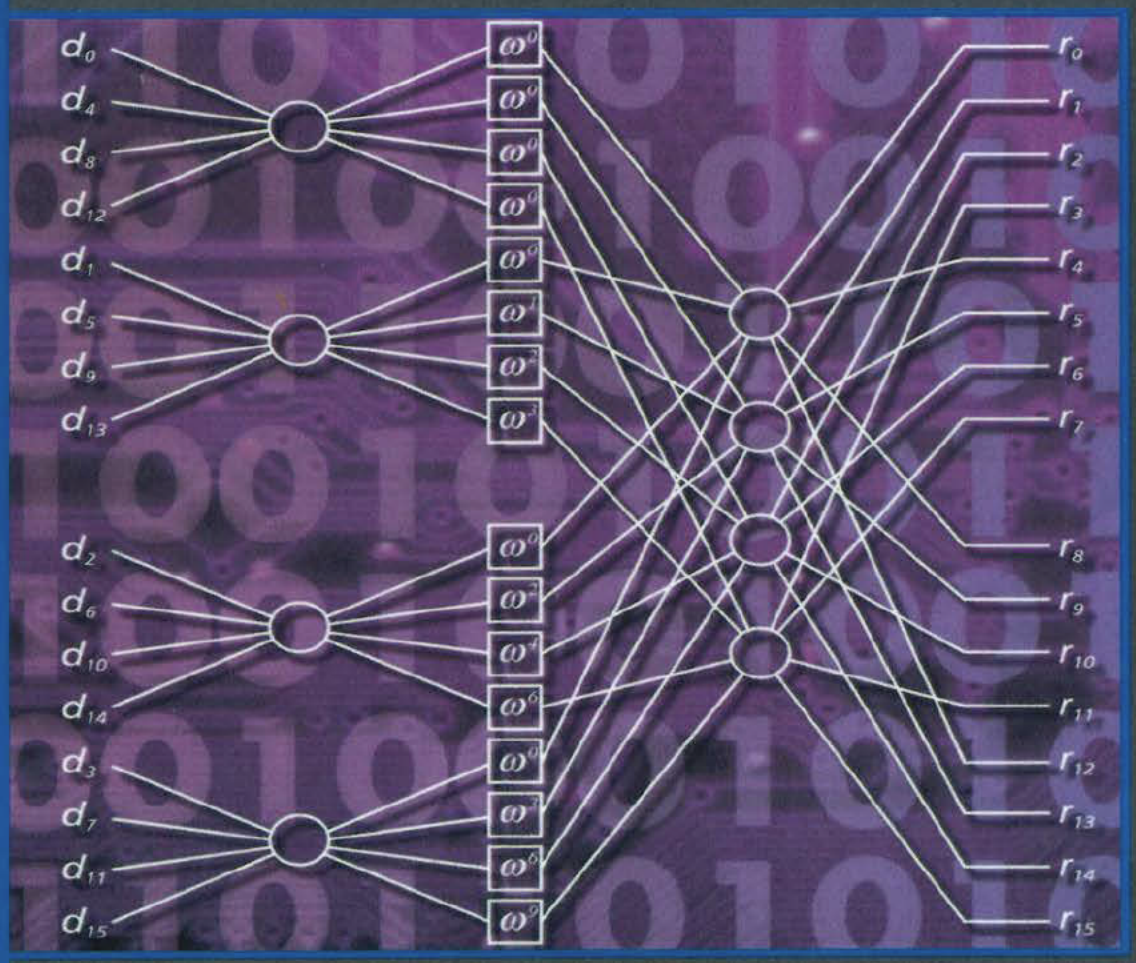
UNIVERSAL PERSONAL COMMUNICATIONS

OFDM FOR WIRELESS MULTIMEDIA COMMUNICATIONS



RICHARD VAN NEE
RAMJEE PRASAD

Artech House Publishers BOSTON • LONDON

OFDM Wireless Multimedia Communications

For a recent listing of titles in the *Artech House Universal Personal Communications Library*, turn to the back of this book.

OFDM Wireless Multimedia Communications

Richard van Nee
Ramjee Prasad



Artech House
Boston • London

ZYXEL-1015
Page 004 (of 277)

Library of Congress Cataloging-in-Publication Data
Nee, Richard van.
OFDM for wireless multimedia communications / Richard van Nee, Ramjee Prasad
p. cm. — (Artech House universal personal communications library)
Includes bibliographical references and index.
ISBN 0-89006-530-6 (alk. paper)
1. Wireless communications systems. 2. Multimedia systems. 3. Multiplexing. I. Prasad, Ramjee. II. Title III. Series

TK5103.2.N44 2000 99-052312
621/3845—dc21 CIP

British Library Cataloguing in Publication Data
Nee, Richard van
OFDM wireless multimedia communications. — (Artech House universal personal communications library)
1.Wireless communications systems 2. Multimedia systems
I. Title II. Prasad, Ramjee
621.3'82

ISBN 0-89006-530-6

Cover design by Igor Vladman

© 2000 Richatd van Nee and Ramjee Prasad

All rights reserved. Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the authors.

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Artech House cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

International Standard Book Number: 0-89006-530-6
Library of Congress Catalog Card Number: 99-052312

10 9 8 7 6 5 4 3 2 1

To my wife Iris, to our son Floris, to our daughters Roselinde and Mirrelijn, and to our newly born baby
—Richard van Nee

To my wife Jyoti, to our daughter Neeli, and to our sons Anand and Rajeev
—Ramjee Prasad

Contents

Preface		xiii
Acknowledgments		xvii
Chapter 1	Introduction	1
1.1	Standardization and Frequency Bands	4
1.2	Multimedia Communications	7
	1.2.1 The Need for High Data Rates	8
	1.2.2 Services and Applications	9
	1.2.3 Antennas and Batteries	9
	1.2.4 Safety Considerations	10
	1.2.5 ATM-Based Wireless (Mobile) Broadband Multimedia Systems	12
1.3	Multipath Propagation	15
	1.3.1 Multipath Channel Models	16
	1.3.2 Delay Spread Values	17
1.4	Time Variation of the Channel	19
1.5	History of OFDM	20
1.6	Preview of the Book	24
References		25
Chapter 2	OFDM Basics	33
2.1	Introduction	33
2.2	Generation of Subcarriers using the IFFT	33
2.3	Guard Time and Cyclic Extension	39
2.4	Windowing	42
2.5	Choice of OFDM Parameters	46
2.6	OFDM Signal Processing	47
2.7	Implementation Complexity of OFDM Versus Single Carrier Modulation	48

ZYXEL-1015
Page 007 (of 277)

References		51
Chapter 3	Coding and Modulation	53
3.1	Introduction	53
3.2	Forward Error Correction Coding	54
	3.2.1 Block Codes	54
	3.2.2 Convolutional Codes	55
	3.2.3 Concatenated Codes	58
3.3	Interleaving	59
3.4	Quadrature Amplitude Modulation	60
3.5	Coded Modulation	62
References		70
Chapter 4	Synchronization	73
4.1	Introduction	73
4.2	Sensitivity to Phase Noise	74
4.3	Sensitivity to Frequency Offset	77
4.4	Sensitivity to Timing Errors	78
4.5	Synchronization using the Cyclic Extension	80
4.6	Synchronization using Special Training Symbols	86
4.7	Optimum Timing in the Presence of Multipath	88
References		92
Chapter 5	Coherent and Differential Detection	95
5.1	Introduction	95
5.2	Coherent Detection	95
	5.2.1 Two Dimensional Channel Estimators	96
	5.2.2 One Dimensional Channel Estimators	103
	5.2.3 Special Training Symbols	104
	5.2.4 Decision Directed Channel Estimation	106
5.3	Differential Detection	107
	5.3.1 Differential Detection in the Time Domain	107

	5.3.2 Differential Detection in the Frequency Domain	112
	5.3.3 Differential Amplitude and Phase Shift Keying	115
References		117
Chapter 6	The Peak Power Problem	119
6.1	Introduction	119
6.2	Distribution of the Peak-to-Average Power Ratio	120
6.3	Clipping and Peak Windowing	123
	6.3.1 Required Backoff with a Non-Ideal Power Amplifier	127
	6.3.2 Coding and Scrambling	130
6.4	Peak Cancellation	131
6.5	PAP Reduction Codes	138
	6.5.1 Generating Complementary Codes	141
	6.5.2 Minimum Distance of Complementary Codes	144
	6.5.3 Maximum Likelihood Decoding of Complementary Codes	145
	6.5.4 Suboptimum Decoding of Complementary Codes	147
	6.5.5 Large Code Lengths	150
6.6	SYMBOL Scrambling	150
References		153
Chapter 7	Basics of CDMA	155
7.1	Introduction	155
7.2	CDMA: Past, Present, and Future	156
7.3	CDMA Concepts	157
	7.3.1 Pure CDMA	161
7.4	Basic DS-CDMA Elements	171
	7.4.1 RAKE Receiver	171
	7.4.2 Power Control	172
	7.4.3 Soft Handover	173
	7.4.4 Interfrequency Handover	175
	7.4.5 Multiuser Detection	175

References 176

Chapter 8 Multi - Carrier CDMA 179
8.1 Introduction 179
8.2 Channel Model 180
8.3 DS-CDMA and MC-CDMA Systems 182
8.3.1 DS-CDMA System 182
8.3.2 MC-CDMA System 185
8.4 MC-CDMA System Design 189
8.5 BEP LOWER Bound 194
8.5.1 DS-CDMA System 194
8.5.2 MC-CDMA System 195
8.5.3 BEP Lower Bound Equivalence 196
8.6 Numerical Results 197
8.6.1 MC-CDMA System Design 197
8.6.2 Down - Link BEP Performance 199
8.6.3 Up - Link BER Performance 203
8.7 Conclusions 206
Appendix 8A 208
References 209

Chapter 9 Orthogonal Frequency Division Multiple Access 213
9.1 Introduction 213
9.2 Frequency Hopping OFDMA 213
9.3 Differences between OFDMA and MC-CDMA 215
9.4 OFDMA System Description 217
9.4.1 Channel Coding 220
9.4.2 Modulation 220
9.4.3 Time and Frequency Synchronization 221
9.4.4 Initial Modulation Timing Synchronization 221
9.4.5 Initial Frequency Offset Synchronization 222

	9.4.6 Synchronization Accuracy	222
	9.4.7 Power Control	223
	9.4.8 Random Frequency Hopping Operation	224
	9.4.9 Dynamic Channel Allocation (Fast DCA)	225
	9.4.10 Dynamic Channel Allocation (Simple DCA)	227
	9.4.11 Capacity of OFDMA	227
9.5	Conclusions	227
References		228
Chapter 10	Applications of OFDM	229
10.1	Introduction	229
10.2	Digital Audio Broadcasting	229
10.3	Terrestrial Digital Video Broadcasting	231
10.4	Magic WAND	233
	10.4.1 Magic WAND Physical Layer	234
	10.4.2 Coding	236
	10.4.3 Simulated Error Probabilities	236
	10.4.4 Effects of Clipping	237
	10.4.5 Magic WAND Medium Access Control Layer	238
10.5	IEEE 802.11, HIPERLAN/2, and MMAC Wireless LAN Standards	241
	10.5.1 OFDM Parameters	243
	10.5.2 Channelization	244
	10.5.3 OFDM Signal Processing	245
	10.5.4 Training	246
	10.5.5 Differences between IEEE 802.11, HIPERLAN/2 and MMAC	249
	10.5.6 Simulation Results	250
References		252
About the Authors		255
Index		257

Preface

सर्वद्वारेषु देहेऽस्मिन् प्रकाश उपजायते ।
ज्ञानं यदा तदा विद्याद् विवृद्धं सत्त्वमित्युत

sarva-dvāreṣu dehe 'smin
prakāśa upajāyate
jñānaṁ yadā tadā vidyād
vivṛddhaṁ sattvam ity uta

The manifestations of the mode of goodness can be experienced when all the gates of the body are illuminated by knowledge

The Bhagavad Gita (14.11)

During the joint supervision of a Master's thesis "The Peak-to-Average Power Ratio of OFDM," of Arnout de Wild from Delft University of Technology, The Netherlands, we realized that there was a shortage of technical information on orthogonal frequency division multiplexing (OFDM) in a single reference. Therefore, we decided to write a comprehensive introduction to OFDM. This is the first book to give a broad treatment to OFDM for mobile multimedia communications. Until now, no such book was available in the market. We have attempted to fill this gap in the literature.

Currently, OFDM is of great interest by the researchers in the Universities and research laboratories all over the world. OFDM has already been accepted for the new wireless local area network standards from IEEE 802.11, High Performance Local Area Network type 2 (HIPERLAN/2) and Mobile Multimedia Access Communication (MMAC) Systems. Also, it is expected to be used for the wireless broadband multimedia communications.

ZYXEL-1015
Page 012 (of 277)

OFDM for Wireless Multimedia Communications is the first book to take a comprehensive look at OFDM, providing the design guidelines one needs to maximize benefits from this important new technology. The book gives engineers a solid base for assessing the performance of wireless OFDM systems. It describes the new OFDM-based wireless LAN standards; examines the basics of direct-sequence and frequency-hopping CDMA, helpful in understanding combinations of OFDM and CDMA. It also looks at applications of OFDM, including digital audio and video broadcasting, and wireless ATM. Loaded with essential figures and equations, it is a must-have for practicing communications engineers, researchers, academics, and students of communications technology.

Chapter 1 presents a general introduction to wireless broadband multimedia communication systems (WBMCS), multipath propagation, and the history of OFDM. A part of this chapter is based on the contributions of Luis Correia from the Technical University of Lisbon, Portugal, Anand Raghawa Prasad from Lucent Technologies, and Hiroshi Harada from the Communications Research Laboratory, Ministry of Posts and Telecommunications, Yokosuka, Japan.

Chapters 2 to 5 deal with the basic knowledge of OFDM including modulation and coding, synchronization, and channel estimation, that every post-graduate student as well as practicing engineers must learn. Chapter 2 contains contributions of Rob Kopmeiners from Lucent Technologies on the FFT design.

Chapter 6 describes the peak-to-average power problem, as well as several solutions to it. It is partly based on the contribution of Arnout de Wild.

Basic principles of CDMA are discussed in Chapter 7 to understand multi carrier CDMA and frequency-hopping OFDMA, which are described in Chapters 8 and 9. Chapter 8 is based on the research contributions from Shinsuke Hara from the University of Osaka, Japan, a postdoctoral student at Delft University of Technology during 1995–96. Chapter 9 is based on a UMTS proposal, with main contributions of Ralf Böhnke from Sony, Germany, David Bhatoolaul and Magnus Sandell from Lucent Technologies, Matthias Wahlquist from Telia Research, Sweden, and Jan-Jaap van de Beek from Lulea University, Sweden.

Chapter 10 was witten from the viewpoint of top technocrats from industries, government departments, and policy-making bodies. It describes several applications of OFDM, with the main focus on wireless ATM in the Magic WAND project, and the new wireless LAN standards for the 5 GHz band from IEEE 802.11, HIPERLAN/2 and MMAC. It is partly based on contributions from Geert Awater from Lucent Technologies, and Masahiro Morikura and Hitoshi Takanashi from NTT in Japan and California, respectively.

We have tried our best to make each chapter quite complete in itself. This book will help generate many new research problems and solutions for future mobile multimedia communications. We cannot claim that this book is errorless. Any remarks to improve the text and correct any errors would be highly appreciated.

Acknowledgments

The material in this book originates from several projects at Lucent Technologies and research activities at Delft University of Technology, The Netherlands. A great deal of OFDM-knowledge was acquired during the Magic WAND project, in which several companies jointly managed to build a wireless ATM network demonstrator based on OFDM. We wish to thank all Magic WAND members who were involved in the design of the OFDM modem, in particular Geert Awater, James Hopper, Rob Kopmeiners, Erik Busking, Han Schmitz, Theo Kleijne, Martin Janssen, Jan Kruys, Urs Bernhard, Urs Lott, Alex Grant, James Aldis, Thomas Mark, Rodolfo Mannpelz, and Ari Vaisanen. Another fruitful source of information was the OFDM-based proposal for UMTS, with main contributions coming from Ralf Böhnke, David Bhatoolaul, Magnus Sandell, and Jan-Jaap van de Beek.

Richard wishes to thank Masahiro Morikura and Hitoshi Takanashi for all their contributions and the pleasant cooperation on the joint OFDM proposal for the IEEE 802.11 wireless LAN standard. All members of IEEE 802.11 and HiperLAN/2 are thanked for numerous contributions, that greatly improved the quality of the final standards and helped to gain more insight in various OFDM techniques.

Geert Awater and Rob Kopmeiners made contributions and corrected numerous mistakes. Neeli Rashmi Prasad helped to prepare the complete manuscript, freeing us from the enormous burden of editorial requirements. Shinsuke Hara from the University of Osaka, Japan, Hiroshi Harada from the Communications Research Laboratory, Ministry of Posts and Telecommunications, Yokosuka, Japan, Luis Correia from the Technical University of Lisbon, Portugal, Anand Raghawa Prasad from Lucent Technologies, and Arnout de Wild from Siemens, The Netherlands, are deeply acknowledged for their valuable contributions.

After the Magic WAND project, we studied several OFDM options for new high-rate wireless LAN products. A circle was completed by the selection of OFDM for the high rate extension of the IEEE 802.11 wireless LAN standard in July 1998. We hope this book will help to gain insight in the principles and design of OFDM-based systems, in particular the new OFDM-based wireless LAN standards.

Richard van Nee
Ramjee Prasad
October 1999

ZYXEL-1015
Page 015 (of 277)

Chapter 1

Introduction

The spectacular growth of video, voice, and data communication over the Internet, and the equally rapid pervasion of mobile telephony, justify great expectations for mobile multimedia. Research and development are taking place all over the world to define the next generation of wireless broadband multimedia communications systems (WBMCS) that may create the "global information village." Figure 1.1 illustrates the basic concept of the global information village, which consists of various components at different scales ranging from global to picocellular size. As we know, the demand for wireless (mobile) communications and Internet/multimedia communications is growing exponentially. Therefore, it is imperative that both wireless and Internet/multimedia should be brought together. Thus, in the near future, wireless Internet Protocol (IP) and wireless asynchronous transfer mode (ATM) will play an important role in the development of WBMCS.

While present communications systems are primarily designed for one specific application, such as speech on a mobile telephone or high-rate data in a wireless local area network (LAN), the next generation of WBMCS will integrate various functions and applications. WBMCS is expected to provide its users with customer premises services that have information rates exceeding 2 Mbps. Supporting such large data rates with sufficient robustness to radio channel impairments, requires careful choosing of modulation technique. The most suitable modulation choice seems to be orthogonal frequency division multiplexing (OFDM). Before going into the details of OFDM, however, first we give some background information on the systems that will be using it.

The theme of WBMCS is to provide its users a means of radio access to broadband services supported on customer premises networks or offered directly by public fixed networks. WBMCS will provide a mobile/movable wireless extension to

ZYXEL-1015
Page 016 (of 277)

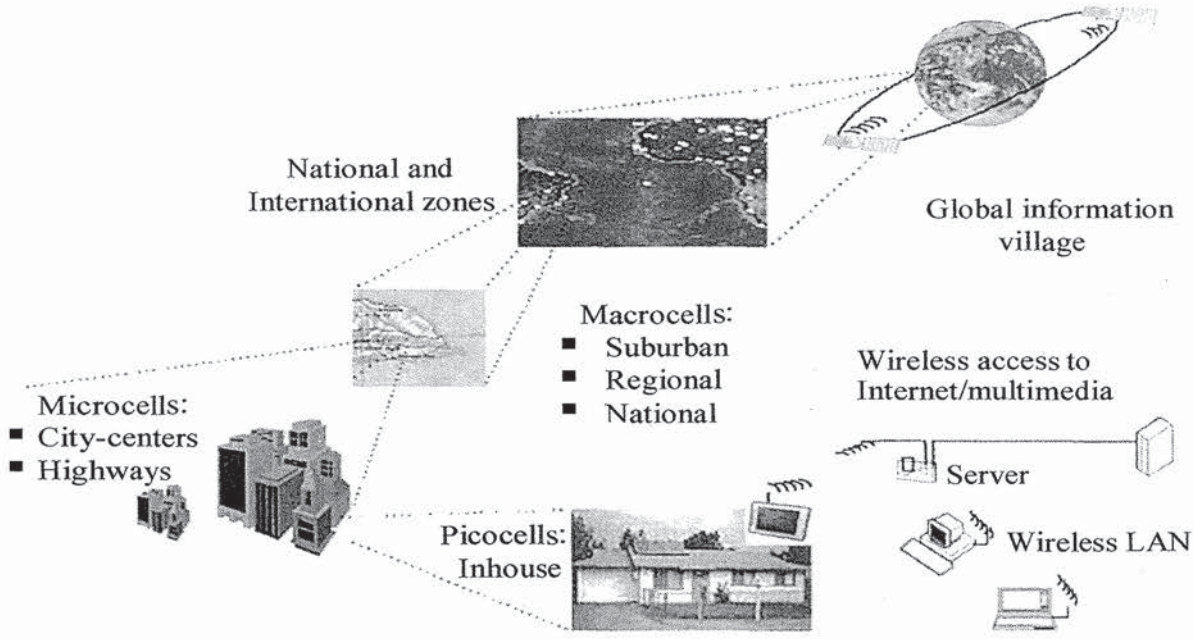


Figure 1.1 Global information village.

WBMCS is under investigation in North America, Europe, and Japan in the microwave and millimeter-wave bands to accommodate the necessary bandwidth. The research in the field of WBMCS has drawn much attention because of the increasing role of multimedia and computer applications in communications. There is a major thrust in three research areas: (1) microwave and millimeter-wave bands for fixed access in outdoor, public commercial networks, (2) evolution of WLAN for inbuilding systems, and (3) use of LAN technology outdoors rather than indoors. In short, WBMCS will provide novel multimedia and video mobile communications services, also related to wireless customer premises network (WCPN) and wireless local loop (WLL).

To implement the wireless broadband communication systems, the following challenges must be considered:

- Frequency allocation and selection;
- Channel characterization;
- Application and environment recognition, including health hazard issues;
- Technology development;
- Air interface multiple access techniques;
- Protocols and networks; and
- Systems development with efficient modulation, coding, and smart antenna techniques.

- Protocols and networks; and
- Systems development with efficient modulation, coding, and smart antenna techniques.

A significant number of research and development (R&D) projects are set up in the area of WBMCS. Within the European Advanced Communication Technologies and Services (ACTS) program are four European Union-funded R&D projects, namely Magic Wand (Wireless ATM Network Demonstrator), ATM Wireless Access Communication System (AWACS), System for Advanced Mobile Broadband Applications (SAMBA), and wireless broadband CPN/LAN for professional and residential multimedia applications (MEDIAN) [1–17]. Table 1.1 summarizes the European projects [11].

In the United States, seamless wireless network (SWAN) and broadband adaptive homing ATM architecture (BAHAMA), as well as two major projects in Bell Laboratories and a wireless ATM network (WATMnet), are being developed in the computer and communications (C&C) research laboratories of Nippon Electric Company (NEC) in the United States [2–6].

In Japan, Communication Research Laboratory (CRL) is working on several R&D projects, such as a broadband mobile communication system in the super high frequency (SHF) band (from 3 to 10 GHz) with a channel bit rate up to 10 Mbps and an indoor high speed WLAN in SHF band with a target bit rate of up to 155 Mbps [12].

