, we present a (by no means comprehensive) list.

The BER analysis presented in Chapters 2 and 3 can be extended to MIMO-OFDM systems using space-time codes like V-BLAST [63]. The analysis can then be used to design new algorithms for these systems.

The analysis presented in Chapter 2 can be extended to include the effects of practical power amplifiers into the error rate analysis. Previous analysis of this effect (1) makes simplified assumptions about the nonlinear distortions and the transmission channel [64] and (2) considers only uncoded OFDM. The first step is to use our analysis to approximate the error rate for transmission over nonlinear channels. In the second step, PAPR reduction methods which are apt for coded transmission can be designed using the analysis. For example, clipping of a single large signal peak is less detrimental for coded transmission than clipping of several moderately large peaks.

A further open area is the use of compressed sensing in telecommunications. There has been recently a plethora of research on this topic. The interested reader can refer to [65] for a complete list. More specifically, our work in Chapter 4 can be extended to the case when clipping is not a rare event and therefore clipping vector is not sparse. Another area is to use the CS technique for cross-talk cancellation in xDSL systems.

As a more theoretical extension, it would be very useful to derive a deterministic method for selecting the rows of the Fourier matrix so that the resulting matrix satisfies UUP for practical values of the number of rows like  $2^6$ ,  $2^7$ , or  $2^8$ . As stated before, the methods proposed in [56, 62] cannot be used for practical values of the number of rows.

# Bibliography

- [1] S.-W. Lei and V.K.L. Lau, "Performance Analysis of Adaptive Interleaving for OFDM Systems," *IEEE Trans. Veh. Technol.*, vol. 51, no. 3, pp. 435–444, May 2002.
- [2] C. Stierstorfer and R.F.H. Fischer, "Adaptive Interleaving for Bit-Interleaved Coded Modulation," in *Proc. 7th International ITG Conference on Source and Channel Coding*, Ulm, Germany, Jan. 2008.
- [3] R. W. Chang, "Synthesis of Band-Limited Orthogonal Signals for Multichannel Data Transmission," in *Bell Systems Tech. J.*, Dec. 1966, pp. 1775–1796.
- [4] S. B. Weinstein and P. Ebert, "Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform," *IEEE Trans. Commun. Technol.*, vol. 19, pp. 628–634, Oct. 1971.
- [5] A. Peled and A. Ruiz, "Frequency Domain Data Transmission Using Reduced Computational Complexity Algorithms," in *Proc. IEEE ICASSP*, Denver, Colorado, Apr. 1980, pp. 964–967.
- [6] J.A.C. Bingham, "Multicarrier Modulation for Data Transmission: An Idea Whose Time Has Come," *IEEE Commun. Mag.*, pp. 5–14, May 1990.
- [7] W. Y. Zou and Y. Wu, "COFDM: An Overview," *IEEE Trans. Broadcast.*, vol. 41, no. 1, pp. 1–8, Mar. 1995.
- [8] IEEE 802.11, "Wireless LAN Medium Access Control (MAC) Physical Layer (PHY) Specifications, Amendment 1: High-Speed Physical Layer in the 5 GHz Band," Jul. 1999.
- [9] ECMA, "Standard ECMA-368: High Rate Ultra Wideband PHY and MAC Standard," Dec. 2005, [Online]: <http://www.ecma-international.org/publications/standards/Ecma-368.htm>.
- [10] R. Gallager, *Information Theory and Reliable Communication*. New York, NY: Wiley, 1968.
- [11] S. Wei, D. L. Goeckel, and P. A. Kelly, "Convergence of the Complex Envelope of Bandlimited OFDM Signal," *IEEE Trans. Inform. Theory*, vol. 56, no. 10, pp. 4893–4904, Oct. 2010.