In the Netherlands, Delft University of Technology has been busy with a multi-disciplinary research project, "Mobile Multimedia Communication (MMC)," since April 1996. The team consists of experts from the telecommunications and traffic control and information theory groups of the department of Electrical Engineering, the Product Ergonomics group of the department of Industrial Design Engineering, and the Organizational Psychology group of the department of Technology and Society.

The MMC has the following objectives to achieve at 60 GHz:

- Wireless access of 155 Mbps using OFDM;
- Both indoor and outdoor use;
- Less complex, inexpensive mobile stations by moving most functionality to the access points;
- Modified OFDM; and
- Constant bit rate (CBR), variable bit rate (VBR), and available bit rate (ABR) services.

Table 1.1
Summary of European ACTS Projects.

ACTS project / Parameter	WAND	AWACS	SAMBA	MEDIAN
Frequency	5 GHz	19 GHz	40 GHz	61.2 GHz
Data rate	20 Mbps	70 Mbps	34 Mbps	155 Mbps
Modulation	OFDM, 16 subcarriers with 8-PSK (phase shift keying)	offset quadrature PSK (OQPSK)	OQPSK	OFDM, 512 carriers, differential QPSK (DQPSK)
Cell radius	20–50m (omnidirectional antennas)	50–100m (directional antennas, line-of-sight only)	10-50m (directional antennas at access point)	10m (directional antennas)
Radio access	time division multiple access/ time division duplex (TDMA/TDD)	TDMA/TDD	time division multiple access /frequency division duplex (TDMA/FDD)	TDMA/TDD

1.1 STANDARDIZATION AND FREQUENCY BANDS

There are three main forums for the standardization of wireless broadband communication systems; namely, IEEE 802.11 [18], European Telecommunication Standards Institute Broadband Radio Access Networks (ETSI BRAN) [19], and Multimedia Mobile Access Communications (MMAC) [20]. IEEE 802.11 made the first WLAN standard for the 2.4-GHz Industrial, Scientific, and Medical band (ISM). It specifies the medium access control and three different physical layers—direct-sequence spread spectrum, frequency hopping, and infrared—which give a data rate of 2 Mbps. Products based on this standard became available in 1998. Figure 1.2 shows an example of an IEEE 802.11 modem in a PCMCIA card. Following the initial 1- and 2-Mbps standard, IEEE 802.11 developed two new physical layer standards. One delivers data rates of up to 11 Mbps in the 2.4-GHz band, using complementary code keying [21,22]. Products based on this standard—with the old 1 and 2 Mbps as fallback rates—are available since mid 1999. An industry alliance called the Wireless Ethernet Compatibility Alliance (WECA) has been established to promote the high rate IEEE

available since mid 1999. An industry alliance called the Wireless Ethernet Compatibility Alliance (WECA) has been established to promote the high rate IEEE 802.11 technology and to certify interoperability of products from different vendors [23]. The second IEEE 802.11 standard extension targets a range of data rates from 6 up to 54 Mbps using OFDM in the 5-GHz band [24]. The OFDM standard was developed jointly with ETSI BRAN and MMAC, making OFDM effectively a worldwide standard for the 5-GHz band.

Table 1.2 lists the main characteristics of the IEEE 802.11 and the ETSI High Performance Local Area Network type 2 (HIPERLAN/2) standards. More details about these standards can be found in Chapter 10.



Figure 1.2 IEEE 802.11 modem for the 2.4-GHz band (WaveLAN™ from Lucent Technologies [25]).

Figure 1.3 shows that 2-, 5- and 60-GHz are the commercially important frequency bands because of geographically wide spectrum allocations in Europe, the United States (U.S.), and Japan for the wireless broadband multimedia communications networks. The 2.4-GHz band is an ISM band, which can be used for many types of transmission systems as long as they obey certain power, spectral density, and spreading gain requirements. The 5-GHz band is designated specifically for WBMCS. In Europe, only HIPERLAN devices are currently allowed in this band. HIPERLAN actually consists of a family of standards, one of which is an OFDM-based standard that is very similar to the IEEE 802.11 5-GHz standard. In Japan, MMAC supports both the IEEE 802.11 and the HIPERLAN standards. Notice that Japan only has 100 MHz available in the 5-GHz band, while the United States and Europe provide 300 and 455 MHz, repectively. In Europe, extra spectrum for HIPERLAN is available in the 17-GHz band, while Japan has allocated spectrum from 10- to 16-GHz to mobile broadband systems (MBS). An analysis of the propagation aspects at the bands foreseen for WBMCS microwaves, millimeterwaves, and infrared is presented in [26–31].

Table 1.2
Comparison of IEEE and HIPERLAN standards.

Parameter	IEEE 802.11 2 GHz	IEEE 802.11 5 GHz	HIPERLAN/2
Configurations	Centralized system with access points connected to wired network, or peer-to-peer networking	Centralized system with access points connected to wired network, or peer-to-peer networking	Centralized system with access points connected to wired network
Range	Up to 60m at 11 Mbps and up to 100m at 2 Mbps with omnidirectional antennas	Up to 30m at 24 Mbps and up to 60m at 6 Mbps with omnidirectional antennas	Up to 30m at 24 Mbps and up to 60m at 6 Mbps with omnidirectional antennas
Channel access	CSMA/CA, variable size data packets (up to 8192 bytes)	CSMA/CA, variable size data packets (up to 8192 bytes)	Reservation based access, scheduled by access point. Contention slots for making slot reservations
Frequency bands	2.4-2.4835 GHz	5.150–5.350 GHz; 5.725–5.825 GHz	5.150–5.350 GHz; 5.470–5.725 GHz;
Duplexing	TDD	TDD	TDD
Data rate	1, 2 Mbps (BPSK/QPSK) 5.5, 11 Mbps (CCK)	6, 9 Mbps (BPSK) 12, 18 Mbps (QPSK) 24, 36 Mbps (16-QAM) 54 Mbps (64-QAM)	6, 9 Mbps (BPSK) 12, 18 Mbps (QPSK) 24, 36 Mbps (16-QAM) 54 Mbps (64-QAM)

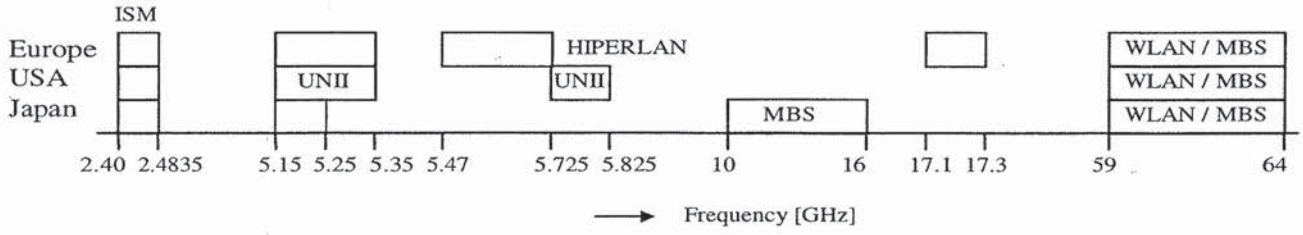


Figure 1.3 Frequency band for wireless broadband communications.

1.2 MULTIMEDIA COMMUNICATIONS

Multimedia and computer communications are playing an increasing role in today's society, creating new challenges to those working in the development of telecommunications systems. Besides that, telecommunications is increasingly relying upon wireless links. Thus, the pressure for wireless systems to cope with increasing data rates is enormous, and WBMCSs with data rates higher than 2 Mbps are emerging rapidly, even if at this moment applications for very high transmission rates do not exist.

Several WBMCSs are being considered for different users with different needs. They may accommodate data rates ranging between 2 and 155 Mbps; terminals can be mobile (moving while communicating) or portable (static while communicating); moving speeds can be as high as that of a fast train; users may or may not be allowed to use more than one channel if their application requires so; the system bandwidth may be fixed, or dynamically allocated according to the user's needs; communication between terminals may be direct or must go through a base station; possible ATM technology use; and so on. Many other cases can be listed as making the difference between various perspectives of a WBMCS, but two major approaches are emerging: WLANs directed to communication among computers, from which IEEE 802.11 [16, 18] and HIPERLAN [19, 20] are examples, with MBS [17] intended as a cellular system providing full mobility to B-ISDN users.

The different requirements imposed by the various approaches to WBMCSs have consequences on system design and development. The tradeoffs between maximum flexibility on one hand and complexity and cost on the other are always difficult to decide, as they have an impact not only on the deployment of a system, but also on its future evolution and market acceptance. GSM is a good example of a system foreseen to accommodate additional services and capacities to those initially offered, and the fact that operators are already implementing phase 2+ is proof of that.

This means that many decisions must be made on the several WBMCSs that will appear on the market. For IEEE 802.11, for example, those decisions have already been made, as the system will be commercialized in the very near future, but for other systems there are still many undecided aspects. Of course this depends on what are the applications intended to be supported by the systems, and whether these applications are targeted to the mass market or only to some niches. The former (from which mobile telephones are a good example) will certainly include WLANs, because the expansion of personal computers will dictate this application as a great success in WBMCSs; the latter will possibly have television broadcasters among their users (to establish links between HDTV cameras and the central control room).

Not only are market aspects at stake in the development and deployment of WBMCSs, but many technical challenges are posed as well. The transmission of such high data rates over radio in a mobile environment creates additional difficulties,

compared with that of existing systems; these difficulties are augmented by the fact that frequencies higher than UHF are needed to support the corresponding bandwidths, thus pushing mobile technology challenges (size and weight among other things) to frequencies where these aspects were not much considered until now. However, additional challenges are posed to those involved in WBMCSs development: in today's world, where consumers are in the habit of using a communications system that is available in different places (e.g., GSM roaming capability, because users can make and receive telephone calls in an increasing number of countries worldwide), or being able to exchange information among different systems (e.g., the exchange of files between different computer applications and systems), for future use it does not make sense to consider systems that offer a high data rate but do not support these capabilities to some extent.

1.2.1 Need for High Data Rates

Data rate is really what broadband is about. For example, the new IEEE and HIPERLAN standards specify bit rates of up to 54 Mbps, although 24 Mbps will be the typical rate used in most applications. Such high data rates impose large bandwidths, thus pushing carrier frequencies for values higher than the UHF band: HIPERLAN has frequencies allocated in the 5- and 17-GHz bands; MBS will occupy the 40- and 60-GHz bands; and even the infrared band is being considered for broadband WLANs. Many people argue whether there is a need for such high-capacity systems, however, bearing in mind all the compression algorithms developed and the type of applications that do require tens of megabits per second. We can examine this issue from another perspective.

The need for high-capacity systems is recognized by the "Visionary Group" [32], put together by the European Commission, to give a perspective of what should be the hot topics in the telecommunications for research in the next European programs (following R&D in Advanced Communications Technologies for Europe (RACE) and ACTS). In this visionary perspective, to meet the needs of society in the years to come as far as communications is concerned, capacity is one of the major issues to be developed because of the foreseen increase in demand for new services (especially those based on multimedia). Along with this, mobility will impose new challenges to the development of new personal and mobile communications systems.

We can conclude the following: even if at a certain point it may look academic to develop a system for a capacity much higher than what seems reasonable (in the sense that there are no applications requiring such high capacity), it is worthwhile to do so, as almost certainly in the future (which may be not very far off) applications will need those capacities and even more. The story of fiber optics is an example.

1.2.2 Services and Applications

The system concept of a WLAN such as IEEE 802.11 and of a mobile broadband cellular system such as MBS is totally different: each is directed to services and applications that differ in many aspects. A comparison of several systems, based on two of the key features (mobility and data rate), is shown in Figure 1.4 [33], where it is clear that no competition exists between the different approaches.

The applications and services of the various systems are also different. IEEE 802.11 is mainly intended for communications between computers (thus being an extension of wired LANS); nevertheless, it can support real-time voice and image signals, and users are allowed some mobility and can have access to public networks.

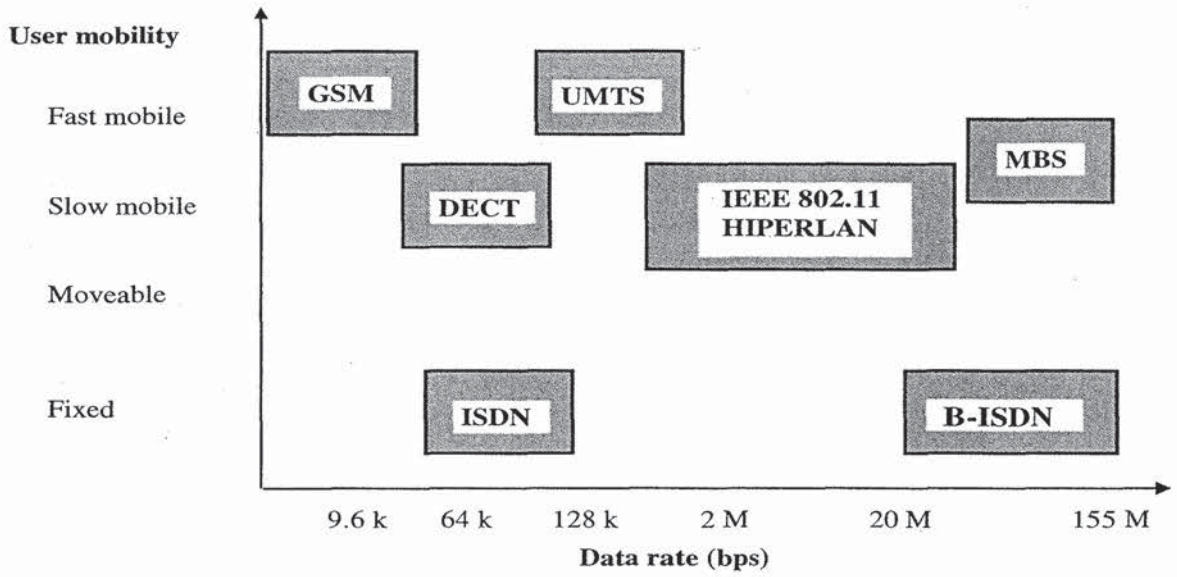


Figure 1.4 Comparison of mobility and data rates for several systems.

1.2.3 Antennas and Batteries

Antennas and batteries play a key role in wireless systems. With the advent of microelectronics and signal processing, antennas and batteries tend to impose the size and weight of mobile terminals. Of course, the higher one goes in frequency, the less developed the technology, and many problems are still found in size and weight at the millimeter-wave band. Power consumption is one example. Though these are likely to be solved in the near future. The number of hours battery-powered equipment can work or operate on stand by, and the percentage of its weight corresponding to the battery, is

not a specific problem of WBMCS. Laptop computers and cellular telephones are the most common terminals relying on batteries these days. This technology continues to demand huge R&D, to extend working time and reduce weight. Although a 100g mobile telephone battery corresponding to several hours of continuous work and a few days on standby is already on the market, users still want more. Mobile multimedia terminals, certainly those to be used in some applications of WBMCSs, will be an extension of the current cellular telephones. Therefore, we foresee that the current problems associated with batteries will be transposed on WBMCS terminals; the same applies to laptops.

Antennas (size, type, technology, etc.) are not a specific problem of WBMCSs as well, but again they are very much related to the type of systems that will be made available for users. It does not make sense to impose restrictions (e.g., pointing in a certain direction or avoiding someone to pass in between) on the type of mobility a mobile terminal can have. Even for portable terminals (e.g., computers), these restrictions make no sense. Hence, there are only two options as far as antennas are concerned: either an omnidirectional antenna (dipole type) or an adaptive array antenna is used. Either way, patch antennas seem a very promising solution for general use in WBMCS. For the frequency bands we consider, isolated patches or adaptive arrays (with many elements) can be made with a small size (e.g., a credit card), thus enabling the terminal not to be limited in size by the antenna system.

The role of antenna radiation patterns is not negligible when discussing system performance, nor is their influence on parameters associated with wave propagation. Although the tradeoff between an omnidirectional and a narrowbeam antenna is not particular of WBMCS, it assumes particular importance at microwave and millimeter waves because of the characteristics of wave propagation at these bands. Using an omnidirectional antenna means a lower gain, but also the possibility of receiving signals from various directions, without the requirement for knowing where the base station is, and allowing the received rays coming from reflections on the propagation scenario. On the other hand, the use of a very directive antenna provides a higher gain, but it must be pointed at the base station and does not receive reflected waves coming from directions very different from the one to which it is pointed. The need for pointing it can be very discouraging, if not a drawback, when a line of sight does not exist. On the other hand, omnidirectional antennas lead to high values of delay spread, but they may ensure that the link still exists, relying on reflections if the line of sight is lost.

1.2.4 Safety Considerations

Until a few years ago, the analysis of possible harmful effects of electromagnetic radiation on people was devoted mainly to power lines and radars, because of the huge power levels involved in those systems. Even when mobile telephone systems appeared, there was no major concern, as the antennas were installed on the roofs of cars. With

the development of personal communication systems, in which users carry mobile telephones inside their coat pockets, and the antenna radiates a few centimeters from the head, safety issues gained great importance and a new perspective. Much research in the literature focuses not only on the absorption of power inside the head, but also on the influence of the head on the antenna's radiation pattern and input impedance. However, these works have addressed only the frequency bands used in today's systems; that is, up to 2-GHz (mainly on the 900- and 1,800-MHz bands), and only very few references are made to systems working at higher frequencies, as it is in the case of WBMCSs.

The problems associated with infrared are different from those posed by microwaves and millimeter waves. Eye safety, rather than power absorption inside the head, is the issue here, because the eye acts as a filter to the electromagnetic radiation, allowing only light and near-frequency radiation to enter into it, and the amount of power absorption inside the human body is negligible. Exposure of the eye to high levels of infrared radiation may cause cataractlike diseases, and the maximum allowed transmitter power seems to limit the range to a few meters [31]. If this is the case, safety restrictions will pose severe limitations to the use of infrared in WBMCSs, as far as general applications are concerned. The question in this case is not that there are always problems during system operation (e.g., mobile telephones), but the damage that may be caused if someone looks at the transmitter during operation.

Microwaves and millimeter waves have no special effect on eyes, other than power absorption. In WLANs, antennas do not radiate very near (1 or 2 cm) to the user as in the mobile telephone case, thus enabling power limitations to be less restrictive; also if mobile multimedia terminals are used as they are in PDAs. But if terminals are used in the same form as mobile telephones, then maximum transmitter powers have to be established, similar to those for the current personal communication systems. The standards for safety levels have already been set in the United States and Europe, as the ones used for UHF extend up to 300 MHz (IEEE/ANSI and CENELEC recommendations are the references). Thus, it is left to researchers in this area to extend their work to higher frequencies, by evaluating SAR (the amount of power dissipated per unit of mass) levels inside the head (or other parts of the human body very near the radiating system), from which maximum transmitter powers will be established. This may not be as straightforward as it seems, however, because the calculation of SAR is usually done by solving integral or differential equations using numerical methods (method of moments or finite difference), which require models of the head made of small elements (e.g., cubes) with dimensions of the order of a tenth of the wavelength. This already requires powerful computer resources (in memory and CPU time) for frequencies in the high UHF band, and may limit the possibility of analyzing frequencies much higher than UHF. On the other hand, the higher the frequency, the smaller the penetration of radio waves into the human body, hence making it possible to have models of only some centimeters deep. This is an area for further research.

1.2.5 ATM-Based Wireless (Mobile) Broadband Multimedia Systems

Recent study of MBS, ATM, and ATM-oriented MBS has drawn the attention of several researchers [34–44]. With enormous complexity of managing and operating the many different types of networks now in use, the door is open for finding a common platform—a network on which all established services can be supported and which will allow new services to be introduced without needing new networks on which to run them. The answer seems to be ATM. This is the technology being defined and standardized for B-ISDN. Thus, ATM, when adequately modified, is also an answer for the future mobile wireless broadband multimedia systems.

ATM is a packet-oriented transmission scheme. The transmission path of the packets of constant length, the so-called ATM cells, is established during connection setup between the two endpoints by assigning a virtual channel. At this time, the necessary resources are provided and the logical channels are assigned. All packets of a virtual channel are carried over the same path. The transmission capacity of the virtual channel is characterized by the parameters, mean bit rate and peak bit rate during connection setup. ATM cells are generated according to the need of the data source. Thus, ATM is a very good method to meet the dynamic requirements of connections with variable data rates.

MBS is the interface between the fixed ATM net at the base station side and mobile ATM net at the mobile station side. Normally, the ATM net at the mobile station side only consists of one end system. For every end station, it is possible to operate several virtual channels with different data rates at the same time.

A conceptual view of the ATM-type broadband communications network is shown in Figure 1.5. The most important benefit of ATM is its flexibility; it is used for the new high bit rate services, which are either VBR or burst traffic. Several factors that tend to favor the use of ATM cell transport in MBS are as follows:

- Flexible bandwidth allocation and service-type selection for a range of applications;
- Efficient multiplexing of traffic from bursty data and multimedia sources;
- End-to-end provisioning of broadband services over wired and wireless networks;
- Suitability of available ATM switching for intercell switching;
- Improved service reliability with packet-switching techniques; and
- Ease of interfacing with wired B-ISDN systems.

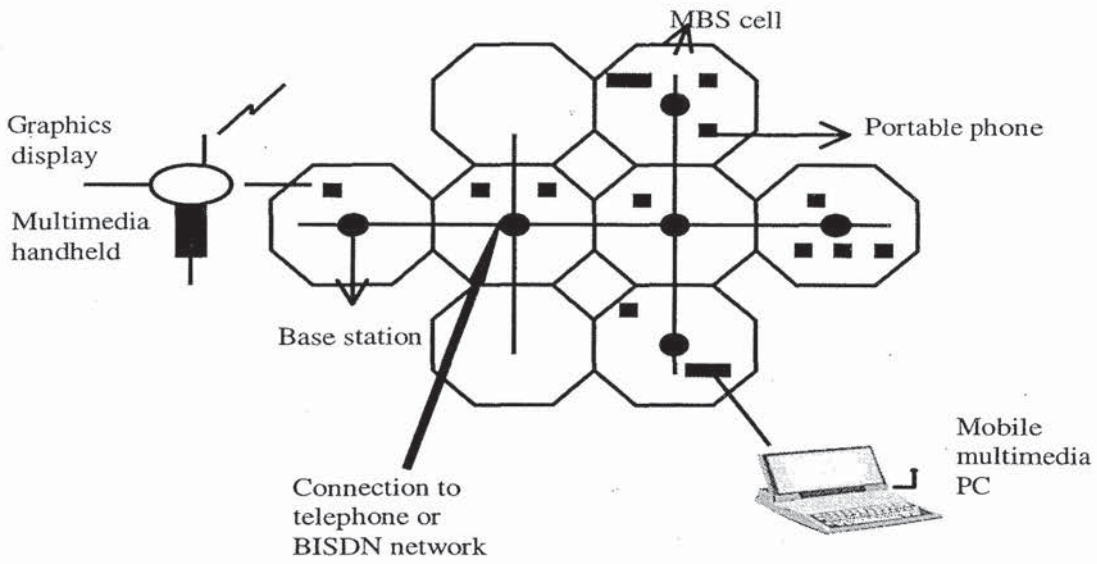


Figure 1.5 Conceptual view of MBS.

Taking the above points into consideration, adoption of ATM-compatible, fixed-length, cell-relay format for MBS is recommended. A possible ATM-compatible MBS approach is shown in Figure 1.6. With this approach, the 48-byte ATM cell payload becomes the basic unit of data within the MBS network. Within MBS, specific protocol layers (e.g., data link and medium-access control layer) are added to the ATM payload as required, and replaced by ATM headers before entering the fixed network [41]. The use of ATM switching for intercell traffic also avoids the crucial problem of developing a new backbone network with sufficient throughput to support intercommunication among large numbers of small cells. ATM multiplexers are used to combine traffic from several base stations into a single ATM port.

For a seamless internetworking mechanism with the wired broadband network, it is vital to have the MBS protocol layering harmonized with the ATM stack. Figure 1.7 shows a protocol reference model. In this approach, new wireless channel-specific physical, medium-access control, and data link layers are added below the ATM network layer. This means that regular network layer and control services such as call setup, virtual channel identifier/virtual path identifier (VCI/VPI) addressing, cell prioritization, and flow-control indication will continue to be used by mobile services. The baseline ATM network and signaling protocol are specified to particular mobility-related functions, such as address registration (roaming), broadcasting, handoff, and so forth.

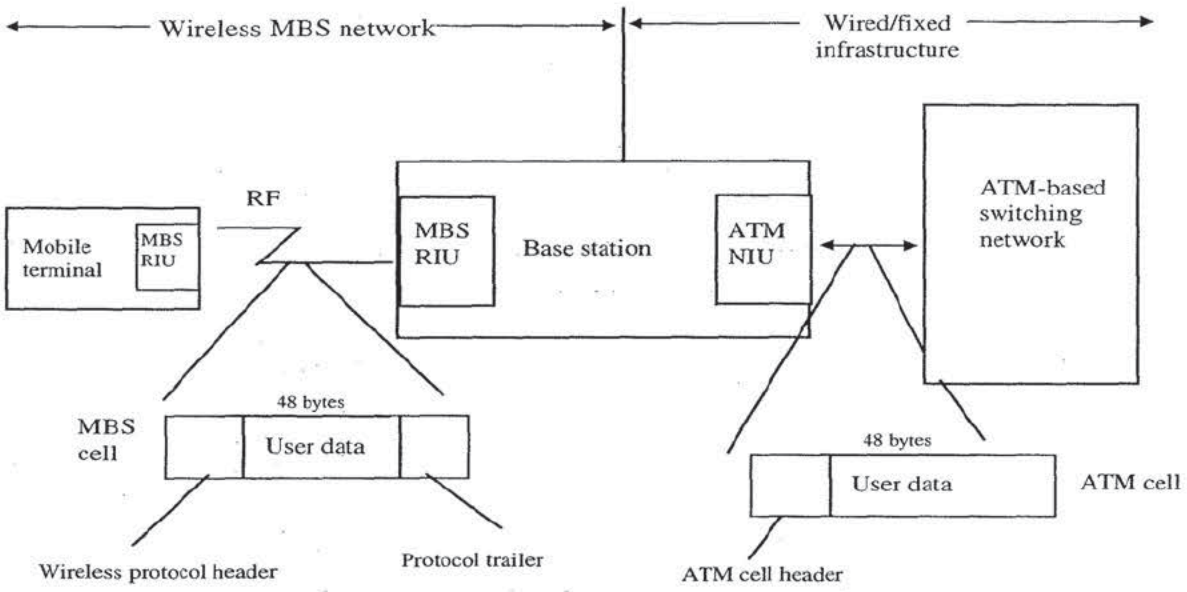


Figure 1.6 ATM-compatible MBS approach. Terms: radio interface unit (RIU); network interface unit (NIU).

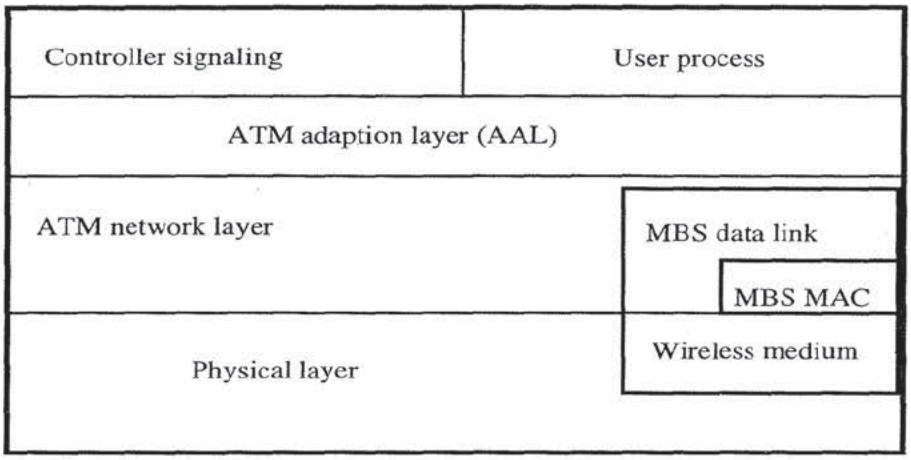


Figure 1.7 Relation of wireless network protocol layers.

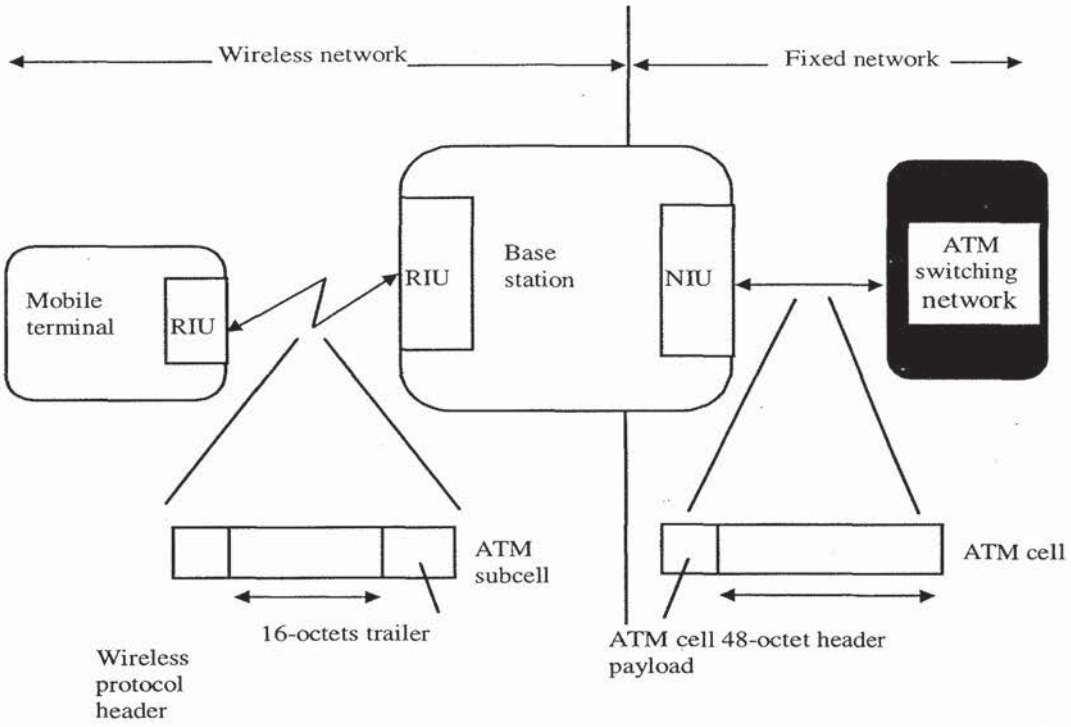


Figure 1.8 Conceptual view of an ATM-compatible air interface.

The payload of an ATM cell consists of 48 octets, meaning that without any segmentation, packet lengths of 384 bits would have to be transmitted over the mobile radio channel. Higher throughput levels over a fading mobile radio channel are achieved when using a packet of smaller lengths. So a suitable integer submultiple of a cell is chosen as the basic unit of data over the wireless radio medium. Based on the system parameters, a submultiple cell size of 16 octets or 128 bits is used as an appropriate value.

The air interface is terminated directly at the base station, as illustrated in Figure 1.8. Also shown is the segmentation of an ATM cell into its three subcells and the addition of an extra wireless ATM header.

1.3 MULTIPATH PROPAGATION

One of the basic reasons to use OFDM is the efficient way it can handle multipath propagation. To understand the effects of multipath fading, this section contains an overview of the most relevant parameters and models.

1.3.1 Multipath Channel Models

Fundamental work on modeling of the indoor radio channel is published in [45, 46]. One of the results of this work, which is also supported by many other measurements reported in the literature, is that the average received multipath power is an exponentially decaying function of the excess delay. Further, the amplitudes of individual multipath components are Rayleigh distributed. This observation has led to simplified channel models as used in [47, 48]. These models assume a fixed number of paths with equidistant delays. The path amplitudes are independent Rayleigh variables, while the path phases are uniformly distributed. Figure 1.9 shows an example of an average and an instantaneous power delay profile that were generated using this approach.

Compared to the more extensive models in [45, 46], the simplified model of [47, 48] may give somewhat optimistic results because the number of multipath components is fixed to the maximum possible amount. In the models of [45, 46], the number of paths is random. Paths arrive in clusters with Poisson distributed arrival times. Within a cluster, the path amplitudes are independent Rayleigh variables. The average power delay profile, averaged over a large number of channels, is an exponentially decaying function, just as for the models in [47, 48]. Thus, the only difference between the models is that the instantaneous power delay profiles have a slightly different shape, and channels generated by the method of [47, 48] generally show more multipath components than channels generated according to [45, 46]. This may give a somewhat optimistic diversity effect, because diversity is proportional to the number of paths. It probably does not make a difference when the models are used to determine the delay spread tolerance of a particular transmission system, however, because that is not depending on the number of paths, but rather on the amount of power in paths exceeding a certain excess delay. So, the channel models of [47, 48] seem a good basis to determine the delay spread tolerance of different modulation schemes. An additional advantage is that the simulation complexity is much lower than that of the models of [45, 46].

One of the key parameters in the design of a transmission system is the maximum delay spread value that it has to tolerate. For an idea of what typical delay spread figures are, the next section presents some measurement results obtained from the literature [49- 63].

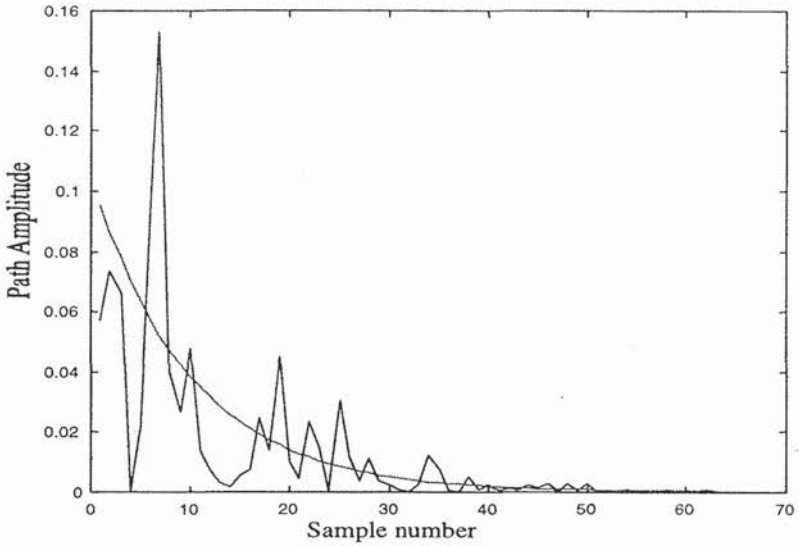


Figure 1.9 Example of generated average and instantaneous power delay profiles for a delay spread of 10 sampling intervals.

1.3.2 Delay Spread Values

Tables 1.3 to 1.5 summarize some delay spread results obtained from literature for frequencies from 800 MHz to 6 GHz. Two delay spread values are given; the *median* delay spread is the 50% value, meaning that 50% of all channels has a delay spread that is lower than the median value. Clearly, the median value is not so interesting for designing a wireless link, because there you want to guarantee that the link works for at least 90% or 99% of all channels. Therefore, the second column gives the measured *maximum* delay spread values. The reason to use maximum delay spread instead of a 90% or 99% value is that many papers only mention the maximum value. From the papers that do present cumulative distribution functions of their measured delay spreads, we can deduce that the 99% value is only a few percent smaller than the maximum measured delay spread.

Table 1.3
Measured delay spreads in frequency range of 800 MHz to 1.5 GHz.

Median delay spread [ns]	Maximum delay spread [ns]	Reference	Remarks
25	50	[46]	Office building
30	56	[58]	Office building
27	43	[59]	Office building
11	58	[60]	Office building
35 40 80 120	80 90 120 180	[61]	Office building Shopping center Airport Factory
50 120	129 300	[62]	Warehouse Factory

Table 1.4
Measured delay spreads in frequency range of 1.8 to 2.4 GHz

Median delay spread [ns]	Maximum delay spread [ns]	Reference	Remarks
40	120	[49]	Large building (New York stock exchange)
40	95	[50]	Office building
40	150	[51]	Office building
60 106	200 270	[54]	Shopping center Laboratory
19	30	[55]	Office building: single room only
20 30 105	65 75 170	[56]	Office building Cafeteria Shopping center
30	56	[58]	Office building
25	30	[63]	Office building: single room only

Table 1.5
Measured delay spreads in frequency range of 4 to 6 GHz

Median delay spread [ns]	Maximum delay spread [ns]	Reference	Remarks
40	120	[49]	Large building (e.g., Stock Exchange)
50 35 10	60 55 35	[52]	Office building Meeting room (5m x 5m) with metal walls Single room with stone walls
40	130	[51]	Office building
40 65 25	120 125 65	[53]	Indoor sports arena Factory Office building
20	30	[63]	Office building: single room only

Interesting results that can be derived from the reported measurements are:

- Measurements done simultaneously on different frequencies show that there is no significant difference in delay spread in the frequency range of 800 MHz to 6 GHz.
- The delay spread is related to the building size; largest delay spreads (up to 270 ns) were measured in large buildings like shopping centers and factories. The reason for this phenomenon is that reflections can have larger delays if the distances between transmitter and reflectors (walls or other objects) are larger.
- For most office buildings, the maximum delay spread is in the range of 40 to 70 ns. Smaller delay spreads around 30 ns occur when both transmitter and receiver are within the same room. Delay spreads of 100 ns or more are seen only in office buildings with large rooms that are some tens of meters in diameter.
- Even small rooms (5m by 5m) can give significant delay spreads around 50 ns when there are metal walls [52]. The reason for this is that reflections from a metal wall only experience a small attenuation, so multiple reflections with significant delays still have considerable power, which increases the delay spread.

1.4 TIME VARIATION OF THE CHANNEL

To classify the time characteristics of the channel [1], the coherence time and Doppler spread are important parameters. The coherence time is the duration over which the channel characteristics do not change significantly. The time variations of the channel are evidenced as a Doppler spread in the frequency domain, which is determined as the width of the spectrum when a single sinusoid (constant envelope) is transmitted. Both

the time correlation function $\varphi_c(\Delta t)$ and the Doppler power spectrum $S_c(f)$ can be related to each other by applying the Fourier transform.

The range of values of the frequency f over which $S_c(f)$ is essentially nonzero is called the Doppler spread B_d of the channel. Because $S_c(f)$ is related to $\varphi_c(\Delta t)$ by the Fourier transform, the reciprocal of B_d is a measure of the coherence time $(\Delta t)_c$ of the channel; that is, $1/B_d$.

The coherence time $(\Delta t)_c$ is a measure of the width of the time correlation function. Clearly, a slow-changing channel has a large coherence time, or equivalently, a small Doppler spread. The rapidity of the fading can now be determined either from the correlation function $\varphi_c(\Delta t)$ or from the Doppler power spectrum $S_c(f)$. This implies that either the channel parameters $(\Delta t)_c$ or B_d can be used to characterize the rapidity of the fading. If the bit time T_b is large compared with the coherence time, then the channel is subject to fast fading. When selecting a bit duration that is smaller than the coherence time of the channel, the channel attenuation and phase shift are essentially fixed for the duration of at least one signaling interval. In this case, the channel is slowly fading or quasistatic. Like coherence bandwidth, there is no exact relationship between coherence time and Doppler spread. If the coherence time $(\Delta t)_c$ is defined as the time over which the time correlation function is above 0.5, $(\Delta t)_c$ is approximately given by $9/(16\pi B_d)$ [79, 80].

Doppler spread is caused by the differences in Doppler shifts of different components of the received signal, if either the transmitter or receiver is in motion. The frequency shift is related to the spatial angle between the direction of arrival of that component and the direction of vehicular motion. If a vehicle is moving at a constant speed V along the X axis, the Doppler shift f_m of the mth plane-wave component is given by $f_m = (V/\lambda)\cos(\alpha_m)$ [81], where α_m is the arrival angle of the mth plane-wave component relative to the direction of movement of the vehicle. It is worth noting here that waves arriving from ahead of the vehicle experience a positive Doppler shift, while those arriving from behind the vehicle have a negative shift. The maximum Doppler shift occurs at $\alpha_m = 0$, assuming $T_b >> (\Delta t)_c$ for fast-fading channel and $T_b << (\Delta t)_c$ for slow-fading channel.

Indoor measurements [82] show that in any fixed location, temporal variations in the received signal envelope caused by movement of personnel and machinery are slow, having a maximum Doppler spread of about 6 Hz.

1.5 HISTORY OF OFDM

OFDM is a special case of multicarrier transmission, where a single datastream is transmitted over a number of lower rate subcarriers. It is worth mentioning here that OFDM can be seen as either a modulation technique or a multiplexing technique. One

of the main reasons to use OFDM is to increase the robustness against frequency selective fading or narrowband interference. In a single carrier system, a single fade or interferer can cause the entire link to fail, but in a multicarrier system, only a small percentage of the subcarriers will be affected. Error correction coding can then be used to correct for the few erroneous subcarriers. The concept of using parallel data transmission and frequency division multiplexing was published in the mid-1960s [64, 65]. Some early development is traced back to the 1950s [66]. A U.S. patent was filed and issued in January, 1970 [67].

In a classical parallel data system, the total signal frequency band is divided into N nonoverlapping frequency subchannels. Each subchannel is modulated with a separate symbol and then the N subchannels are frequency-multiplexed. It seems good to avoid spectral overlap of channels to eliminate interchannel interference. However, this leads to inefficient use of the available spectrum. To cope with the inefficiency, the ideas proposed from the mid-1960s were to use parallel data and FDM with overlapping subchannels, in which each carrying a signaling rate b is spaced b apart in frequency to avoid the use of high-speed equalization and to combat impulsive noise and multipath distortion, as well as to fully use the available bandwidth.

Figure 1.10 illustrates the difference between the conventional nonoverlapping multicarrier technique and the overlapping multicarrier modulation technique. As shown in Figure 1.10, by using the overlapping multicarrier modulation technique, we save almost 50% of bandwidth. To realize the overlapping multicarrier technique, however we need to reduce crosstalk between subcarriers, which means that we want orthogonality between the different modulated carriers.

The word orthogonal indicates that there is a precise mathematical relationship between the frequencies of the carriers in the system. In a normal frequency-division multiplex system, many carriers are spaced apart in such a way that the signals can be received using conventional filters and demodulators. In such receivers, guard bands are introduced between the different carriers and in the frequency domain which results in a lowering of spectrum efficiency.

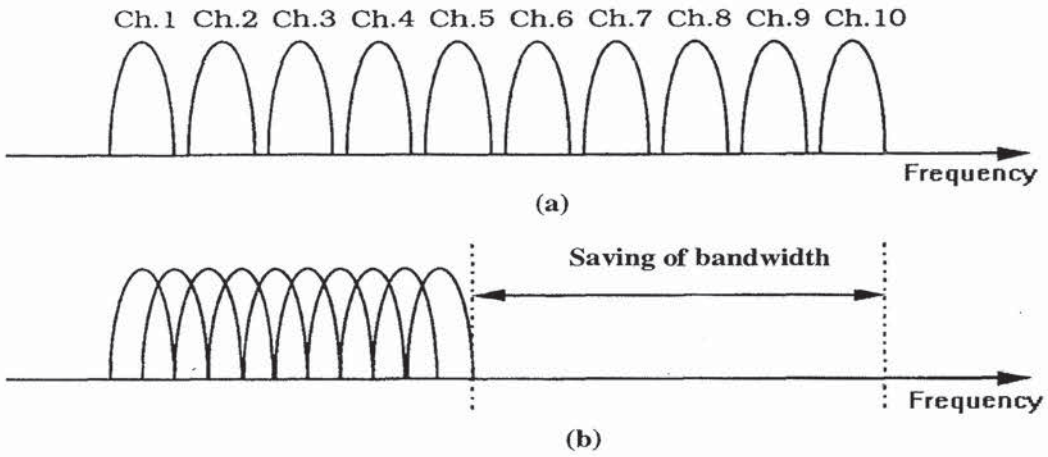


Figure 1.10 Concept of OFDM signal: (a) Conventional multicarrier technique, and (b) orthogonal multicarrier modulation technique.

It is possible, however, to arrange the carriers in an OFDM signal so that the sidebands of the individual carriers overlap and the signals are still received without adjacent carrier interference. To do this the carriers must be mathematically orthogonal. The receiver acts as a bank of demodulators, translating each carrier down to DC, with the resulting signal integrated over a symbol period to recover the raw data. If the other carriers all beat down the frequencies that, in the time domain, have a whole number of cycles in the symbol period T, then the integration process results in zero contribution from all these other carriers. Thus, the carriers are linearly independent (i.e., orthogonal) if the carrier spacing is a multiple of $1/T$. Chapter 2 presents in detail the basic principle of OFDM.

Much of the research focuses on the high efficient multicarrier transmission scheme based on "orthogonal frequency" carriers. In 1971, Weinstein and Ebert [68] applied the discrete Fourier transform (DFT) to parallel data transmission systems as part of the modulation and demodulation process. Figure 1.11 (a) shows the spectrum of the individual data of the subchannel. The OFDM signal, multiplexed in the individual spectra with a frequency spacing b equal to the transmission speed of each subcarrier, is shown in Figure 1.11(b). Figure 1.11 shows that at the center frequency of each subcarrier, there is no crosstalks from other channels. Therefore, if we use DFT at the receiver and calculate correlation values with the center of frequency of each subcarrier, we recover the transmitted data with no crosstalk. In addition, using the DFT-based multicarrier technique, frequency-division multiplex is achieved not by bandpass filtering but by baseband processing.

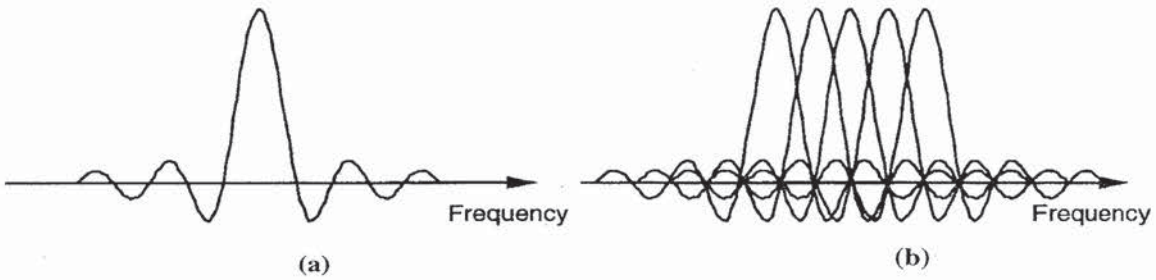


Figure 1.11 Spectra of (a) an OFDM subchannel and (b) an OFDM signal.

Moreover, to eliminate the banks of subcarrier oscillators and coherent demodulators required by frequency-division multiplex, completely digital implementations could be built around special-purpose hardware performing the fast Fourier transform (FFT), which is an efficient implementation of the DFT. Recent advances in very-large-scale integration (VLSI) technology make high-speed, large-size FFT chips commercially affordable. Using this method, both transmitter and receiver are implemented using efficient FFT techniques that reduce the number of operations from N^2 in DFT down to $N\log N$ [69].