- 
- [12] S. Han and J. Lee, "An Overview of Peak-to-Average Power Ratio Reduction Techniques for Multicarrier Transmission," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 56–65, Apr. 2005.
- [13] A. Kenarsari-Anhari and L. Lampe, "An Analytical Approach for Performance Evaluation of BICM Transmission Over Nakagami-m Fading Channels," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1090–1101, April 2010.
- [14] Z. Song, K. Zhang, and Y. Guan, "Statistical Adaptive Modulation for QAM-OFDM Systems," in *Proc. IEEE Globecom*, Taipei, Taiwan, R.O.C., Nov. 2002, pp. 706–710.
- [15] P. Xia, S. Zhou, and G. B. Giannakis, "MIMO-OFDM with ST Coding and Beamforming Adapted to Partial CSI," in *Proc. 37th Conf. Information Sciences Systems*, Baltimore, MD, Mar. 2003.
- [16] S. Ye, R. Blum, and L. Cimini, "Adaptive Modulation for Variable Rate OFDM Systems with Imperfect Channel Information," in *Proc. IEEE Vehicular Technology Conf.*, vol. 2, Birmingham, AL, May 2002, pp. 767–771.
- [17] C. Snow, L. Lampe, and R. Schober, "Error Rate Analysis for Coded Multicarrier Systems over Quasi-Static Fading Channels," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1736–1746, Sep. 2007.
- [18] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
- [19] G. Caire, G. Taricco, and E. Biglieri, "Bit-Interleaved Coded Modulation," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.
- [20] K.-B. Song and A. Ekbal and S.T. Chung and J. M. Cioffi, "Adaptive Modulation and Coding (AMC) for Bit-Interleaved Coded OFDM (BIC-OFDM)," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1685–1694, Jul. 2006.
- [21] C.K. Sung and S.-Y. Chung and J. Heo and I. Lee, "Adaptive Bit-Interleaved Coded OFDM With Reduced Feedback Information," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1649–1655, Sep. 2007.
- [22] A. Martinez, A. Guillén i Fàbregas, and G. Caire, "Error Probability Analysis of Bit-Interleaved Coded Modulation," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 262–271, Jan. 2006.
- [23] B. Canpolat and Y. Tanik, "Performance Analysis of Adaptive Loading OFDM Under Rayleigh Fading," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1105–1115, Jul. 2004.
- [24] A. Seyedi and G. J. Saulnier, "Symbol-Error Rate Analysis of Fischer's Bit-Loading Algorithm," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1480–1483, Sep. 2004.

- [25] R.F.H. Fischer and J. B. Huber, "A New Loading Algorithm for Discrete Multitone Transmission," in *Proc. IEEE GLOBECOM*, vol. 1, London, England, Nov. 1996, pp. 724–728.
- [26] I. Stupia, L. Vandendorpe, J. Louveaux, F. Giannetti, V. Lottici, and N. D'Andrea, "Power Allocation for Goodput Optimization in BICM-OFDM Systems," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 3604–3608.
- [27] H. Moon and D. C. Cox, "Efficient Power Allocation for Coded OFDM Systems," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 943–947, Apr. 2009.
- [28] S. Nagaraj, "Improving Power Efficiency of Coded OFDM With Feedback," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2541–2544, Sept. 2009.
- [29] D. Matas and M. Lamarca, "Optimum Power Allocation and Bit Loading for BICM Systems," *IEEE Trans. Commun.*, vol. 58, no. 8, pp. 2314–2323, Aug. 2010.
- [30] J. Campello, "Optimal Discrete Bit Loading for Multicarrier Modulation Systems," in *Proc. IEEE ISIT*, Cambridge, MA, USA, Aug. 1998, p. 193.
- [31] J. Campello, "Practical bit loading for DMT," in *Proc. IEEE International Conference on Communications (ICC)*, vol. 2, Vancouver, Canada, June 1999, pp. 801–805.
- [32] Y. George and O. Amrani, "Bit Loading Algorithms for OFDM," in *Proc. IEEE ISIT*, Chicago, USA, June 2004, pp. 391–391.
- [33] A. Mahmood and J. C. Belfiore, "An Efficient Algorithm for Optimal Discrete Bit-Loading in Multicarrier Systems," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1627–1630, June 2010.
- [34] P. Ormeci, X. Liu, D. Goeckel, and R. Wesel, "Adaptive Bit-Interleaved Coded Modulation," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1572–1581, Sept 2001.
- [35] P. S. Chow and J. M. Cioffi and J. A.C. Bingham, "A Practical Discrete Multitone Transceiver Loading Algorithm for Data Transmission over Spectrally Shaped Channels," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 773–775, Feb./Mar./Apr. 1995.
- [36] X. Liu and R. D. Wesel, "Profile Optimal 8-QAM and 32-QAM Constellations," in *Proc. 36th Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, USA, 1998, pp. 136–145.
- [37] A. Batra, J. Balakrishnan, G. Aiello, J. Foerster, and A. Dabak, "Design of a Multi-band OFDM System for Realistic UWB Channel Environments," *IEEE Trans. Microwave Theory Tech.*, vol. 52, no. 9, pp. 2123–2138, Sep. 2004.
- [38] A. F. Molisch, J. R. Foerster, and M. Pendergrass, "Channel Models for Ultrawideband Personal Area Networks," *IEEE Wireless Commun. Mag.*, pp. 14–21, Dec. 2003.