In the 1960s, the OFDM technique was used in several high-frequency military systems such as KINEPLEX [66], ANDEFT [70], and KATHRYN [71]. For example, the variable-rate data modem in KATHRYN was built for the high-frequency band. It used up to 34 parallel low-rate phase-modulated channels with a spacing of 82 Hz.

In the 1980s, OFDM was studied for high-speed modems, digital mobile communications, and high-density recording. One of the systems realized the OFDM techniques for multiplexed QAM using DFT [72], and by using pilot tone, stabilizing carrier and clock frequency control and implementing trellis coding are also implemented [73]. Moreover, various-speed modems were developed for telephone networks [74].

In the 1990s, OFDM was exploited for wideband data communications over mobile radio FM channels, high-bit-rate digital subscriber lines (HDSL; 1.6 Mbps), asymmetric digital subscriber lines (ADSL; up to 6 Mbps), very-high-speed digital subscriber lines (VDSL; 100 Mbps), digital audio broadcasting (DAB), and high-definition television (HDTV) terrestrial broadcasting [75–80].

The OFDM transmission scheme has the following key advantages:

- OFDM is an efficient way to deal with multipath; for a given delay spread, the implementation complexity is significantly lower than that of a single carrier system with an equalizer.
- In relatively slow time-varying channels, it is possible to significantly enhance the capacity by adapting the data rate per subcarrier according to the signal-to-noise ratio of that particular subcarrier.

- OFDM is robust against narrowband interference, because such interference affects only a small percentage of the subcarriers.
- OFDM makes single-frequency networks possible, which is especially attractive for broadcasting applications.

On the other hand, OFDM also has some drawbacks compared with single-carrier modulation:

- OFDM is more sensitive to frequency offset and phase noise.
- OFDM has a relatively large peak-to-average power ratio, which tends to reduce the power efficiency of the RF amplifier.

1.6 PREVIEW OF THE BOOK

This book consists of 10 chapters. It covers all the necessary elements to achieve the OFDM-based WBMCS. In Chapter 2, the basics of OFDM are presented. It is explained how an OFDM signal is formed using the inverse fast Fourier transform, how the cyclic extension helps to mitigate the effects of multipath, and how windowing can limit the out-of-band radiation. Basic design rules are given how to choose the OFDM parameters, given a required bandwidth, multipath delay spread, and maximum Doppler spread.

In Chapter 3, we explain how coding and interleaving can be used to mitigate the effects of frequency-selective fading channels. *Quadrature amplitude modulation* is introduced as an appropriate modulation technique for the OFDM subcarriers.

Synchronization of the symbol clock and carrier frequency is the subject of Chapter 4. First, we discuss the sensitivity of OFDM to synchronization errors. Then, we describe different sychronization techniques. Special attention is given to packet transmission, which requires rapid synchronization at the beginning of each packet with a minimum of training overhead.

Chapter 5 describes channel estimation; that is, estimating the reference phases and amplitudes of all subcarriers. Various coherent and differential techniques are explained in conjunction with their relative merits on signal-to-noise ratio performance, overhead, and buffering delay.

The relatively large peak-to-average power (PAP) ratio of OFDM is the subject of Chapter 6. The distribution of the PAP ratio is shown, and various methods to reduce the PAP ratio are explained. This chapter demonstrates that for an arbitrary number of subcarriers, the PAP ratio can be reduced to about 5 dB without a significant loss in performance. This means that the required backoff—and hence the efficiency—of an RF power amplifier for an OFDM system is not much different from that for a single-carrier QPSK system.

In Chapter 7, we explain the basics of direct-sequence and frequency-hopping CDMA, which is especially helpful in understanding the following chapters on combinations of OFDM and CDMA.

Chapter 8 describes multicarrier CDMA. Different techniques with their transmitter and receiver architectures are introduced. Advantages and disadvantages compared with other CDMA techniques are discussed.

Orthogonal FDMA and frequency-hopping CDMA are the subjects of Chapter 9. It shows how OFDM and frequency hopping can be combined to get a multiple-access system with similar advantages as direct-sequence CDMA.

Finally, some applications of OFDM systems are described in Chapter 10: digital audio broadcasting, digital video broadcasting, wireless ATM in the Magic WAND project, and the new IEEE 802.11 and ETSI BRAN OFDM standards.

REFERENCES

[1] Prasad, R., *Universal Wireless Personal Communications*, Norwood, MA: Artech House, 1998.

[2] Prasad, R., "Wireless Broadband Communication Systems," *IEEE Comm. Mag.*, Vol. 35, p. 18, Jan. 1997.

[3] Honcharenko, W., J. P. Kruys, D. Y. Lee, and N.J. Shah, "Broadband Wireless Access," *IEEE Comm. Mag.,* Vol. 35, pp. 20–26, Jan. 1997.

[4] Correia, L. M., and R. Prasad, "An Overview of Wireless Broadband Communications," *IEEE Comm. Mag.,* Vol. 35, pp. 28–33, Jan. 1997.

[5] Morinaga, M., M. Nakagawa, and R. Kohno, "New Concepts and Technologies for achieving Highly Reliable and High Capacity Multimedia Wireless Communications System," *IEEE Comm. Mag.* Vol. 35, pp. 34–40, Jan. 1997.

[6] Da Silva, J. S., B. Arroyo-Fernandez, B. Barani, J. Pereira, and D. Ikonomou, "Mobile and Personal Communications: ACTS and Beyond," *Proc. PIMRC'97,* Helsinki, Finland, Sep. 1997.

[7] Prisoli, F. D., and R. Velt, "Design of Medium Access Control and Logical Link Control Functions for ATM Support in the MEDIAN System," *Proc. ACTS Mobile Comm. Summit'97,* pp. 734–744, Aalborg, Denmark, Oct. 1997.

[8] Rheinschmitt, R., A. de Haz, and M. Umehinc, "AWACS MAC and LLC Functionality," *Proc. ACTS Mobile Comm. Summit'97,* pp. 745–750, Aalborg, Denmark, Oct. 1997.

[9] van Nee, R., "An OFDM Modem for Wireless ATM," *Proceedings of IEEE 4th Symposium on Communications and Vehicular Technology in the Benelux*, Gent, Belgium, pp. 36-41, Oct. 7-8, 1996.

[10] Aldis, J., E. Busking, T. Kleijne, R. Kopmeiners, and R. van Nee, "Magic Into Reality, Building the WAND Modem," *Proc. ACTS Mobile Comm. Summit'97*, pp. 734–744, Aalborg, Denmark, Oct. 1997.

[11] Mikkonen, J., C. Corrado, C. Erci, and M. Progler, "Emerging Wireless Broadband Networks," *IEEE Comm. Mag.*, Feb. 1988.

[12] Wu, G., Y. Hase, K. Taira, and K. Iwasak, "Emerging Wireless ATM Oriented MAC Protocol for High-Speed Wireless LAN," Proc. PIMRC'97, pp. 198-203, Helsinki, Finland, Sep. 1997.

[13] Prasad, R., and L. M. Correia, "Wireless Broadband Multimedia Communications," *International Wireless and Telecommunications Symposium*, Shah Alam, Malaysia, Vol. 2, pp. 55–70 May 14–16,1997.

[14] Prasad, R., and L. M. Correia, "An Overview of Wireless Broadband Multimedia Communications," *Proc. MoMuC'97*, Seoul, Korea, pp. 17–31, Sep.-Oct. 1997.

[15] Chelouche, M., S. Hethuin, and L. Ramel, "Digital Wireless Broadband Corporate and Private Networks: RNET Concepts and Applications," *IEEE Comm. Mag.*, Vol. 35, pp. 42–51, Jan. 1997.

[16] IEEE 802.11, *IEEE Standard for Wireless LAN Medium Access Control and Physical Layer Specifications*, Nov. 1997.

[17] Fernandes, L., "R2067 MBS - Mobile Broadband System," in *Proc. ICUPC'93 — 2nd IEEE International Conference on Universal Personal Communications*, Ottawa, Canada, Oct.,1993.

[18] IEEE 802.11 Web URL: http://grouper.ieee.org/groups/802/11/.

[19] ETSI home page: http://www.etsi.org/.

[20] Kruys, J., "Standardization of Wireless High Speed Premises Data Networks," *Wireless ATM Workshop*, Espoo, Finland, Sep. 1996.

[21] van Nee, R., G. Awater, M. Morikura, H. Takanashi, M. Webster, and K. Halford, "New High Rate Wireless LAN Standards," *IEEE Communications Magazine*, Dec. 1999.

[22] IEEE, "Supplement to Standard for Telecommunications and Information Exchange Between Systems—LAN/MAN Specific Requirements—Part 11: Wireless MAC and PHY Specifications: Higher Speed Physical Layer Extension in the 2.4-GHz Band," P802.11b/D7.0, July 1999.

[23] WECA home page: http://www.wirelessethernet.com.

[24] IEEE, "Supplement to Standard for Telecommunications and Information Exchange Between Systems—LAN/MAN Specific Requirements—Part 11: Wireless MAC and PHY Specifications: High Speed Physical Layer in the 5-GHz Band," P802.11a/D7.0, July 1999.

[25] WaveLAN home page: http://www.wavelan.com.

[26] Prasad, R., "Research Challenges in Future Wireless Personal Communications: Microwave Perspective," *Proc. 25th European Microwave Conference* (Keynote opening address), Bologna, Italy, pp. 4–11, Sept. 4–7, 1995.

[27] Prasad, R., "Overview of Wireless Personal Communications: Microwave Perspectives," *IEEE Comm. Mag.*, Vol. 35, pp. 104–108, April 1997.

[28] Smulders, P. F. M., "Broadband Wireless LANs: a feasibility study," *Ph.D. Thesis*, Eindhoven Univ. of Technology, Eindhoven, The Netherlands, Dec. 1995.

[29] Correia, L. M., and P. O. Frances, "A Propagation Model for the Estimation of the Average Received Power in an Outdoor Environment at the Millimeter Wave Band," *Proc. VTC'94 — IEEE VTS 44th Vehicular Technology Conference*, Stockholm, Sweden, pp. 1785–1788, July 1994.

[30] Lovnes, G., J. J. Reis, and R. H. Raekken, "Channel Sounding Measurements at 59 GHz in City Streets," *Proc. PIMRC'94 — 5th IEEE International Symposium on Personal, Indoor; and Mobile Radio Communications*, The Hague, The Netherlands, pp. 496–500, Sep. 1994.

[31] Fernandes, J. J., P. A. Watson, and J. C. Neves, "Wireless LANs: Physical Properties of Infrared Systems versus mmWave Systems," *IEEE Comm. Mag.*, Vol. 32, No.8, pp. 68–73, Aug. 1994.

[32] European Commission, "Communications for Society Visionary Research," European Commission, DG XIII/B, Brussels, Belgium, Feb. 1997.

[33] Rokitansky, C. –H., and M. Scheibenbogen (eds.), "Updated Version of System Description Document," *RACE Deliverable R2067/UA/WP215/DS/P/068.b1*, RACE Central Office, European Commission, Brussels, Belgium, Dec. 1995.

[34] Prasad, A. R., "Asynchronous Transfer Mode Based Mobile Broadband System," *Proc. IEEE 3rd Symposium on Comm. & Vech. Technology in Benelux*, Eindhoven, The Netherlands, pp. 143–148, Oct. 1995.

[35] Prasad, R., and L. Vandendorpe, "An Overview of Millimeter Wave Indoor Wireless Communication System," *Proc. 2nd Int. Conf. on Universal Personal Communications*, pp. 885–889, Ottawa, Canada, 1993.

[36] Prasad, R., and L. Vandendorpe, "Cost 231 Project: Performance Evaluation of a Millimetric-Wave Indoor Wireless Communications System," *RACE Mobile Telecommunications Workshop*, Metz, France, pp. 137–144, June 16–18, 1993.

[37] Zauner, K., "On the ATM Road to Broadband," *Telecom Report International*, No.3, pp. 26–29, 1993.

[38] Tubtiang, H.I. Kwon and O. Pujolle, "A Simple ATM Switching Architecture for Broadband-ISDN and Its Performance," *Modeling and Performance Evaluation of ATM Technology (C-I5)*, pp. 361–367, IFIP 1993.

[39] Stallings, W., *Advances in ISDN and Broadband ISDN*, IEEE Computer Society Press, 1992.

[40] Leslie, I. M., D. R. McAuley, and D. L. Tennenhouse, "ATM Everywhere?" *IEEE Network*, pp. 40–46, March 1993.

[41] McTiffin, M. J., A. P. Hulbert, T. J. Ketseoglou, W. Heimsch, and O. Crisp, "Mobile access to an ATM Network Using A CDMA Air Interface," *IEEE J. Selected Areas in Comm.*, Vol. 12, No. 5, pp. 900–908, June 1994.

[42] Raychaudhri, D., "ATM-based Transport Architecture for Multiservices Wireless Personal Communication Networks," *IEEE J. Selected Areas in Comm.*, Vol.12, No. 8, pp. 1401–1414, Oct. 1994.

[43] Geij, R. R., "Mobile Multimedia Scenario using ATM and Microcellular Technologies," *Vehicular Technology*, Vol. 43, No. 3, pp. 699–703, Aug. 1994.

[44] Bakker, J. D., and R. Prasad, "Wireless Multimedia Communications Using a CDMA-based ATM Air Interface," *Proc. IEEE ISSSTA'96*, Mainz, Germany, pp. 1128–1132, Sept. 1996.

[45] Hashemi, H., "The Indoor Radio Propagation Channel," *Proceedings of the IEEE*, Vol. 81, no. 7, pp. 943-968, July 1993.

[46] Saleh, A. A. M., and R. A. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, no. 2, pp. 128-137, Feb. 1987.

[47] Halls, G., "HIPERLAN Radio Channel Models and Simulation Results," RES10TTG 93/58.

[48] Chayat, N., "Tentative Criteria for Comparison of Modulation Methods," IEEE P802.11-97/96.

[49] Devasirvatham, D. M. J., "Multi-Frequency Propagation Measurements and Models in a Large Metropolitan Commercial Building for Personal Communications," *IEEE PIMRC '91*, pp. 98-103, London, UK, Sept. 23-25, 1991.

[50] Devasirvatham, D. M. J., "Time Delay Spread Measurements at 850 MHz and 1.7 GHz inside a Metropolitan Office Building," *Electronics Letters*, pp. 194–196, Feb. 2, 1989.

[51] Devasirvatham, D. M. J., M. J. Krain and D. A. Rappaport, "Radio Propagation Measurements at 850 MHz, 1.7 GHz and 4 GHz Inside Two Dissimilar Office buildings," *Electronics Letters*, Vol. 26, No. 7, pp. 445-447, March 29, 1990.

[52] Hafesi, P., D. Wedge, M. Beach, M. Lawton, "Propagation Measurements at 5.2 GHz in Commercial and Domestic Environments," *IEEE PIMRC '97*, Helsinki, pp. 509-513, Sept. 1-4, 1997.

[53] Hawbaker, D. A., and T. S. Rappaport, "Indoor Wideband Radiowave Propagation Measurements at 1.3 GHz and 4.0 GHz," *Electronics Letters*, vol. 26, no. 21, pp. 1800-1802, Oct. 11, 1990.

[54] Johnson, I. T., and E. Gurdenelli, "Measurements of Wideband Channel Characteristics in Cells Within Man Made Structures of Area Less Than 0.2 km^2," COST 231 TD(90)083.

[55] Lahteenmaki, J., "Indoor Measurements and Simulation of Propagation at 1.7 GHz," COST 231 TD(90)084.

[56] Anderson, P. C., O. J. M. Houen, K. Kladakis, K. T. Peterson, H. Fredskild and I. Zarnoczay, "Delay Spread Measurements at 2 GHz," COST 231 TD(91)029.

[57] Bultitude, R. J. C., et al., "The Dependance of Indoor Radio Channel Multipath Characteristics on Transmit/Receive Ranges," *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 7, pp. 979 - 990, Sep. 1993.

[58] Bultitude, R. J. C., S. A. Mahmoud and W. A. Sullivan, "A Comparison of Indoor Radio Propagation Characteristics at 910 MHz and 1.75 GHz," *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 1, pp. 20-30, Jan. 1989.

[59] Pahlavan, K., and S. J. Howard, "Frequency Domain Measurements of Indoor Radio Channels," *Electronics Letters*, vol. 25, no. 24, pp. 1645 - 1647, Nov. 23, 1989.

[60] Davies, R., and J. P. McGeehan, "Propagation Measurements at 1.7 GHz for Microcellular Urban Communications," *Electronics Letters*, vol. 26, no. 14, pp. 1053-1055, July 5, 1990.

[61] Zollinger, E., and A. Radovic, "Measured Time Variant Characteristics of Radio Channels in the Indoor Environment," COST 231 TD(91)089.

[62] Rappaport, T. S., and C. D. McGillem, "Characterization of UHF Multipath Radio Channels in Factory Buildings," *IEEE Transactions on Antennas and Propagation*, Vol. 37, No. 8, pp. 1058-1069, Aug. 1989.

[63] Nobles, P., and F. Halsall, "Delay Spread and Received Power Measurements Within a Building at 2 GHz, 5 GHz, and 17 GHz', *IEE Tenth International Conference on Antennas and Propagation*, Edinburgh, UK, April 14-17, 1997.

[64] R. W. Chang, "Synthesis of Band Limited Orthogonal Signals for Multichannel Data Transmission," *Bell Syst. Tech. J.,* Vol. 45, pp. 1775–1796, Dec. 1996.

[65] Salzberg, B. R., "Performance of an efficient parallel data transmission system," *IEEE Trans. Comm.,* Vol. COM-15, pp. 805 – 813, Dec. 1967.

[66] Mosier, R. R., and R.G. Clabaugh, "Kineplex, a Bandwidth Efficient Binary Transmission System," *AIEE Trans.,* Vol. 76, pp. 723 – 728, Jan. 1958.

[67] "Orthogonal Frequency Division Multiplexing," U.S. Patent No. 3, 488,4555, filed November 14, 1966, issued Jan. 6, 1970.

[68] Weinstein, S. B., and P .M. Ebert, "Data Transmission by Frequency Division Multiplexing Using the Discrete Fourier Transform," *IEEE Trans. Comm.,* Vol. COM-19, pp. 628 – 634, Oct. 1971.

[69] Zou, W. Y., and Y. Wu, "COFDM: an overview," *IEEE Trans. Broadc.,* Vol. 41, No. 1, pp. 1 – 8, March 1995.

[70] Porter, G. C., "Error Distribution and Diversity Performance of a Frequency Differential PSK HF modem," *IEEE Trans. Comm.,* Vol., COM-16, pp. 567–575, Aug. 1968.

[71] Zimmerman, M. S., and A. L. Kirsch, "The AN/GSC-10 (KATHRYN) variable rate data modem for HF radio," *IEEE Trans. Comm.,* Vol., COM-15, pp. 197 – 205, April 1967.

[72] Hirosaki, B., "An Orthogonally Multiplexed QAM system Using the Discrete Fourier Transform," *IEEE Trans. Comm.,* Vol., COM-29, pp. 982 – 989, July 1981.

[73] Hirosaki, B., "A 19.2 kbits Voice Band Data Modem Based on Orthogonality Multiplexed QAM Techniques," *Proc. of IEEE ICC'85,* pp. 21.1.1 – 5, 1985.

[74] Keasler, W. E., and D. L. Bitzer, "High speed modem suitable for operating with a switched network," U.S. Patent No. 4,206,320, June 1980.

[75] Chow, P. S., J. C. Tu and J. M. Cioffi, "Performance Evaluation of a Multichannel Transceiver System for ADSL and VHDSL services," *IEEE J. Selected Area,* Vol., SAC-9, No. 6, pp. 909 – 919, Aug. 1991.

[76] Chow, P. S., J. C. Tu, and J. M. Cioffi, "A Discrete Multitone Transceiver System for HDSL Applications," *IEEE J. Selected Areas in Comm.,* Vol. SAC-9, No. 6, pp. 909 – 919, Aug. 1991.

[77] Paiement, R. V., "Evaluation of Single Carrier and Multicarrier Modulation Techniques for Digital ATV Terrestrial Broadcasting," *CRC Report,* No. CRC-RP-004, Ottawa, Canada, Dec. 1994.

[78] Sari, H., G. Karma, and I. Jeanclaude, "Transmission Techniques for Digital Terrestrial TV Broadcasting," *IEEE Comm. Mag.*, Vol.33, pp. 100–109, Feb. 1995.

[79] Oppenheim, A.V., and R.W. Schaffer, *Discrete-time Signal Processing*, Prentice-Hall International, ISBN 0-13-216771-9, 1989.

[80] Hara, S., M. Mouri, M. Okada, and N. Morinaga, "Transmission Performance Analysis of Multi-Carrier Modulation in Frequency Selective Fast Rayleigh Fading Channel," in *Wireless Personal Communications*, Kluwer Academic Publishers, Vol.2, pp. 335 – 356, 1996.

[79] Rappaport, T. S., *Wireless Communication Principles & Practice*, Prentice-Hall PTR, New Jersey, 1996.

[80] Steel, R., (ed.), *Mobile Radio Communications*, IEEE Press, 1994.

[81] Macrario, R. C. V., (ed.), "Personal & Mobile Communications," Peter Peregrinus Ltd., 1991.

[82] Howard, S. J., and K. Rappaport, "Performance of a DFE Modem Evaluted From Measured Indoor Radio Multipath Profiles," Proc. ICC'90, Vol. 4, pp. 335.2.1 – 335.2.5, April 1990.

Chapter 2

OFDM Basics

2.1 INTRODUCTION

The basic principle of OFDM is to split a high-rate datastream into a number of lower rate streams that are transmitted simultaneously over a number of subcarriers. Because the symbol duration increases for the lower rate parallel subcarriers, the relative amount of dispersion in time caused by multipath delay spread is decreased. Intersymbol interference is eliminated almost completely by introducing a guard time in every OFDM symbol. In the guard time, the OFDM symbol is cyclically extended to avoid intercarrier interference. This whole process of generating an OFDM signal and the reasoning behind it are described in detail in sections 2.2 to 2.4.

In OFDM system design, a number of parameters are up for consideration, such as the number of subcarriers, guard time, symbol duration, subcarrier spacing, modulation type per subcarrier, and the type of forward error correction coding. The choice of parameters is influenced by system requirements such as available bandwidth, required bit rate, tolerable delay spread, and Doppler values. Some requirements are conflicting. For instance, to get a good delay spread tolerance, a large number of subcarriers with a small subcarrier spacing is desirable, but the opposite is true for a good tolerance against Doppler spread and phase noise. These design issues are discussed in Section 2.5. Section 2.6 gives an overview of OFDM signal processing functions, while Section 2. ends this chapter with a complexity comparison of OFDM versus single-carrier systems.

2.2 GENERATION OF SUBCARRIERS USING THE IFFT

An OFDM signal consists of a sum of subcarriers that are modulated by using *phase shift keying (PSK)* or *quadrature amplitude modulation (QAM)*. If d_i are the complex

ZYXEL-1015
Page 048 (of 277)

QAM symbols, N_s is the number of subcarriers, T the symbol duration, and f_c the carrier frequency, then one OFDM symbol starting at $t = t_s$ can be written as

$$s(t) = \mathrm{Re}\left\{ \sum_{i=-\frac{N_s}{2}}^{\frac{N_s}{2}-1} d_{i+N_s/2} \exp\left(j2\pi \left(f_c - \frac{i+0.5}{T} \right)(t-t_s) \right) \right\} \quad , \; t_s \le t \le t_s + T \tag{2.1}$$

$$s(t) = 0 \, , \quad t < t_s \;\wedge\; t > t_s + T$$

In the literature, often the equivalent complex baseband notation is used, which is given by (2.2). In this representation, the real and imaginary parts correspond to the in-phase and quadrature parts of the OFDM signal, which have to be multiplied by a cosine and sine of the desired carrier frequency to produce the final OFDM signal. Figure 2.1 shows the operation of the OFDM modulator in a block diagram.

$$s(t) = \sum_{i=-\frac{N_s}{2}}^{\frac{N_s}{2}-1} d_{i+N_s/2} \exp\left(j2\pi \frac{i}{T}(t-t_s) \right) \quad , \; t_s \le t \le t_s + T \tag{2.2}$$

$$s(t) = 0 \, , \quad t < t_s \;\wedge\; t > t_s + T$$

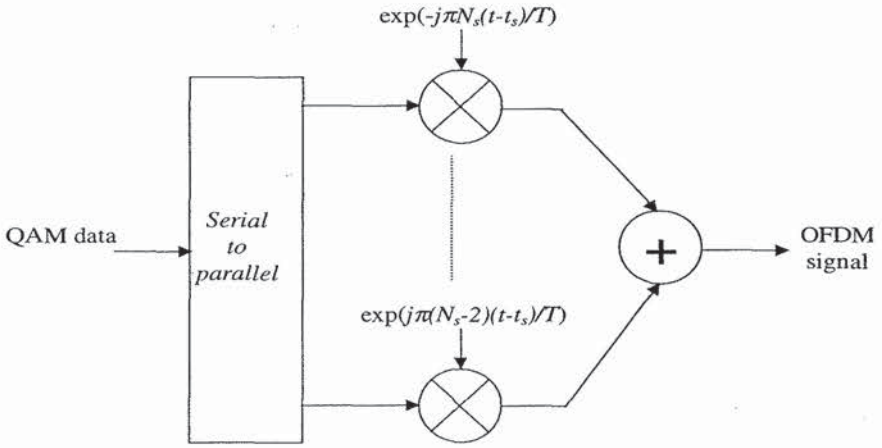


Figure 2.1 OFDM modulator.

As an example, Figure 2.2 shows four subcarriers from one OFDM signal. In this example, all subcarriers have the same phase and amplitude, but in practice the amplitudes and phases may be modulated differently for each subcarrier. Note that each subcarrier has exactly an integer number of cycles in the interval T, and the number of cycles between adjacent subcarriers differs by exactly one. This property accounts for

the orthogonality between the subcarriers. For instance, if the jth subcarrier from (2.2) is demodulated by downconverting the signal with a frequency of j/T and then integrating the signal over T seconds, the result is as written in (2.3). By looking at the intermediate result, it can be seen that a complex carrier is integrated over T seconds. For the demodulated subcarrier j, this integration gives the desired output $d_{j+N/2}$ (multiplied by a constant factor T), which is the QAM value for that particular subcarrier. For all other subcarriers, the integration is zero, because the frequency difference $(i\text{-}j)/T$ produces an integer number of cycles within the integration interval T, such that the integration result is always zero.

$$\int_{t_s}^{t_s+T} \exp(-j2\pi \frac{j}{T}(t-t_s)) \sum_{i=-\frac{N_s}{2}}^{\frac{N_s}{2}-1} d_{i+N_s/2} \exp(j2\pi \frac{i}{T}(t-t_s))\,\mathrm{dt}$$
$$= \sum_{i=-\frac{N_s}{2}}^{\frac{N_s}{2}-1} d_{i+N_s/2} \int_{t_s}^{t_s+T} \exp(j2\pi \frac{i-j}{T}(t-t_s))dt = d_{j+N_s/2}T \tag{2.3}$$

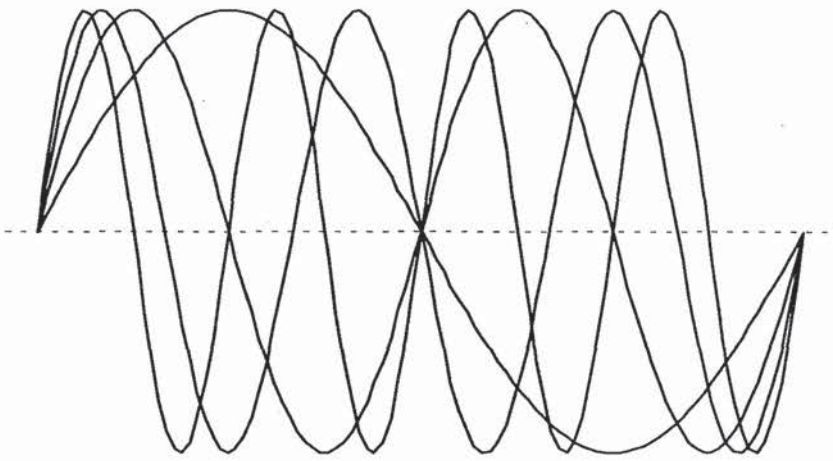


Figure 2.2 Example of four subcarriers within one OFDM symbol.

The orthogonality of the different OFDM subcarriers can also be demonstrated in another way. According to (2.1), each OFDM symbol contains subcarriers that are nonzero over a T-second interval. Hence, the spectrum of a single symbol is a convolution of a group of Dirac pulses located at the subcarrier frequencies with the spectrum of a square pulse that is one for a T-second period and zero otherwise. The amplitude spectrum of the square pulse is equal to $\mathrm{sinc}(\pi fT)$, which has zeros for all frequencies f that are an integer multiple of $1/T$. This effect is shown in Figure 2.2, which shows the overlapping sinc spectra of individual subcarriers. At the maximum of each subcarrier spectrum, all other subcarrier spectra are zero. Because an OFDM

receiver essentially calculates the spectrum values at those points that correspond to the maxima of individual subcarriers, it can demodulate each subcarrier free from any interference from the other subcarriers. Basically, Figure 2.3 shows that the OFDM spectrum fulfills Nyquist's criterium for an intersymbol interference free pulse shape. Notice that the pulse shape is present in the frequency domain and not in the time domain, for which the Nyquist criterium usually is applied. Therefore, instead of inter-symbol interference (ISI), it is intercarrier interference (ICI) that is avoided by having the maximum of one subcarrier spectrum correspond to zero crossings of all the others.

The complex baseband OFDM signal as defined by (2.2) is in fact nothing more than the inverse Fourier transform of N_s QAM input symbols. The time discrete equivalent is the inverse discrete Fourier transform (IDFT), which is given by (2.4), where the time t is replaced by a sample number n. In practice, this transform can be implemented very efficiently by the inverse fast Fourier transform (IFFT). An N point IDFT requires a total of N^2 complex multiplications—which are actually only phase rotations. Of course, there are also additions necessary to do an IDFT, but since the hardware complexity of an adder is significantly lower than that of a multiplier or phase rotator, only the multiplications are used here for comparison. The IFFT drastically reduces the amount of calculations by exploiting the regularity of the operations in the IDFT. Using the radix-2 algorithm, an N-point IFFT requires only $(N/2)\cdot\log_2(N)$ complex multiplications [1]. For a 16-point transform, for instance, the difference is 256 multiplications for the IDFT versus 32 for the IFFT—a reduction by a factor of 8! This difference grows for larger numbers of subcarriers, as the IDFT complexity grows quadratically with N, while the IFFT complexity only grows slightly faster than linear.

$$s(n) = \sum_{i=0}^{N_s-1} d_i \exp(j2\pi \frac{in}{N}) \tag{2.4}$$

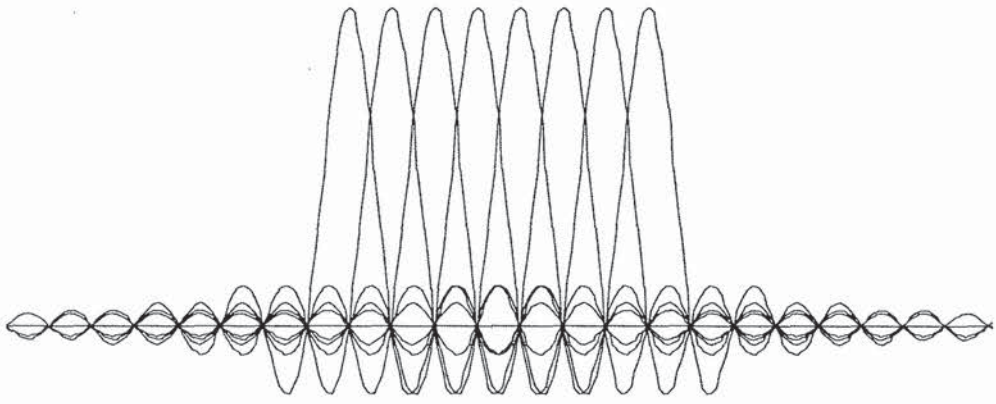


Figure 2.3 Spectra of individual subcarriers.

The number of multiplications in the IFFT can be reduced even further by using a radix-4 algorithm. This technique makes use of the fact that in a four-point IFFT, there are only multiplications by $\{1,-1,j,-j\}$, which actually do not need to be implemented by a full multiplier, but rather by a simple add or subtract and a switch of real and imaginary parts in the case of multiplications by j or $-j$. In the radix-4 algorithm, the transform is split into a number of these trivial four-point transforms, and non-trivial multiplications only have to be performed between stages of these four-point transforms. In this way, an N-point FFT using the radix-4 algorithm requires only $(3/8)N(\log_2 N\text{-}2)$ complex multiplications or phase rotations and $N\log_2 N$ complex additions [1]. For a 64-point FFT, for example, this means 96 rotations and 384 additions, or 1.5 and 6 rotations and additions per sample, respectively.

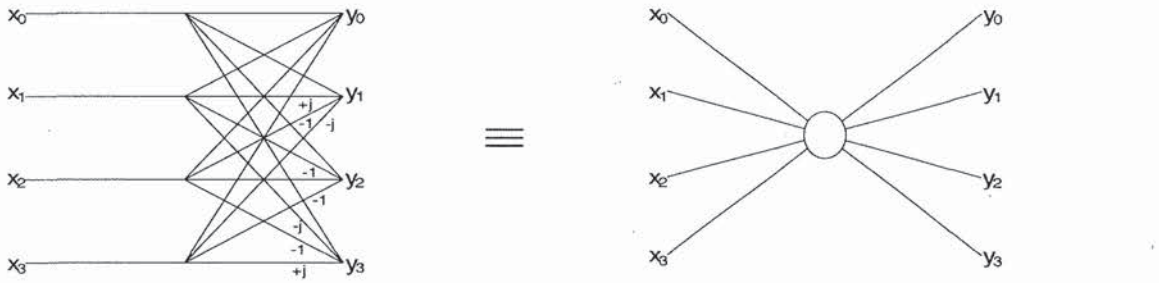


Figure 2.4 The radix-4 butterfly.

Figure 2.4 shows the four-point IFFT, which is known as the *radix-4 butterfly* that forms the basis for constructing larger IFFT sizes [1]. Four input values x_0 to x_3 are transformed into output values y_0 to y_3 by simple additions and trivial phase rotations. For instance, y_1 is given by $x_0 + jx_1 - x_2 - jx_3$, which can be calculated by doing four additions plus a few additional I/Q swappings and inversions to account for the multiplications by j and -1.

The radix-4 butterfly can be used to efficiently build an IFFT with a larger size. For instance, a 16-point IFFT is depicted in Figure 2.5. The 16-point IFFT contains two stages with four radix-4 butterflies, separated by an intermediate stage where the 16 intermediate results are phase rotated by the twiddle factor ω^i, which is defined as $exp(j2\pi i/N)$. Notice that for N=16, rotation by the twiddle factor ω^i reduces to a trivial operation for i=0, 4, 8 and 12, where ω^i is 1, j, -1 and $-j$, respectively. Taking this into account, the 16-point IFFT actually contains only eight non-trivial phase rotations, which is a factor of 32 smaller than the amount of phase rotations for the IDFT. These non-trivial phase rotations largely determine the implementation complexity, because the complexity of a phase rotation or complex multiplication is an order of magnitude larger than the complexity of an addition.

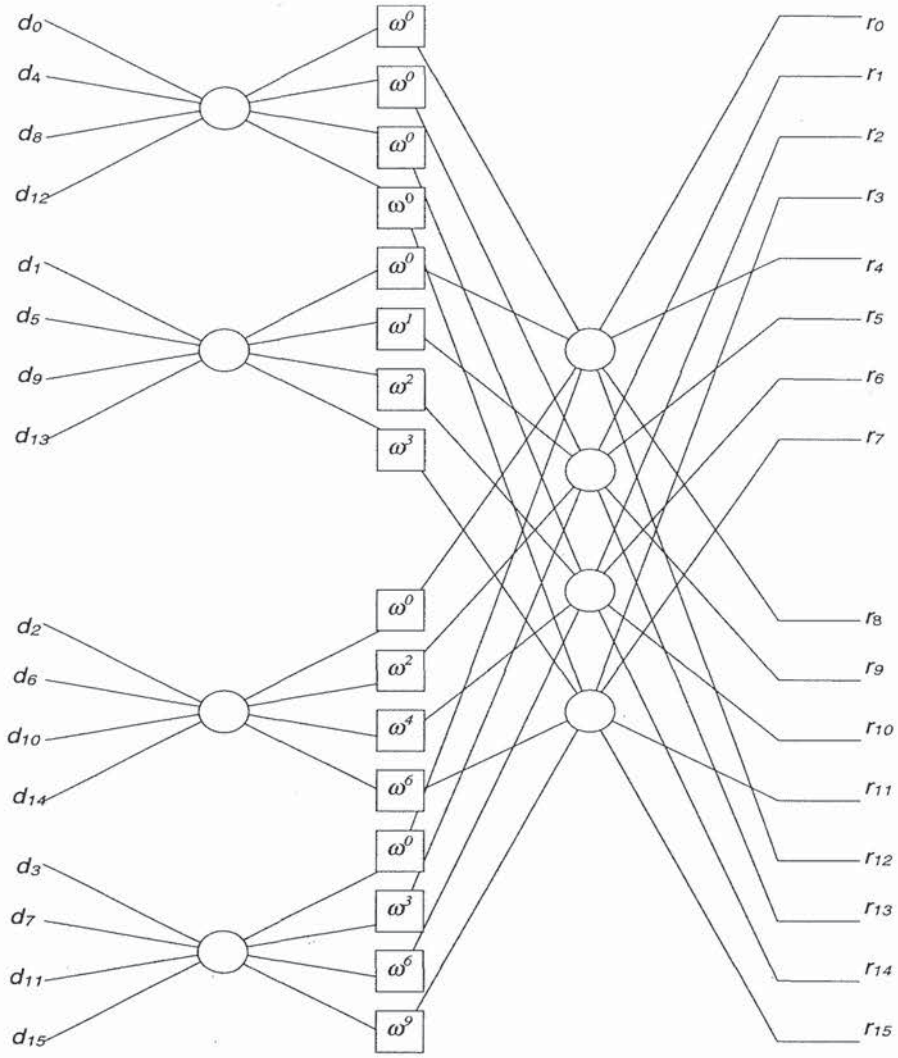


Figure 2.5 16-point IFFT using the radix-4 algorithm.

As an example of how to generate an OFDM symbol, let us assume that we want to transmit eight binary values {1 1 1 –1 1 1 –1 1} on eight subcarriers. The IDFT or IFFT then has to calculate:

$$\frac{1}{8}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \frac{1}{2}\sqrt{2}(1+j) & j & \frac{1}{2}\sqrt{2}(-1+j) & -1 & \frac{1}{2}\sqrt{2}(-1-j) & -j & \frac{1}{2}\sqrt{2}(1-j) \\ 1 & j & -1 & -j & 1 & j & -1 & -j \\ 1 & \frac{1}{2}\sqrt{2}(-1+j) & -j & \frac{1}{2}\sqrt{2}(1+j) & -1 & \frac{1}{2}\sqrt{2}(1-j) & j & \frac{1}{2}\sqrt{2}(-1-j) \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & \frac{1}{2}\sqrt{2}(-1-j) & j & \frac{1}{2}\sqrt{2}(1-j) & -1 & \frac{1}{2}\sqrt{2}(1+j) & -j & \frac{1}{2}\sqrt{2}(-1+j) \\ 1 & -j & -1 & j & 1 & -j & -1 & j \\ 1 & \frac{1}{2}\sqrt{2}(1-j) & -j & \frac{1}{2}\sqrt{2}(-1-j) & -1 & \frac{1}{2}\sqrt{2}(-1+j) & j & \frac{1}{2}\sqrt{2}(1+j) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \\ 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} 4 \\ \sqrt{2}(1+j(\sqrt{2}-1)) \\ 2+2j \\ -\sqrt{2}(1+j(\sqrt{2}+1)) \\ 0 \\ -\sqrt{2}(1-j(\sqrt{2}+1)) \\ 2-2j \\ \sqrt{2}(1-j(\sqrt{2}-1)) \end{bmatrix}$$

(2.5)

The left-hand side of (2.5) contains the IDFT matrix, where every column corresponds to a complex subcarrier with a normalized frequency ranging from -4 to +3. The right-hand side of (2.5) gives the eight IFFT output samples that form one OFDM symbol. In practice, however, these samples are not enough to make a real OFDM signal. The reason is that there is no oversampling present, which would introduce intolerable aliasing if one would pass these samples though a digital-to-analog converter. To introduce oversampling, a number of zeros can be added to the input data. For instance, eight zeros could be added to the eight input samples of the previous example, after which a 16-point IFFT can be performed to get 16 output samples of a twice-oversampled OFDM signal. Notice that in the complex IFFT as in (2.5), the first half of the rows correspond to positive frequencies while the last half correspond to negative frequencies. Hence, if oversampling is used, the zeros should be added in the middle of the data vector rather than appending them at the end. This ensures the zero data values are mapped onto frequencies close to plus and minus half the sampling rate, while the nonzero data values are mapped onto the subcarriers around 0 Hz. For the data of the previous example, the oversampled input vector would become {1 1 1 −1 0 0 0 0 0 0 0 0 1 1 −1 1}.

2.3 GUARD TIME AND CYCLIC EXTENSION

One of the most important reasons to do OFDM is the efficient way it deals with multipath delay spread. By dividing the input datastream in N_s subcarriers, the symbol duration is made N_s times smaller, which also reduces the relative multipath delay spread, relative to the symbol time, by the same factor. To eliminate intersymbol interference almost completely, a guard time is introduced for each OFDM symbol. The guard time is chosen larger than the expected delay spread, such that multipath components from one symbol cannot interfere with the next symbol. The guard time could consist of no signal at all. In that case, however, the problem of intercarrier interference (ICI) would arise. ICI is crosstalk between different subcarriers, which means they are no longer orthogonal. This effect is illustrated in Figure 2.6. In this example, a subcarrier 1 and a delayed subcarrier 2 are shown. When an OFDM receiver

tries to demodulate the first subcarrier, it will encounter some interference from the second subcarrier, because within the FFT interval, there is no integer number of cycles difference between subcarrier 1 and 2. At the same time, there will be crosstalk from the first to the second subcarrier for the same reason.

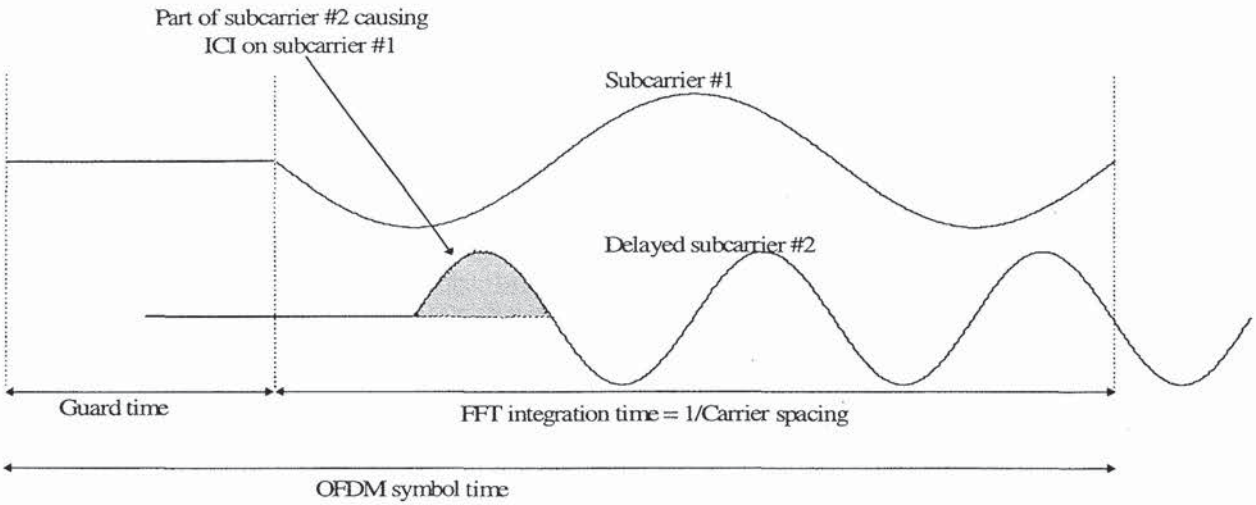


Figure 2.6 Effect of multipath with zero signal in the guard time; the delayed subcarrier 2 causes ICI on subcarrier 1 and vice versa.

To eliminate ICI, the OFDM symbol is cyclically extended in the guard time, as shown in Figure 2.7. This ensures that delayed replicas of the OFDM symbol always have an integer number of cycles within the FFT interval, as long as the delay is smaller than the guard time. As a result, multipath signals with delays smaller than the guard time cannot cause ICI.

As an example of how multipath affects OFDM, Figure 2.8 shows received signals for a two-ray channel, where the dotted curve is a delayed replica of the solid curve. Three separate subcarriers are shown during three symbol intervals. In reality, an OFDM receiver only sees the sum of all these signals, but showing the separate components makes it more clear what the effect of multipath is. From the figure, we can see that the OFDM subcarriers are BPSK modulated, which means that there can be 180-degree phase jumps at the symbol boundaries. For the dotted curve, these phase jumps occur at a certain delay after the first path. In this particular example, this multipath delay is smaller than the guard time, which means there are no phase transitions during the FFT interval. Hence, an OFDM receiver "sees" the sum of pure sine waves with some phase offsets. This summation does not destroy the orthogonality between the subcarriers, it only introduces a different phase shift for each subcarrier. The orthogonality does become lost if the multipath delay becomes larger than the

guard time. In that case, the phase transitions of the delayed path fall within the FFT interval of the receiver. The summation of the sine waves of the first path with the phase modulated waves of the delayed path no longer gives a set of orthogonal pure sine waves, resulting in a certain level of interference.