- [39] A. Saleh and R. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 2, pp. 128–137, Feb. 1987.
- [40] D. Hughes-Hartogs, "Ensemble Modem Structure for Imperfect Transmission Media," U.S. Patents Nos. 4,679,227 (July 1987), 4,731,816 (Mar. 1988), and 4,833,706 (May 1989).
- [41] L. Piazzo, "Fast Algorithm for Power and Bit Allocation in OFDM Systems," *IEE Electronics Letters*, vol. 35, no. 25, pp. 2173–2174, Dec. 1999.
- [42] Z. Luo and H. Gao and Y. Liu, "Adaptive Transmission With Linear Computational Complexity in MIMO-OFDM Systems," *IEEE Trans. Commun.*, vol. 55, no. 10, pp. 1873–1877, Oct. 2007.
- [43] E. Zehavi, "8-PSK Trellis Codes for a Rayleigh Channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873–884, May 1992.
- [44] A. Viterbi, "An Intuitive Justification and a Simplified Implementation of the MAP Decoder for Convolutional Codes," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 260–264, Feb 1998.
- [45] B. Classon, K. Blankenship, and V. Desai, "Channel Coding for 4G Systems With Adaptive Modulation and Coding," *IEEE Wireless Communications*, vol. 9, no. 2, pp. 8–13, April 2002.
- [46] A. Guillén i Fàbregas, A. Martinez, and G. Caire, "Bit-Interleaved Coded Modulation," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 1-2, pp. 1–153, 2008.
- [47] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications, 1972.
- [48] K.-B. Song, private communication, 2009.
- [49] D. Kim and G. L. Stuber, "Clipping Noise Mitigation for OFDM by Decision-Aided Reconstruction," *IEEE Commun. Lett.*, vol. 3, no. 1, pp. 4–6, Jan. 1999.
- [50] H. Chen and A. Haimovich, "Iterative Estimation and Cancellation of Clipping Noise for OFDM Signals," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 305–307, Jul. 2003.
- [51] J. Tellado, L. Hoo, and J. Cioffi, "Maximum-Likelihood Detection of Nonlinearly Distorted Multicarrier Symbols by Iterative Decoding," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 218–228, Feb. 2003.
- [52] H. Saeedi, M. Sharif, and F. Marvasti, "Clipping Noise Cancellation in OFDM Systems Using Oversampled Signal Reconstruction," *IEEE Commun. Lett.*, vol. 6, no. 2, pp. 73–75, Feb. 2002.