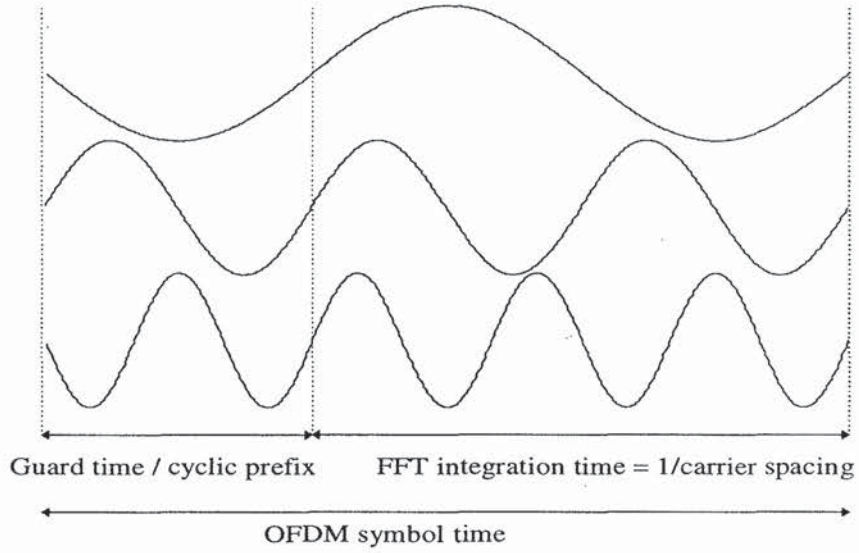


Figure 2.7 OFDM symbol with cyclic extension.

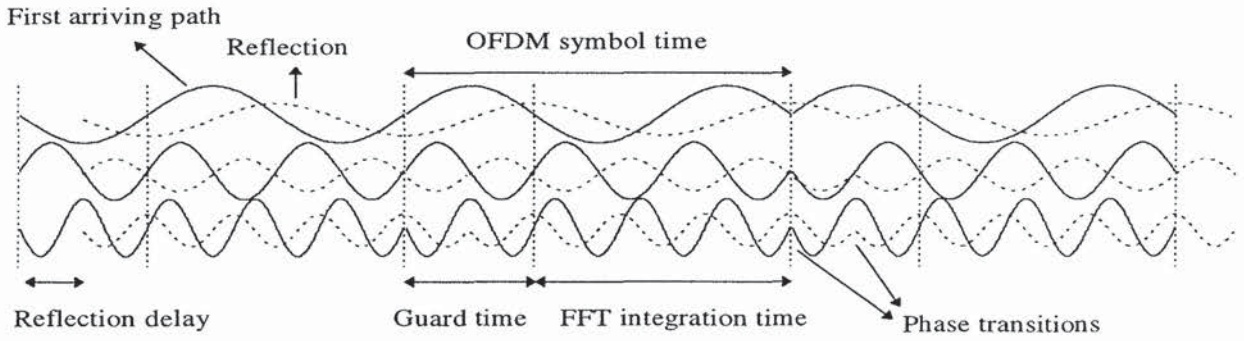
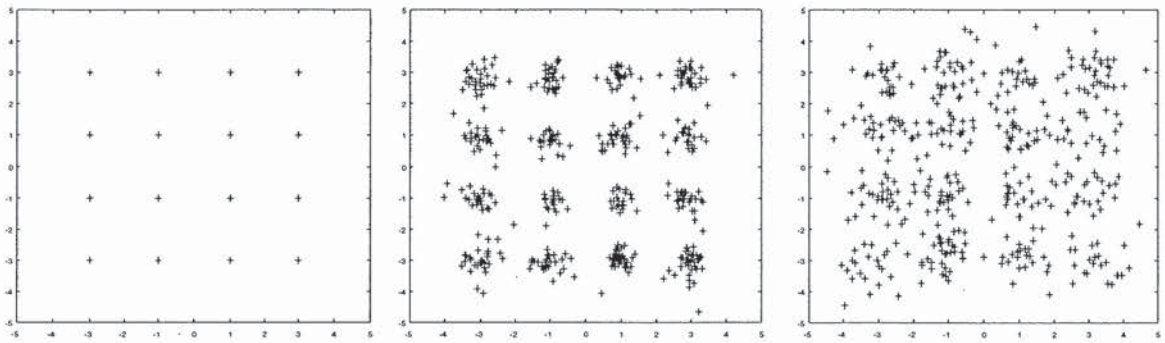


Figure 2.8 Example of an OFDM signal with three subcarriers in a two-ray multipath channel. The dashed line represents a delayed multipath component.

To get an idea what level of interference is introduced when the multipath delay exceeds the guard time, Figure 2.9 depicts three constellation diagrams that were derived from a simulation of an OFDM link with 48 subcarriers, each modulated by using 16-QAM. Figure 2.9(a) shows the undistorted 16-QAM constellation, which is observed whenever the multipath delay is below the guard time. In Figure 2.9(b), the multipath delay exceeds the guard time by a small 3% fraction of the FFT interval.

Hence, the subcarriers are not orthogonal anymore, but the interference is still small enough to get a reasonable received constellation. In Figure 2.9(c), the multipath delay exceeds the guard time by 10% of the FFT interval. The interference is now so large that the constellation is seriously blurred, causing an unacceptable error rate.



(a) (b) (c)

Figure 2.9 16-QAM constellation for a 48-subcarrier OFDM link with a two-ray multipath channel, the second ray being 6 dB lower than the first one. (a) delay < guard time; (b) delay exceeds guard time by 3% of the FFT interval; (c) delay exceeds guard time by 10% of the FFT interval.

2.4 WINDOWING

In the previous sections, it was explained how an OFDM symbol is formed by performing an IFFT and adding a cyclic extension. Looking at an example OFDM signal like in Figure 2.8, sharp phase transitions caused by the modulation can be seen at the symbol boundaries. Essentially, an OFDM signal like the one depicted in Figure 2.8 consists of a number of unfiltered QAM subcarriers. As a result, the out-of-band spectrum decreases rather slowly, according to a sinc function. As an example of this, the spectra for 16, 64, and 256 subcarriers are plotted in Figure 2.10. For larger number of subcarriers, the spectrum goes down more rapidly in the beginning, which is caused by the fact that the sidelobes are closer together. However, even the spectrum for 256 subcarriers has a relatively large –40-dB bandwidth that is almost four times the –3-dB bandwidth.

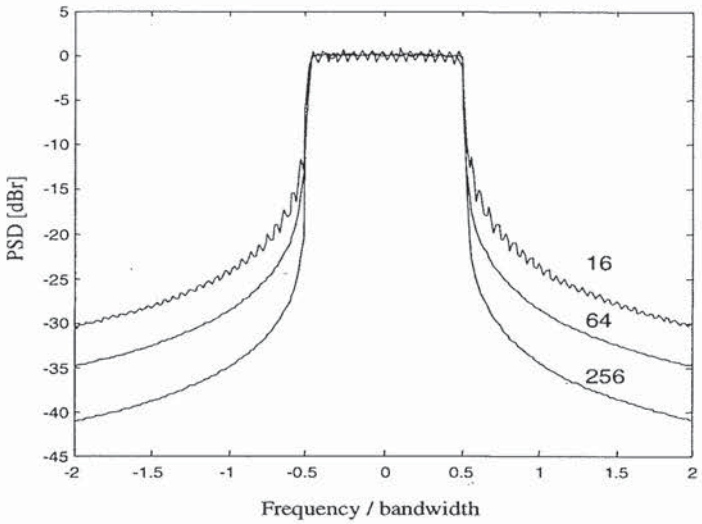


Figure 2.10 Power spectral density (PSD) without windowing for 16, 64, and 256 subcarriers.

To make the spectrum go down more rapidly, windowing can be applied to the individual OFDM symbols. Windowing an OFDM symbol makes the amplitude go smoothly to zero at the symbol boundaries. A commonly used window type is the raised cosine window, which is defined as

$$w(t)=\begin{cases} 0.5+0.5\cos\left(\pi+t\pi/(\beta T_s)\right) & 0 \le t \le \beta T_s \\ 1.0 & \beta T_s \le t \le T_s \\ 0.5+0.5\cos\left((t-T_s)\pi/(\beta T_s)\right) & T_s \le t \le (1+\beta)T_s \end{cases} \quad (2.6)$$

Here, T_s is the symbol interval, which is shorter than the total symbol duration because we allow adjacent symbols to partially overlap in the roll-off region. The time structure of the OFDM signal now looks like Figure 2.11.

In equation form, an OFDM symbol starting at time $t = t_s = kT_s$ is defined as

$$s_k(t)=\mathrm{Re}\left\{ w(t-t_s)\sum_{i=-\frac{N_s}{2}}^{\frac{N_s}{2}-1} d_{i+N_s(k+1/2)}\exp\left(j2\pi\left(f_c-\frac{i+0.5}{T}\right)(t-t_s-T_{prefix})\right)\right\},\ t_s \le t \le t_s+T_s(1+\beta)$$

$$s_k(t)=0,\ t<t_s \ \wedge\ t>t_s+T_s(1+\beta) \quad (2.7)$$

In practice, the OFDM signal is generated as follows: first, N_c input QAM values are padded with zeros to get N input samples that are used to calculate an IFFT. Then, the last T_{prefix} samples of the IFFT output are inserted at the start of the OFDM symbol, and the first $T_{postfix}$ samples are appended at the end. The OFDM symbol is then multiplied by a raised cosine window $w(t)$ to more quickly reduce the power of out-of-band subcarriers. The OFDM symbol is then added to the output of the previous OFDM symbol with a delay of T_s, such that there is an overlap region of βT_s, where β is the rolloff factor of the raised cosine window.

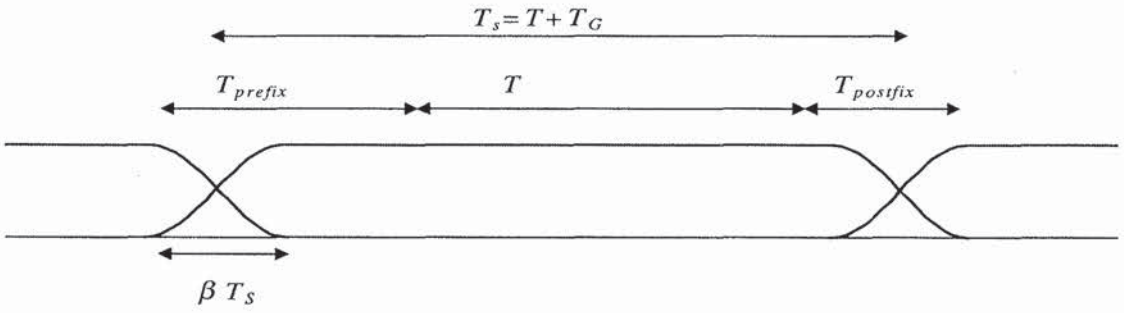


Figure 2.11 OFDM cyclic extension and windowing. T_s is the symbol time, T the FFT interval, T_G the guard time, T_{prefix} the preguard interval, $T_{postfix}$ the postguard interval, and β is the rolloff factor.

Figure 2.12 shows spectra for 64 subcarriers and different values of the rolloff factor β. It can be seen that a rolloff factor of 0.025—so the rolloff region is only 2.5% of the symbol interval—already makes a large improvement in the out-of-band spectrum. For instance, the –40-dB bandwidth is more than halved to about twice the –3-dB bandwidth. Larger rolloff factors improve the spectrum further, at the cost, however, of a decreased delay spread tolerance. The latter effect is demonstrated in Figure 2.13, which shows the signal structure of an OFDM signal for a two-ray multipath channel. The receiver demodulates the subcarriers by taking an FFT over the T-second interval between the dotted lines. Although the relative delay between the two multipath signals is smaller than the guard time, ICI and ISI are introduced because of the amplitude modulation in the gray part of the delayed OFDM symbol. The orthogonality between subcarriers as proved by (2.3) only holds when amplitude and phase of the subcarriers are constant during the entire T-second interval. Hence, a rolloff factor of β reduces the effective guard time by βT_s.

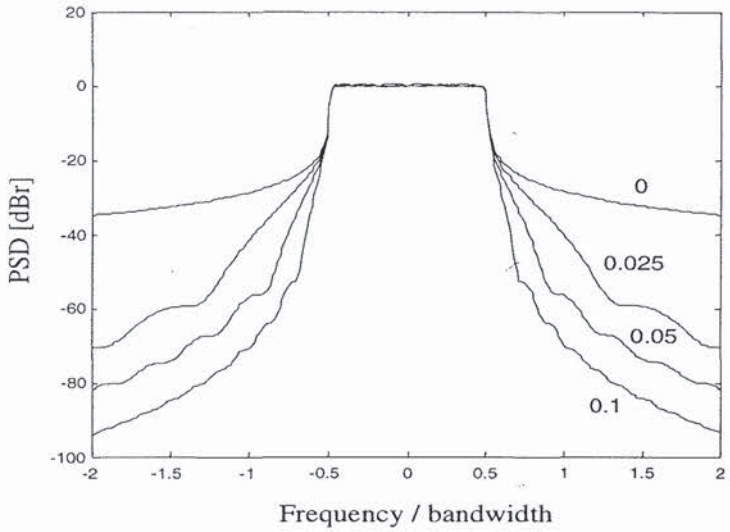


Figure 2.12 Spectra for raised cosine windowing with rolloff factors of 0 (rectangular window), 0.025, 0.05, and 0.1.

Instead of windowing, it is also possible to use conventional filtering techniques to reduce the out-of-band spectrum. Windowing and filtering are dual techniques; multiplying an OFDM symbol by a window means the spectrum is going to be a convolution of the spectrum of the window function with a set of impulses at the subcarrier frequencies. When filtering is applied, a convolution is done in the time domain and the OFDM spectrum is multiplied by the frequency response of the filter. When using filters, care has to be taken not to introduce rippling effects on the envelope of the OFDM symbols over a timespan that is larger than the rolloff region of the windowing approach. Too much rippling means the undistorted part of the OFDM envelope is smaller, and this directly translates into less delay spread tolerance. Notice that digital filtering techniques are more complex to implement than windowing. A digital filter requires at least a few multiplications per sample, while windowing only requires a few multiplications per symbol, for those samples which fall into the rolloff region. Hence, because only a few percent of the samples are in the rolloff region, windowing is an order of magnitude less complex than digital filtering.

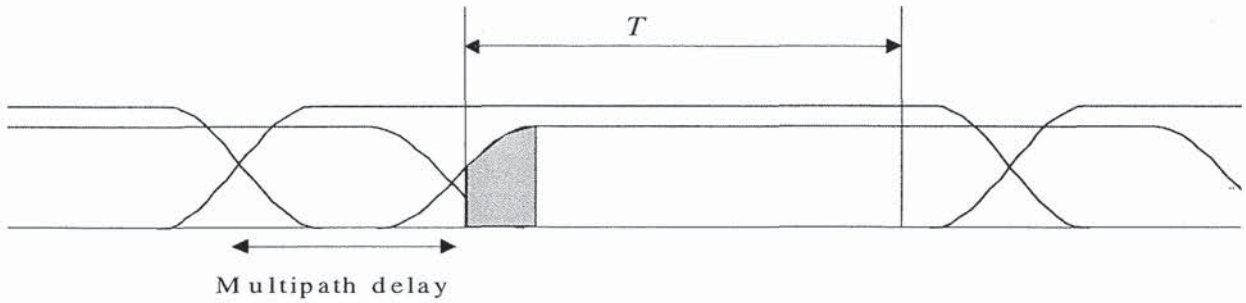


Figure 2.13 OFDM symbol windows for a two-ray multipath channel, showing ICI and ISI, because in the gray part, the amplitude of the delayed subcarriers is not constant.

2.5 CHOICE OF OFDM PARAMETERS

The choice of various OFDM parameters is a tradeoff between various, often conflicting requirements. Usually, there are three main requirements to start with: bandwidth, bit rate, and delay spread. The delay spread directly dictates the guard time. As a rule, the guard time should be about two to four times the root-mean-squared delay spread. This value depends on the type of coding and QAM modulation. Higher order QAM (like 64-QAM) is more sensitive to ICI and ISI than QPSK; while heavier coding obviously reduces the sensitivity to such interference.

Now that the guard time has been set, the symbol duration can be fixed. To minimize the signal-to-noise ratio (SNR) loss caused by the guard time, it is desirable to have the symbol duration much larger than the guard time. It cannot be arbitrarily large, however, because a larger symbol duration means more subcarriers with a smaller subcarrier spacing, a larger implementation complexity, and more sensitivity to phase noise and frequency offset [2], as well as an increased peak-to-average power ratio [3, 4]. Hence, a practical design choice is to make the symbol duration at least five times the guard time, which implies a 1-dB SNR loss because of the guard time.

After the symbol duration and guard time are fixed, the number of subcarriers follows directly as the required –3-dB bandwidth divided by the subcarrier spacing, which is the inverse of the symbol duration less the guard time. Alternatively, the number of subcarriers may be determined by the required bit rate divided by the bit rate per subcarrier. The bit rate per subcarrier is defined by the modulation type (e.g., 16-QAM), coding rate, and symbol rate.

As an example, suppose we want to design a system with the following requirements:

- Bit rate: 20 Mbps
- Tolerable delay spread: 200 ns
- Bandwidth: < 15 MHz

The delay-spread requirement of 200 ns suggests that 800 ns is a safe value for the guard time. By choosing the OFDM symbol duration 6 times the guard time (4.8 μs), the guard time loss is made smaller than 1 dB. The subcarrier spacing is now the inverse of 4.8 – 0.8 = 4 μs, which gives 250 kHz. To determine the number of subcarriers needed, we can look at the ratio of the required bit rate and the OFDM symbol rate. To achieve 20 Mbps, each OFDM symbol has to carry 96 bits of information (96/4.8 μs = 20 Mbps). To do this, there are several options. One is to use 16-QAM together with rate ½ coding to get 2 bits per symbol per subcarrier. In this case, 48 subcarriers are needed to get the required 96 bits per symbol. Another option is to use QPSK with rate ¾ coding, which gives 1.5 bits per symbol per subcarrier. In this case, 64 subcarriers are needed to reach the 96 bits per symbol. However, 64 subcarriers means a bandwidth of 64·250 kHz = 16 MHz, which is larger than the target bandwidth. To achieve a bandwidth smaller than 15 MHz, the number of subcarriers needs to be smaller than 60. Hence, the first option with 48 subcarriers and 16-QAM fulfills all the requirements. It has the additional advantage that an efficient 64-point radix-4 FFT/IFFT can be used, leaving 16 zero subcarriers to provide oversampling necessary to avoid aliasing.

An additional requirement that can affect the chosen parameters is the demand for an integer number of samples both within the FFT/IFFT interval and in the symbol interval. For instance, in the above example we want to have exactly 64 samples in the FFT/IFFT interval to preserve orthogonality among the subcarriers. This can be achieved by making the sampling rate 64/4 μs = 16 MHz. However, for that particular sampling rate, there is no integer number of samples with the symbol interval of 4.8 μs. The only solution to this problem is to change one of the parameters slightly to meet the integer constraint. For instance, the number of samples per symbol can be set to 78, which gives a sampling rate of 78/4.8 μs = 16.25 MHz. Now, the FFT interval becomes 64/16.25 MHz = 3.9385 μs, so both guard time and subcarrier spacing are slightly larger than in the case of the original FFT interval of 4 μs.

2.6 OFDM SIGNAL PROCESSING

The previous sections described how the basic OFDM signal is formed using the IFFT, adding a cyclic extension and performing windowing to get a steeper spectral rolloff. However, there is more to it to build a complete OFDM modem. Figure 2.14 shows the block diagram of an OFDM modem, where the upper path is the transmitter chain and the lower path corresponds to the receiver chain. In the center we see the IFFT, which modulates a block of input QAM values onto a number of subcarriers. In the receiver, the subcarriers are demodulated by an FFT, which performs the reverse operation of an IFFT. An interesting feature of the FFT/IFFT is that the FFT is almost identical to an IFFT. In fact, an IFFT can be made using an FFT by conjugating input and output of the FFT and dividing the output by the FFT size. This makes it possible to use the same hardware for both transmitter and receiver. Of course, this saving in complexity is only possible when the modem does not have to transmit and receive simultaneously.

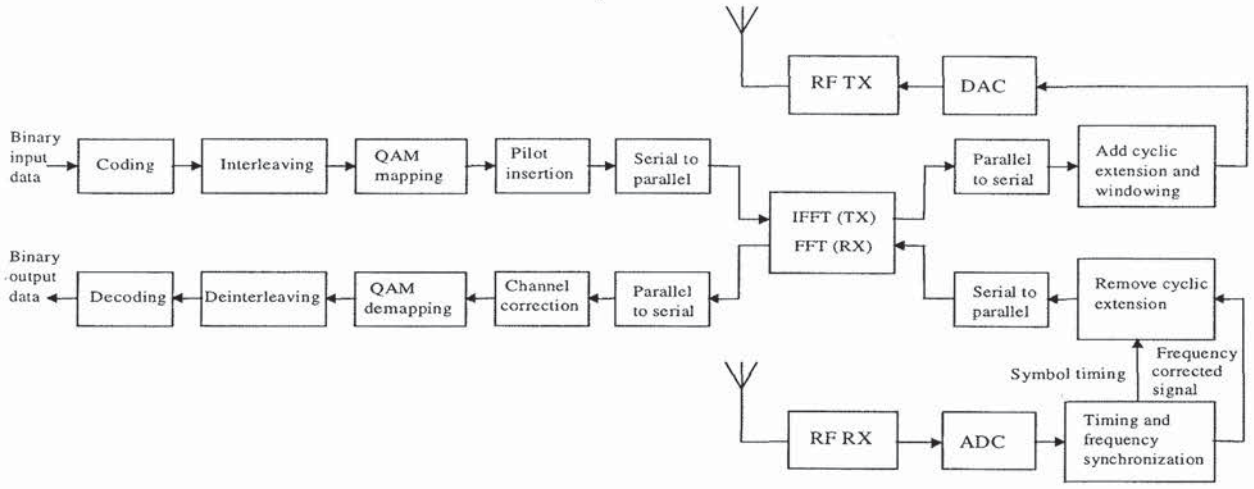


Figure 2.14 Block diagram of an OFDM transceiver.

The functions before the IFFT have not been discussed till now. Binary input data is first encoded by a forward error correction code. The encoded data is then interleaved and mapped onto QAM values. These functions are discussed in more detail in the next chapter.

In the receiver path, after passing the RF part and the analog-to-digital conversion, the digital signal processing starts with a training phase to determine symbol timing and frequency offset. An FFT is used to demodulate all subcarriers. The output of the FFT contains N_s QAM values, which are mapped onto binary values and decoded to produce binary output data. To successfully map the QAM values onto binary values, first the reference phases and amplitudes of all subcarriers have to be acquired. Alternatively, differential techniques can be applied. All of these receiver functions are highlighted in the following chapters.

2.7 IMPLEMENTATION COMPLEXITY OF OFDM VERSUS SINGLE-CARRIER MODULATION

One of the main reasons to use OFDM is its ability to deal with large delay spreads with a reasonable implementation complexity. In a single-carrier system, the implementation complexity is dominated by equalization, which is necessary when the delay spread is larger than about 10% of the symbol duration. OFDM does not require an equalizer. Instead, the complexity of an OFDM system is largely determined by the FFT, which is used to demodulate the various subcarriers. In the following example, it is demonstrated that the processing complexity of an OFDM modem can be significantly less than that of a single carrier modem, given the fact that both can deal with the same amount of delay spread. Notice that some papers mention the use of a

single-tap equalizer for OFDM, when they are referring to the phase correction in case of a coherent OFDM receiver. We will not use this somewhat confusing terminology, because the term equalization suggests that an attempt is made to invert the channel, while an OFDM receiver is doing the opposite; weak subcarriers are not being extra amplified to equalize the channel, but rather, they get a low weight in the decoding. In this way, OFDM avoids the problem of noise enhancement that is present in linear equalization techniques.

Figure 2.15 shows the block diagram of a decision feedback equalizer with symbol-spaced taps. Such an equalizer can be used to combat delay spread in single-carrier systems using quadrature amplitude modulation (QAM) or some form of constant amplitude modulation such as Gaussian Minimum Shift Keying (GMSK) or offset-QPSK. From references [5, 6], it can be learned that at least 8 feedforward and 8 feedback taps are required to handle a delay spread of 100 ns for a GMSK modem at a data rate of 24 Mbps. We can use this information to compare a single carrier system with the 24 Mbps mode of the new IEEE 802.11 OFDM standard, which can handle delay spreads up to 250 ns using a 64-point FFT. In order to increase the delay spread tolerance of a GMSK modem to the same level, the equalizer length has to be increased by a factor of 250/100, so it has 20 feedforward and feedback taps. Fortunately, for GMSK only the real outputs of the complex multiplications are used, so each multiplier has to perform two real multiplications per sample. Hence, the number of real multiplications per second becomes $2 \cdot 20 \cdot 24 \cdot 10^6 = 960 \cdot 10^6$. Notice we only count the feedforward taps here, since the feedback taps consist of trivial rotations, while the feedforward taps consist of full multiplications. For the OFDM system, a 64-point FFT has to be processed by every OFDM symbol duration which is 4 μs, see chapter 10. With a radix-4 algorithm, this requires 96 complex multiplications [1], which gives a processing load of $96 \cdot 10^6$ real multiplications per second. So, in terms of multiplications per second, the equalizer of the single carrier system is 10 times more complex than the FFT of the OFDM system! This complexity difference grows with the bit rate, or rather with the bandwidth – delay spread product, which is a measure of the relative amount of inter-symbol interference. For instance, if we want to double the bit rate in the previous example by doubling the bandwidth, but the delay spread tolerance has to stay the same, then the number of subcarriers and the guard time has to be doubled for OFDM, while for the single-carrier system, both the number of equalizer taps and the sampling rate are doubled. The latter means that the number of multiplications per second is quadrupled, so the equalizer complexity grows quadratically with the bandwidth – delay spread product. For OFDM, a double-sized FFT has to be calculated in the same amount of time in order to double the rate. For the radix-2 algorithm, this means an increase in the number of multiplications of $N\log_2(2N)/ (N/2)\log_2 N = 2(1+1/\log_2 N)$, where N is the FFT size for the half-rate system. The complexity of the FFT grows only slightly faster than linear with the bandwidth – delay spread product, which explains why OFDM is more attractive than a single carrier system with equalization for relatively large bandwidth – delay spread products (values around 1 or larger). It should be noted that the complexity difference between the FFT and the equalizer is less if the equalization is done in the frequency domain, as

described in [7]. In this case, equalization is twice as complex, because both an FFT and an IFFT have to be performed to do frequency domain equalization on a signal block.

Another complexity advantage of OFDM is the fact that the FFT does not really require full multiplications, but rather phase rotations, which can be efficiently implemented by the CORDIC algorithm [8]. Because phase rotations do not change the amplitude, they do not increase the dynamic range of the signals, which simplifies the fixed point design.

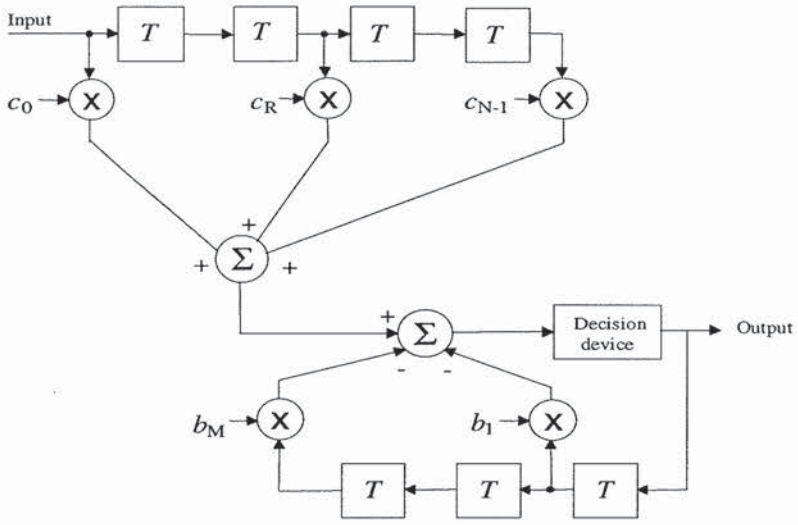


Figure 2.15 Decision-feedback equalizer.

Except for the difference in complexity, OFDM has another advantage over single carrier systems with equalizers. For the latter systems, the performance degrades abruptly if the delay spread exceeds the value for which the equalizer is designed. Because of error propagation, the raw bit error probability increases so quickly that inroducing lower rate coding or a lower constellation size does not significantly improve the delay spread robustness. For OFDM, however, there are no such nonlinear effects as error propagation, and coding and lower constellation sizes can be employed to provide fallback rates that are significantly more robust against delay spread. This is an important consideration, as it enhances the coverage area and avoids the situation that users in bad spots cannot get any connection at all.

REFERENCES

[1] Blahut, R. E., *Fast Algorithms for Digital Signal Processing*. Reading, MA: Addison-Wesley, 1985.

[2] Pollet, T., M. van Bladel and M. Moeneclaey, "BER Sensitivity of OFDM Systems to Carrier Frequency Offset and Wiener Phase Noise," *IEEE Trans. on Comm.*, Vol. 43, No. 2/3/4, pp. 191–193, Feb.–Apr. 1995.

[3] Pauli, M., and H. P. Kuchenbecker, "Minimization of the Intermodulation Distortion of a Nonlinearly Amplified OFDM Signal," *Wireless Personal Communications*, Vol. 4, No. 1, pp. 93–101, Jan. 1997.

[4] Rapp, C., "Effects of HPA-Nonlinearity on a 4-DPSK/OFDM Signal for a Digital Sound Broadcasting System," *Proc. of the Second European Conference on Satellite Communications*, Liège, Belgium, pp.179–184, Oct. 22–24, 1991.

[5] Tellado-Mourelo, J., E. K.Wesel, J. M.Cioffi, "Adaptive DFE for GMSK in Indoor Radio Channels," *IEEE Trans. on Sel. Areas in Comm.*, Vol. 14, No. 3, pp. 492–501, Apr. 1996.

[6] Wales, S. W., "Modulation and Equalization Techniques for HIPERLAN," *Proc. of PIMRC/WCN*, The Hague, The Netherlands, Sept. 21–23, pp. 959–963, 1994.

[7] Sari, H., G. Karam, I. Jeanclaude, "Transmission Techniques for Digital Terrestrial TV Broadcasting," *IEEE Communications Magazine*, Feb. 1995, pp. 100–109.

[8] Parhi, K. K., and T. Nishitani, *Digital Signal Processing for Multimedia Systems*. New York: Marcel Dekker, Inc., 1999.

Chapter 3

Coding and Modulation

3.1 INTRODUCTION

We explained in the previous chapter how OFDM avoids the problem of intersymbol interference by transmitting a number of narrowband subcarriers together with using a guard time. This does give rise to another problem, however, which is the fact that in a multipath fading channel, all subcarriers will arrive at the receiver with different amplitudes. In fact, some subcarriers may be completely lost because of deep fades. Hence, even though most subcarriers may be detected without errors, the overall bit-error ratio (BER) will be largely dominated by a few subcarriers with the smallest amplitudes, for which the bit-error probability is close to 0.5. To avoid this domination by the weakest subcarriers, forward-error correction coding is essential. By using coding across the subcarriers, errors of weak subcarriers can be corrected up to a certain limit that depends on the code and the channel. A powerful coding means that the performance of an OFDM link is determined by the average received power, rather than by the power of the weakest subcarrier.

This chapter starts with a review of block codes and convolutional codes. Then, it introduces interleaving as a way to randomize error bursts that occur when adjacent subcarriers are lost in a deep fade. Section 3.4 describes QAM as the most commonly used modulation technique in OFDM, after which Section 3.5 shows the important relation between coding and modulation. Thus, this chapter presents a brief overview of coding and modulation [1–9].

ZYXEL-1015
Page 068 (of 277)

3.2 FORWARD-ERROR CORRECTION CODING

3.2.1 Block Codes

A *block code* encodes a block of k input symbols into n coded symbols, with n being larger than k. The purpose of adding the redundant n-k symbols is to increase the *minimum Hamming distance*, which is the minimum number of different symbols between any pair of code words. For a minimum Hamming distance of d_{min}, the code can correct t errors where t is given by

$$t \leq floor(\frac{d_{\min} - 1}{2}) \tag{3.1}$$

Here, *floor(x)* denotes the floor function that rounds x downward to the closest integer value. The minimum Hamming distance is upperbound by the number of redundant symbols n-k as

$$d_{\min} \leq n - k + 1 \tag{3.2}$$

For binary codes, only repetition codes and single-parity check codes reach this upperbound. A class of nonbinary codes that does reach the above bound are the *Reed-Solomon codes*. Because of their good distance properties and the availabity of efficient coding and decoding algorithms [6, 7], Reed-Solomon codes are the most popularly used block codes. Reed-Solomon codes are defined for blocks of symbols with m bits per symbol, where the code length n is related to m by

$$n = 2^m - 1 \tag{3.3}$$

The number of input symbols k is related to m and the required minimum Hamming distance d_{min} as

$$k = 2^m - d_{\min} \tag{3.4}$$

There appears to be little flexibility in the available code lengths as indicated by (3.3). However, a Reed-Solomon code can easily be shortened to any arbitrary length

by leaving a number of input bits zero and deleting the same amount of output bits. It is also possible to extend the code length to a power of 2 by adding an extra parity symbol.

According to (3.1) and (3.2), a Reed-Solomon code can correct up to floor((n-k)/2) erroneous symbols. Each symbol contains m bits, so a maximum amount of m·floor((n-k)/2) erroneous bits may be corrected. The latter is only true, however, if all bit errors occur within the maximum amount of correctable symbol errors. So if a Reed-Solomon is designed to correct up to two symbol errors containing 8 bits per symbol, it cannot correct an arbitrary combination of three bit errors, as these errors may occur in three different symbols. This characteristic makes Reed-Solomon codes particularly useful for correcting bursty channels. One example of such a channel is an OFDM link in the presence of multipath fading, which causes the errors to be concentrated in a few subcarriers that are hit by deep fades.

3.2.2 Convolutional Codes

A convolutional code maps each k bits of a continuous input stream on n output bits, where the mapping is performed by convolving the input bits with a binary impulse response. The convolutional encoding can be implemented by simple shift registers and modulo-2 adders. As an example, Figure 3.1 shows the encoder for a rate 1/2 code which is actually one of the most frequently applied convolutional codes. This encoder has a single data input and two outputs A_i and B_i, which are interleaved to form the coded output sequence $\{A_1B_1A_2B_2 \ldots\}$. Each pair of output bits $\{A_i,B_i\}$ depends on seven input bits, being the current input bit plus six previous input bits that are stored in the length 6 shift register. This value of 7—or in general the shift register length plus 1—is called the *constraint length*. The shift register taps are often specified by the corresponding generator polynomials or generator vectors. For the example of Figure 3.1, the generator vectors are {1011011,1111001} or {133,171} octal. The ones in the generator vectors correspond with taps on the shift register.

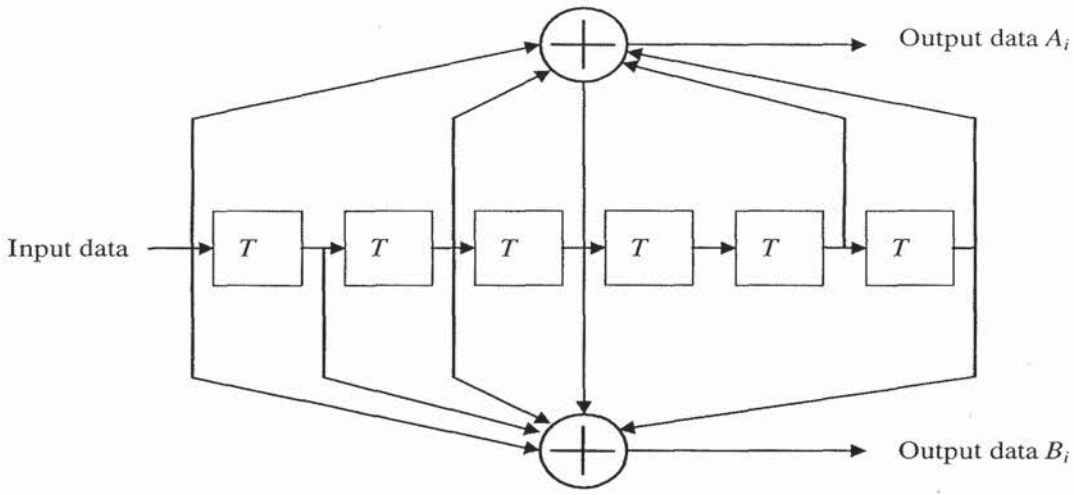


Figure 3.1 Block diagram of a constraint length 7 convolutional encoder.

Decoding of convolutional codes is most often performed by *soft decision Viterbi decoding*, which is an efficient way to obtain the optimal maximum likelihood estimate of the encoded sequence. A description of this decoding technique can be found in [8]. The complexity of Viterbi decoding grows exponentially with the constraint length. Hence, practical implementations do not go further than a constraint length of about 10. Decoding of convolutional codes with larger constraint length is possible by using suboptimal decoding techniques like sequential decoding [8].

Because convolutional codes do not have a fixed length, it is more difficult to specify their performance in terms of Hamming distance and a number of correctable errors. One measure that is used is the *free distance*, which is the minimum Hamming distance between arbitrarily long different code sequences that begin and end with the same state of the encoder, where the state is defined by the contents of the shift registers of the encoder. For example, the code of Figure 3.1 has a free distance of 10. When hard decision decoding is used, this code can correct up to floor((10-1)/2) = 4 bit errors within each group of encoded bits with a length of about 3 to 5 times the constraint length. When soft decision decoding is used, however, the number of correctable errors does not really give a useful measure anymore. A better performance measure is the *coding gain*, which is defined as the gain in the bit energy-to-noise density ratio E_b/N_o relative to an uncoded system to achieve a certain bit error ratio. The E_b/N_o gain is equivalent to the gain in input signal-to-noise ratio (SNR) minus the rate loss in dB because of the redundant bits. As an example, Figure 3.2 shows the bit error ratio versus E_b/N_o for uncoded QPSK and for coded QPSK using the previously mentioned constraint length 7 code. It can be seen from the figure that for a bit-error ratio of 10^{-4}, the coded link needs about 5 dB less E_b/N_o compared with that of the uncoded link. For lower bit error ratios, this coding gain converges to a maximum value of about 5.5 dB.

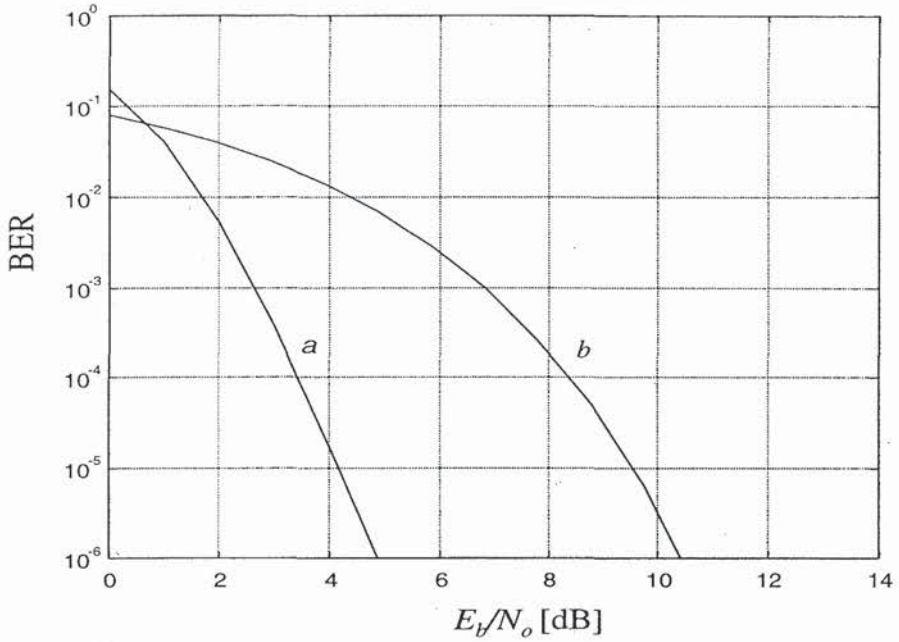


Figure 3.2 BER versus E_b/N_o for coded and uncoded QPSK in AWGN.

In the above curves, we have used the bit energy-to-noise density ratio E_b/N_o, which is equivalent to the ratio of the signal power P and the noise power in a bandwidth equal to the bit rate N_o/T_b, where T_b is the bit time. Some other useful SNR definitions are the input SNR and the symbol energy-to-noise density ratio E_s/N_o, which is equivalent to the ratio of the signal power and the noise power within a bandwidth equal to the symbol rate N_o/T_s, where T_s is the symbol duration. E_s/N_o is equivalent to the SNR for individual subcarriers, plus the guard time loss in dB. It is related to E_b/N_o as

$$\frac{E_s}{N_o} = \frac{E_b}{N_o}\frac{T_s}{T_b} \tag{3.5}$$

The input signal-to-noise ratio SNR_i is related to E_b/N_o as

$$SNR_i = \frac{E_b}{N_o}\frac{1}{BT_b} = \frac{E_b}{N_o}\frac{bN_s r}{BT_s} \tag{3.6}$$

Here, B is the input noise bandwidth, b is the number of coded bits per subcarrier, N_s the number of subcarriers, and r the coding rate. Basically, the SNR_i is equal to E_b/N_o multiplied by the ratio of bit rate and bandwidth. The latter ratio is equivalent to the spectral efficiency in bps/Hz. The spectral efficiency depends on the number of bits per subcarrier b, which is determined by the constellation size, the coding rate r, and the guard time, which appears indirectly in (3.6) as a part of the

symbol duration T_s. The number of subcarriers has no influence on the spectral efficiency, because the noise bandwidth increases linearly with the number of subcarriers.

A convolutional code can be punctured to increase the coding rate. For instance, increasing the rate of the above rate 1/2 code to 3/4 is done by deleting 2 of every 6 bits at the output of the encoder. The punctured output sequence for a rate 3/4 code is $\{A_1B_1A_2B_3A_4B_4A_5B_6A_7B_7 \ldots\}$. For a rate 2/3 code, the punctured output sequence is $\{A_1B_1A_2A_3B_3A_4A_5B_5 \ldots\}$. To decode the punctured sequence, the original rate 1/2 decoder can be used. Before decoding, erasures have to be inserted in the data at the locations of the punctured bits.

3.2.3 Concatenated Codes

Instead of using a single block code or convolutional code, it is also possible to combine or concatenate two codes. The main advantage of a concatenated code is that it can provide a large coding gain with less implementation complexity as a comparable single code. Figure 3.3 shows the block diagram of a concatenated coding scheme. The input bits are first coded and interleaved by an outer coder and interleaver. The coded bits are then again coded and interleaved by an inner coder and interleaver. Usually, the inner code is a convolutional code and the outer code a block code; for instance, a Reed-Solomon code. The motivation behind this is that the convolutional code with soft decision decoding performs better for relatively low-input SNRs. The hard decision block decoder then cleans up the relatively few remaining errors in the decoded output bits of the convolutional decoder. The task of the interleavers is to break up bursts of errors as much as possible. In case of an outer block code, the outer interleaver preferably separates symbols by more than the block length of the outer encoder. Compared with a single-code system, concatenated coding has more delay because of the extra interleaving, which can be a disadvantage for packet communications where the interleaving delay affects turnaround time and throughput. A good overview of achievable performance of concatenated coding is given in [3].

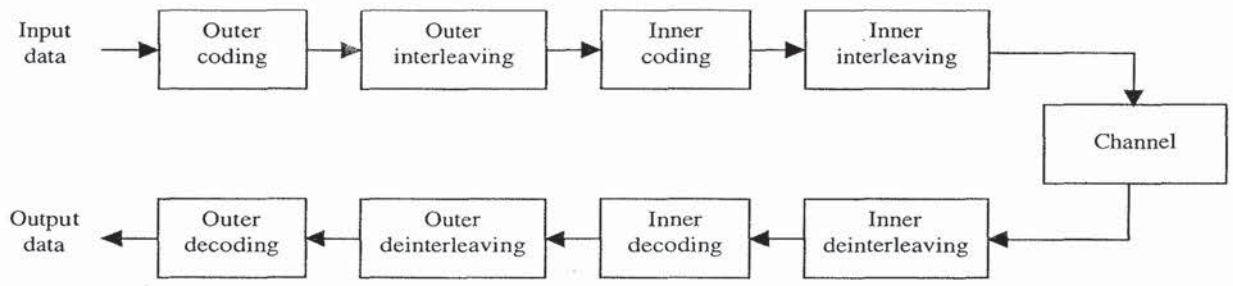


Figure 3.3 Concatenated coding/decoding.

3.3 INTERLEAVING

Because of the frequency selective fading of typical radio channels, the OFDM subcarriers generally have different amplitudes. Deep fades in the frequency spectrum may cause groups of subcarriers to be less reliable than others, thereby causing bit errors to occur in bursts rather than being randomly scattered. Most forward error-correction codes are not designed to deal with error bursts. Therefore, interleaving is applied to randomize the occurrence of bit errors prior to decoding. At the transmitter, the coded bits are permuted in a certain way, which makes sure that adjacent bits are separated by several bits after interleaving. At the receiver, the reverse permutation is performed before decoding. A commonly used interleaving scheme is the block interleaver, where input bits are written in a matrix column by column and read out row by row. As an example of such an interleaver, Figure 3.4 shows the bit numbers of a block interleaver operating on a block size of 48 bits. After writing the 48 bits in the matrix according to the order as depicted in the figure, the interleaved bits are read out row by row, so the output bit numbers are 0, 8, 16, 24, 32, 40, 1, 9, . . . , 47. Instead of bits, the operation can also be applied on symbols; for instance, the matrix can be filled with 48 16-QAM symbols containing 4 bits per symbol, so the interleaving changes the symbol order but not the bit order within each symbol. Interleaving on a symbol-basis is especially useful for Reed-Solomon codes, as these codes operate on symbols rather than bits. A Reed-Solomon code can correct up to a certain number of symbol errors per block length, so interleaving should be done over several block lengths, in order to spread bursts of symbol errors over a number of different Reed-Solomon blocks.

0	8	16	24	32	40
1	9	17	25	33	41
2	10	18	26	34	42
3	11	19	27	35	43
4	12	20	28	36	44
5	13	21	29	37	45
6	14	22	30	38	46
7	15	23	31	39	47

Figure 3.4 Interleaving scheme.

For a general block interleaver with a block size of N_B bits and d columns, the ith interleaved bit is equal to the kth encoded input bit, where k is given by

$$k = id - (N_B - 1)\,floor(\frac{id}{N_B}) \tag{3.7}$$

Instead of a block interleaver, it is also possible to use a convolutional interleaver. An example of this type of interleaver is shown in Figure 3.5. The interleaver cyclically writes each input symbol or bit into one of K shift registers that introduce a delay of 0 to k-1 symbol durations. The shift registers are read out cyclically to produce the interleaved symbols.