- [53] Y. Xiao, S. Li, X. Lei, and Y. Tang, "Clipping Noise Mitigation for Channel Estimation in OFDM Systems," *IEEE Commun. Lett.*, vol. 10, no. 6, pp. 474–476, Jun. 2006.
- [54] A. R. S. Bahai, M. Singh, A. J. Goldsmith, and B. R. Saltzberg, "A New Approach for Evaluating Clipping Distortion in Multicarrier Systems," *IEEE J. Select. Areas Commun.*, vol. 20, no. 5, pp. 1037–1047, Jun. 2002.
- [55] E. J. Candès and M. B. Wakin, "An Introduction to Compressive Sampling," *IEEE Signal Processing Mag.*, pp. 21–30, Mar. 2008.
- [56] G. Caire, T. Y. Al-Naffouri, and A. K. Narayanan, "Impulse Noise Cancellation in OFDM: An Application of Compressed Sensing," in *Proc. IEEE ISIT*, Toronto, Canada, Jul. 2008, pp. 1293–1298.
- [57] D. Dardari, V. Tralli, and A. Vaccari, "A Theoretical Characterization of Nonlinear Distortion Effects in OFDM Systems," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1755–1764, Oct. 2000.
- [58] C. Rapp, "Effects of HPA-Nonlinearity on a 4-DPSK/OFDM-Signal for a Digital Sound Broadcasting System," in *2nd European Conference on Satellite Communications*, Liege, Belgium, Oct. 1991, pp. 179–184.
- [59] M. Friese, "On the Degradation of OFDM Signals due to Peak-Clipping in Optimally Predistorted Power Amplifiers," in *Proc. IEEE Globecom*, Sydney, Australia, Nov. 1998, pp. 939–944.
- [60] E. J. Candès, "Compressive Sampling," in *International Congress on Mathematics*, 2006.
- [61] E. J. Candès and T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [62] N. Y. Yu, "Deterministic Construction of Partial Fourier Compressed Sensing Matrices via Cyclic Difference Set," in <http://arxiv.org/pdf/1008.0885>.
- [63] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multiple Antennas," *Bell Laboratories Technical Journal*, 1(2):41-59, Autumn 1996.
- [64] H. Nikopour and S. Jamali, "On the Performane of OFDM Systems Over a Cartesian Clipping Channel: A Theoretical Approach," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2083–2096, Nov. 2004.
- [65] <http://dsp.rice.edu/cs>.

- [66] E. J. Candès, J. Romberg, and T. Tao, “Stable Signal Recovery from Incomplete and Inaccurate Measurements,” in *Communications on Pure and Applied Mathematics*, vol. 59, Aug. 2006, pp. 1207–1223.
- [67] M. Noh, Y. Lee, and H. Park, “Low Complexity LMMSE Channel Estimation for OFDM,” in *IEE Proc.-Commun*, vol. 153, no. 5, Oct. 2006, pp. 645–650.

# Appendix A

## Closed-form Expression for $\epsilon_1$ and $\epsilon_2$ for CS Reconstruction

In this section, we derive expressions for the values of  $\epsilon_1$  and  $\epsilon_2$  which are used in (4.19) and (4.25). In particular, following [66], we derive estimations for the mean and standard deviation of the additive noise present in the CS-estimation step, based on which  $\epsilon_1$  and  $\epsilon_2$  are determined.

### A.1 Perfect Channel Estimation

This case is similar to the case considered in [66] except that estimation noise  $\mathbf{Z}_p$  in (4.18) is complex-valued. Each element of  $\mathbf{Z}_p$  is complex Gaussian distributed with zero mean and variance  $\sigma^2$ . Therefore,  $\|\mathbf{Z}_p\|_2^2$  has a chi-square distribution with mean  $2L_p\sigma^2$  and standard deviation  $2\sqrt{2L_p}\sigma^2$ . Then, following [66], we choose  $\epsilon_1$  according to

$$\begin{aligned}\epsilon_1^2 &= 2L_p\sigma^2 + \lambda_1 2\sqrt{2L_p}\sigma^2 \\ &= 2\sigma^2 \left( L_p + \lambda_1 \sqrt{2L_p} \right)\end{aligned}\tag{A.1}$$

with  $\lambda_1 = 2$ .

## A.2 Imperfect Channel Estimation

Different from the perfect channel estimation case, in this case the channel estimation error is also contributing to the total noise. According to (4.22),  $\hat{\mathbf{n}} = (\mathbf{\Lambda}_p - \hat{\mathbf{\Lambda}}_p)\mathbf{X}_p + \mathbf{Z}_p$  is the total estimation noise.  $\hat{\mathbf{\Lambda}}_p = \text{diag}(\hat{\mathbf{H}}_p)$  and  $\hat{\mathbf{H}}_p$  is the LMMSE channel estimation given by [67]

$$\hat{\mathbf{H}}_p = \mathbf{R}_{\mathbf{H}_p\mathbf{H}_p} \left( \mathbf{R}_{\mathbf{H}_p\mathbf{H}_p} + \frac{\tau}{\gamma_s} \mathbf{I}_{L_p} \right)^{-1} \hat{\mathbf{H}}_{\text{LS}}, \quad (\text{A.2})$$

where the  $L_p \times L_p$   $\mathbf{R}_{\mathbf{H}_p\mathbf{H}_p}$  is the channel autocorrelation matrix corresponding to the pilot tones,  $\mathbf{I}_{L_p}$  is the identity matrix of size  $L_p$ ,  $\tau$  is a constant which only depends on the modulation,  $\gamma_s$  is the operating SNR, and  $\hat{\mathbf{H}}_{\text{LS}}$  is the  $L_p \times 1$  least square (LS) channel estimation given by

$$\hat{\mathbf{H}}_{\text{LS}} = \left[ \frac{Y_{i_1}}{X_{i_1}}, \dots, \frac{Y_{i_{L_p}}}{X_{i_{L_p}}} \right]^T. \quad (\text{A.3})$$

By defining  $\mathbf{H}_p$  as the diagonal elements of  $\mathbf{\Lambda}_p$ ,  $\hat{\mathbf{n}}$  can be rewritten as  $\hat{\mathbf{n}} = \text{diag}(\mathbf{X}_p)(\mathbf{H}_p - \hat{\mathbf{H}}_p) + \mathbf{Z}_p$ . Assuming that  $\mathbf{e} = \mathbf{H}_p - \hat{\mathbf{H}}_p$  and  $\mathbf{Z}_p$  are independent,  $\hat{\mathbf{n}}$  has a multivariate normal distribution with zero mean and covariance matrix given by

$$\mathbf{R}_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = \text{diag}(\mathbf{X}_p)\mathbf{R}_{\mathbf{e}\mathbf{e}}\text{diag}(\mathbf{X}_p)^H + \sigma^2\mathbf{I}_{L_p} \quad (\text{A.4})$$

where  $\sigma^2$  is the variance of each element of  $\mathbf{Z}_p$  and  $\mathbf{R}_{\mathbf{e}\mathbf{e}}$  is the covariance matrix of  $\mathbf{e}$  given by [67]

$$\mathbf{R}_{\mathbf{e}\mathbf{e}} = \mathbf{R}_{\mathbf{H}_p\mathbf{H}_p} - \mathbf{R}_{\mathbf{H}_p\mathbf{H}_p} \left( \mathbf{R}_{\mathbf{H}_p\mathbf{H}_p} + \frac{\tau}{\gamma_s} \mathbf{I}_{L_p} \right)^{-1} \mathbf{R}_{\mathbf{H}_p\mathbf{H}_p}. \quad (\text{A.5})$$

We denote the element in the  $i$ th row and  $j$ th column of  $\mathbf{R}_{\hat{\mathbf{n}}\hat{\mathbf{n}}}$  by  $\Omega_{i,j}$ . Using (A.4),  $\mathbb{E}(\|\hat{\mathbf{n}}\|_2^2)$  can be calculated as

$$\mathbb{E}(\|\hat{\mathbf{n}}\|_2^2) = \text{trace}(\mathbf{R}_{\hat{\mathbf{n}}\hat{\mathbf{n}}}), \quad (\text{A.6})$$

and  $\mathbb{E}(\|\dot{\mathbf{n}}\|_2^4)$  can be expressed as

$$\mathbb{E}(\|\dot{\mathbf{n}}\|_2^4) = \sum_{i=1}^{L_p} \sum_{j=1}^{L_p} (\Omega_{i,i}\Omega_{j,j} + \|\Omega_{i,j}\|^2) . \quad (\text{A.7})$$

By using (A.6) and (A.7), the standard deviation of  $\|\dot{\mathbf{n}}\|_2^2$  ( $\text{std}(\|\dot{\mathbf{n}}\|_2^2)$ ) can be calculated as

$$\text{std}(\|\dot{\mathbf{n}}\|_2^2) = \sqrt{\mathbb{E}(\|\dot{\mathbf{n}}\|_2^4) - (\mathbb{E}(\|\dot{\mathbf{n}}\|_2^2))^2} . \quad (\text{A.8})$$

Using the same argument as before,  $\epsilon_2$  is chosen as

$$\epsilon_2^2 = \mathbb{E}(\|\dot{\mathbf{n}}\|_2^2) + \lambda_2 \text{std}(\|\dot{\mathbf{n}}\|_2^2) \quad (\text{A.9})$$

with  $\lambda_2 = 2$ .