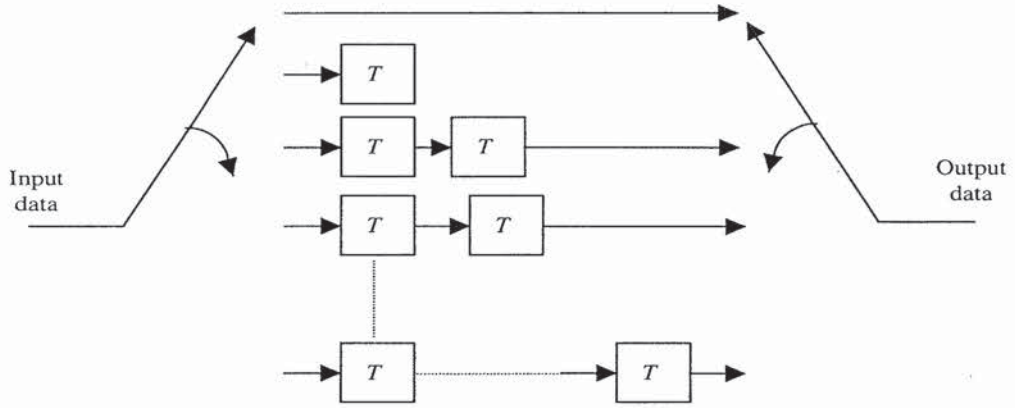


Figure 3.5 Convolutional interleaver.

3.4 QUADRATURE AMPLITUDE MODULATION

Quadrature amplitude modulation (QAM) is the most popular type of modulation in combination with OFDM. Especially rectangular constellations are easy to implement as they can be split into independent *Pulse amplitude modulated* (PAM) components for both the in-phase and the quadrature part. Figure 3.6 shows the rectangular constellations of Quadrature Phase Shift Keying (QPSK), 16-QAM, and 64-QAM. The constellations are not normalized; to normalize them to an average power of one—assuming that all constellation points are equally likely to occur—each constellation has to be multiplied by the normalization factor listed in Table 3.1. The table also mentions Binary Phase Shift Keying (BPSK), which uses two of the four QPSK constellation points (1+j, −1−j).

Table 3.1 also gives the loss in the minimum squared Euclidean distance between two constellation points, divided by the gain in data rate of that particular QAM type relative to BPSK. This value defines the maximum loss in E_b/N_o relative to BPSK that is needed to achieve a certain bit-error ratio in an uncoded QAM link. The BER curves in Figure 3.7 illustrate that the E_b/N_o loss values of Table 3.1 are quite accurate for BER values below 10^{-2}. The difference between QPSK and 16-QAM is about 4 dB. From 16-QAM to 64-QAM, almost 4.5 dB extra E_b/N_o is required. For larger constellation sizes, the E_b/N_o penalty of increasing the number of bits per symbol by 1 converges to 3 dB.

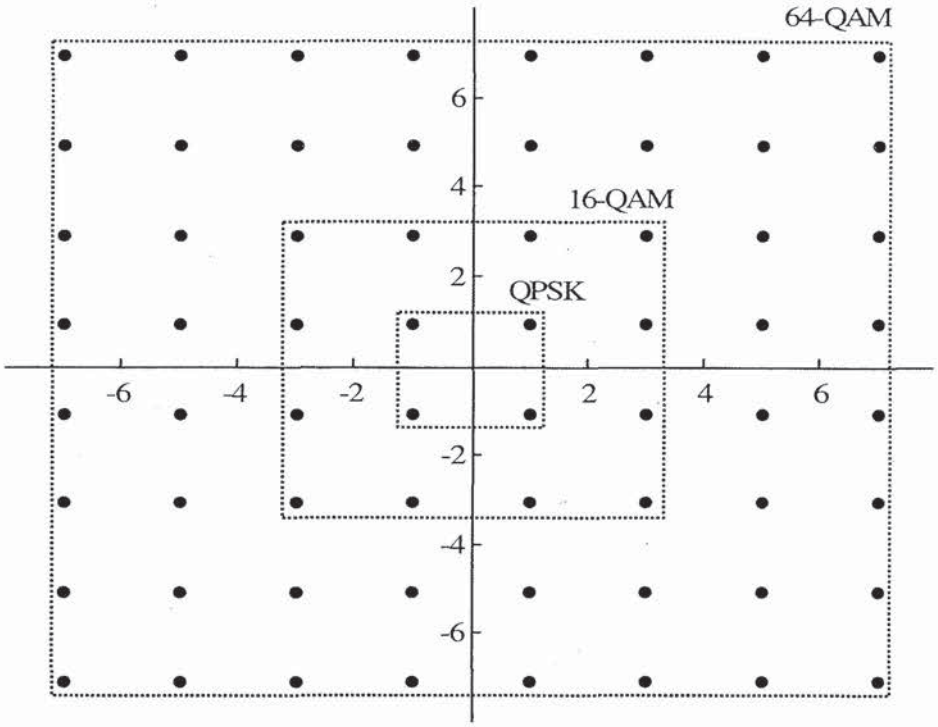


Figure 3.6 QPSK, 16-QAM, and 64-QAM constellations.

Table 3.1
QAM normalization factors and normalized Euclidean distance differences.

Modulation	Normalization factor	Maximum E_b/N_o loss relative to BPSK in dB
BPSK	$1/\sqrt{2}$	0
QPSK	$1/\sqrt{2}$	0
16-QAM	$1/\sqrt{10}$	3.98
64-QAM	$1/\sqrt{42}$	8.45

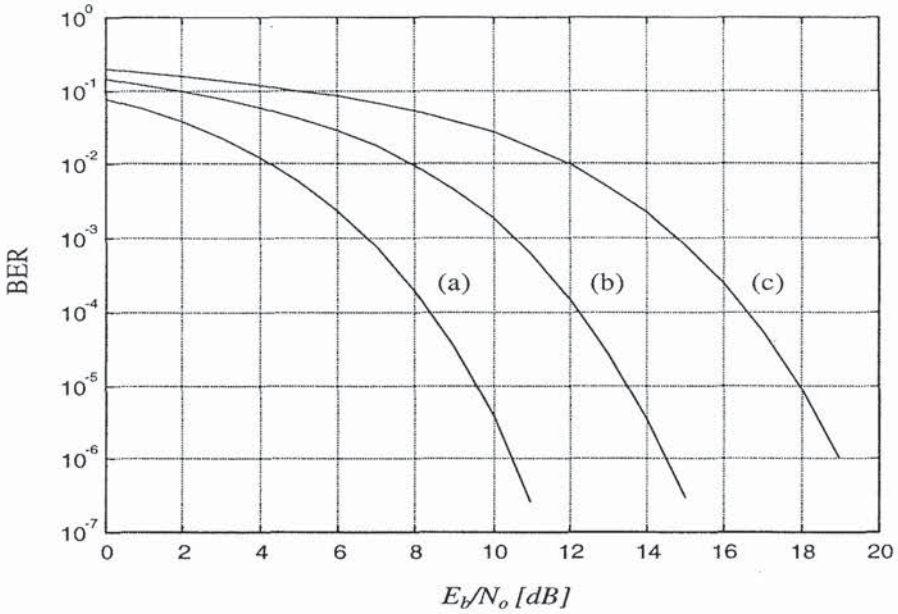


Figure 3.7 BER versus E_b/N_o for (a) BPSK/QPSK, (b) 16-QAM, (c) 64-QAM.

3.5 CODED MODULATION

When coding is applied to a QAM signal, it is important to consider the relation between coding and modulation to obtain the best result. In [9], Trellis coding was introduced as a way to attain coding gain without bandwidth expansion, meaning that all redundancy is obtained by increasing the constellation size. It was demonstrated in [9] that coding gains up to 6 dB can be obtained by going from uncoded QPSK to Trellis-coded 8-PSK using rate 2/3 coding. This technique is based on partitioning the PSK or QAM constellations into subsets with a high Euclidean distance within each subset. For instance, each 2 8-PSK constellation points with a relative phase difference of 180° define a subset with the same Euclidean distance as BPSK. The minimum Euclidean distance between different subsets is much smaller. Hence, coding is applied on the bits defining the subset number, such that the minimum Euclidean distance of the coded signal becomes equal to the distance within each subset.

A disadvantage of the above Trellis coding approach is that although the codes can have a large minimum Euclidean distance, the minimum Hamming distance is only 1, because bits within a subset are left uncoded. Hence, if one Trellis-coded symbol is lost, this immediately results in one or more bit errors. For OFDM this is a very undesirable property, as the data of several subcarriers may be lost by deep fades. This illustrates that the minimum Euclidean distance is not the only relevant parameter when selecting a good code for OFDM. For frequency selective channels, an additional criterion is that the Euclidean distance should be spread over as many symbols as possible, such that a few lost symbols have the smallest possible impact on the

probability of a decoding error [1]. As a consequence, in fading channels it is preferable to use high-order QAM constellations in combination with low-rate coding schemes. It is demonstrated in [1] that a specially designed rate 1/4 code with constraint length 7 together with a 16-QAM constellation can give good performance even on channels where more than half of the subcarriers are lost, with a degradation of less than 2 dB in the required signal-to-noise ratio relative to an ideal AWGN channel.

One of the disadvantages of special Trellis codes is that they are designed for specific constellations, such that you need a different encoder and decoder for different constellations. A practical approach to avoid this problem is to use standard binary codes together with Gray-encoded QAM. In [4], for instance, an efficient way is described to use a standard binary convolutional code together with 16-QAM. This scheme can easily be extended to an arbitrary rectangular QAM constellation. To do this, binary input data are converted into QAM symbols according to a Gray code mapping. For 16-QAM, for instance, the in-phase and quadrature parts are separately formed as 4 level PAM values, determined by two bits $b0$ and $b1$, as shown in Figure 3.8. The vertical lines indicate the regions for which the bit values are 1.

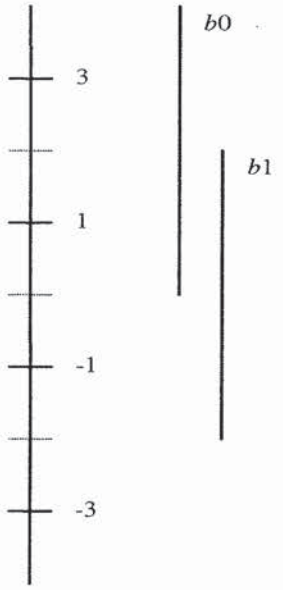


Figure 3.8 Gray mapping of two bits into 4 level PAM.

In the receiver, the incoming QAM symbols have to be demapped into one-dimensional values with corresponding metrics for the Viterbi decoder. For QPSK, the demapping is simply taking the in-phase and quadrature values as the two desired metrics. For the case of 16-QAM, the in-phase and quadrature values are treated as independent 4 level PAM signals, which are demapped into 2 metrics as shown in Figure 3.9. Here, the input values are normalized such that 2 corresponds to a decision

level, above which $b1$ is zero. The scale of the output values depends on the required quantization level of the Viterbi decoder, which typically ranges from 3 to 8 bits. Notice that the metric for b0 can be twice as large as for b1 in case the in-phase or quadrature value is –3 or +3. This indicates the fact that the b0 values are indeed more reliable in this case, as the minimum Euclidean distance from a +3 value to an erroneous $b0$ value of –1 is 4, while the minimum distance between two different $b1$ values is always 2. This difference in reliability of bits becomes even larger for higher order QAM. For instance, Figure 3.10 shows the demapping of 3-bit metrics for the case of 8 level PAM, which is used to make a 64-QAM constellation.

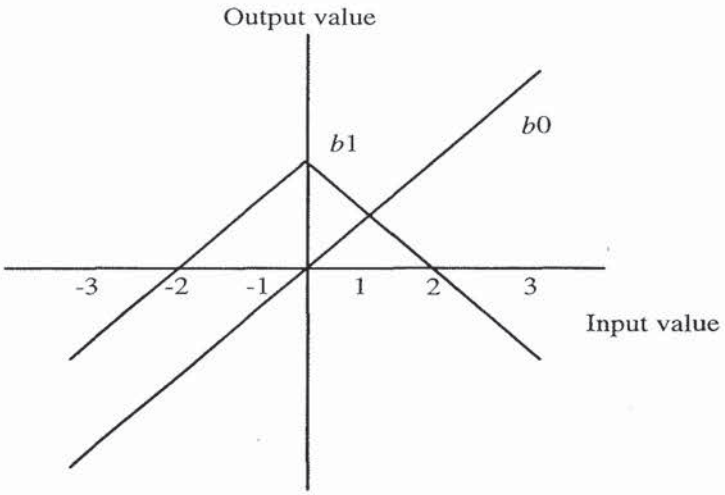


Figure 3.9 Demapping of 4 level PAM into 2 metrics.

Figure 3.11 shows the BERs versus E_b/N_o for several combinations of coding rates and QAM types in an AWGN channel. An interesting effect is that the coding gain relative to the uncoded QAM curves of Figure 3.7 becomes larger for larger QAM constellations. At a BER of 10^{-5}, for instance, the coding gains for a rate 1/2 code are approximately 5.5, 7, and 8.5 dB for QPSK, 16-QAM, and 64-QAM, respectively. This is explained by the fact that the uncoded error probability is determined mainly by the least significant bits in the Gray mapping; for instance, the b3 values in Figure 3.10. When coding is used, then the error probability depends on an average over several coded bits; for instance, several b1, b2, and b3 values in the case of 64-QAM. As a result of this averaging, the minimum squared Euclidean distance between different coded QAM sequences is larger than in the situation where only the least significant bits like b3 in Figure 3.10 are transmitted. Therefore, the coding gain is larger than in the case of QPSK or BPSK, where all bits have the same weight.

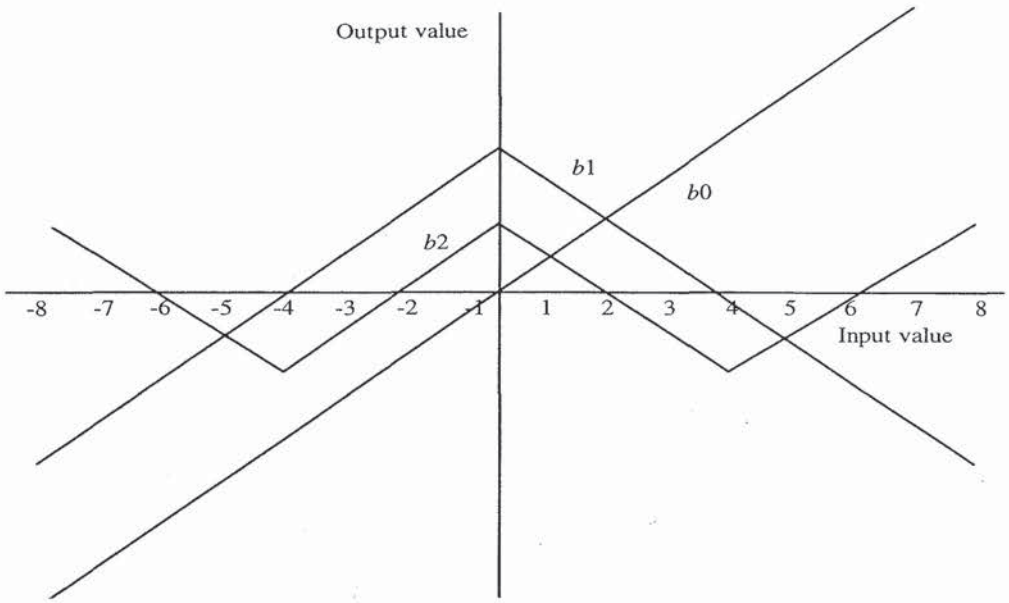


Figure 3.10 Demapping of 8 level PAM into 3 metrics.

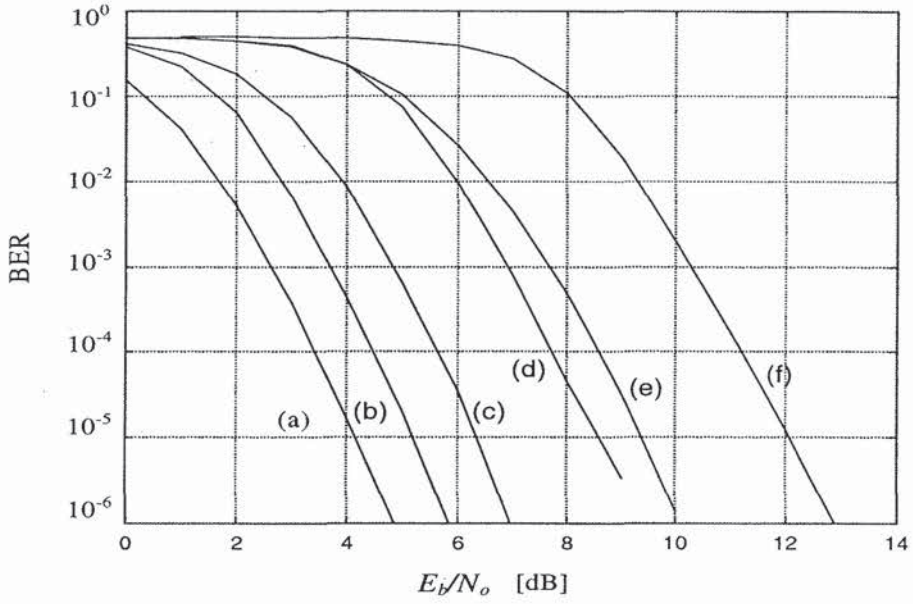


Figure 3.11 BERs versus E_b/N_o in AWGN for a constraint length 7 convolutional code with (a) QPSK, rate ½, (b) QPSK, rate ¾; (c) 16-QAM, rate ½; (d) 16-QAM, rate ¾; (e) 64-QAM, rate ½; (f) 64-QAM, rate ¾.

Other interesting information that can be derived from Figure 3.11 is the coding gain of a particular coded modulation type when compared with the uncoded QAM constellation that gives the same net data rate or the same efficiency in bps/Hz. For instance, the curve for 16-QAM with rate 1/2 coding can be compared with the curve for uncoded QPSK in Figure 3.7, as both have the same efficiency of 2 bps/Hz. It can be seen from the figures that for a BER of 10^{-5}, coded 16-QAM gives a coding gain of about 3 dB, compared with that of uncoded QPSK.

Figure 3.12 shows simulated BERs versus mean E_b/N_o for a Rayleigh fading channel with an exponentially decaying power delay profile. This channel was introduced in Chapter 1. Curves are drawn for various normalized delay spreads $\tau_{rms}N_s/T$, which is the rms delay spread as defined in Chapter 1 multiplied by the bandwidth of the OFDM signal. The normalized delay spread makes it possible to generalize delay spread results, independent of the number of subcarriers or the absolute bandwidth value of an OFDM system. It is required though that the number of subcarriers be significantly larger (a factor of 4 is sufficient) than the constraint length of the convolutional code, such that the code is able to fully benefit from the frequency diversity of the channel. The fact that the performance of an OFDM link depends only on the normalized delay spread $\tau_{rms}N_s/T$ can be understood better by realizing that the rms delay spread is approximately equal to the inverse of the coherence bandwidth of the channel, which determines the characteristics of bandwidth and spacing of fades in the channel frequency response. A small normalized delay spread is equivalent to a small ratio of OFDM signal bandwidth and coherence bandwidth. In such a situation, the channel frequency response is relatively flat within the OFDM signal bandwidth, so if there is a deep fade, all the subcarriers are significantly attenuated. In the case of a large normalized delay spread, a fade only affects a few adjacent subcarriers. There can be several fades within the OFDM signal bandwidth, with relatively strong subcarriers between the fades. As a result, the average signal power is much more constant over several channels than in the case of small delay spreads. The coding benefits from this by using the stronger subcarriers to compensate for the attenuated subcarriers.

Except for the delay spread, the guard time T_G is also normalized in Figure 3.12 and all other figures in this chapter. The normalized guard time is defined here as T_GN_s/T; the same normalization as for the guard time is applied to maintain a fixed ratio between delay spread and guard time, independent of the number of subcarriers N_s and the FFT interval T. Because the E_b/N_o loss caused by the guard time depends on the ratio T_G/T_s rather than the normalization chosen here, the guard time loss is not included in the following figures. This makes it possible to see from Figure 3.12 for what ratio of guard time and delay spread the system breaks down because of ISI and ICI. In Figure 3.12, for instance, an error floor is starting to appear for a normalized delay spread of 4; the ratio of guard time and delay spread is 3 for this case. Hence, for QPSK with rate 3/4 coding, the guard time should be at least three times the delay spread to achieve an average BER less than 10^{-4}.

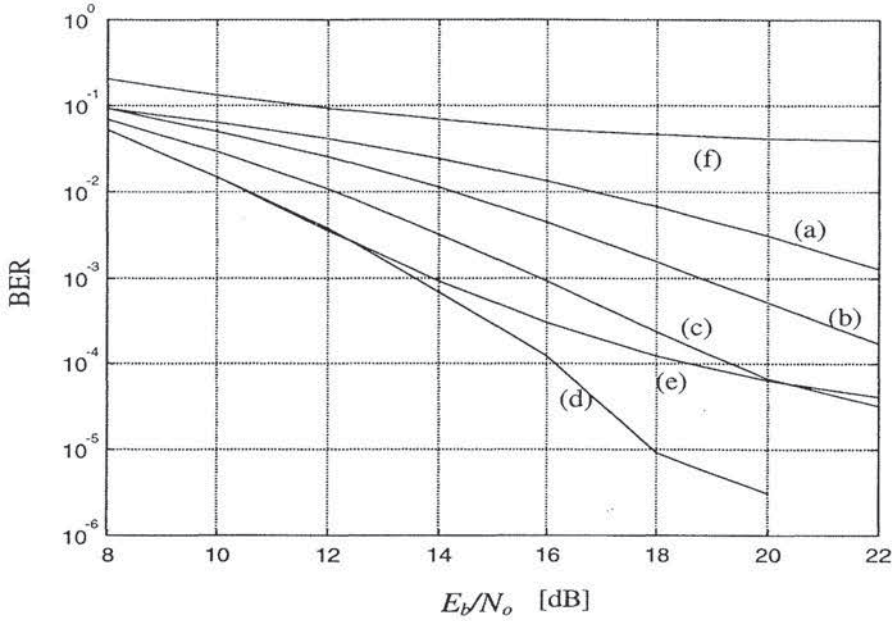


Figure 3.12 BERs versus mean E_b/N_o in a Rayleigh fading channel for QPSK, rate 3/4 convolutional coding, normalized guard time $T_G N_s/T = 12$, normalized delay spread $\tau_{rms} N_s/T$ = (a) 0.25, (b) 0.5, (c) 1, (d) 2, (e) 4, (f) 8.

Notice that Figure 3.12 uses the mean E_b/N_o, which is the average value over a large number of independent channels. The instantaneous E_b/N_o of an individual channel can be significantly smaller or larger than this mean value, especially for low delay spreads where the instantaneous signal power is determined by a single Rayleigh fading path. For larger delay spreads, the variation in the instantaneous signal power becomes much smaller because of the increased frequency diversity of the channel. Low instantaneous E_b/N_o values, which dominate the error ratio, occur much less frequently than at low delay spreads, hence the improved performance. The larger the delay spread, the smaller the E_b/N_o can be, until the delay spread becomes so large that ISI and ICI become limiting factors.

Figure 3.13 shows packet-error ratios (PERs) for 256-byte packets, simulated for the same conditions as Figure 3.12. For relatively slowly time-varying channels, as encountered for instance in indoor wireless LAN applications, the packet-error ratio averaged over a large number of fading channels is equivalent to the coverage outage probability, which is the probability of an unacceptably large packet-error ratio at a certain location within the coverage range. For instance, at an E_b/N_o value of 18 dB and a normalized delay spread between 0.5 and 4, a packet error ratio of 1% means that 1% of the channels generate most of the packet errors—caused by deep fades or ISI/ICI—while the remaining 99% of the channels have a much lower error ratio.

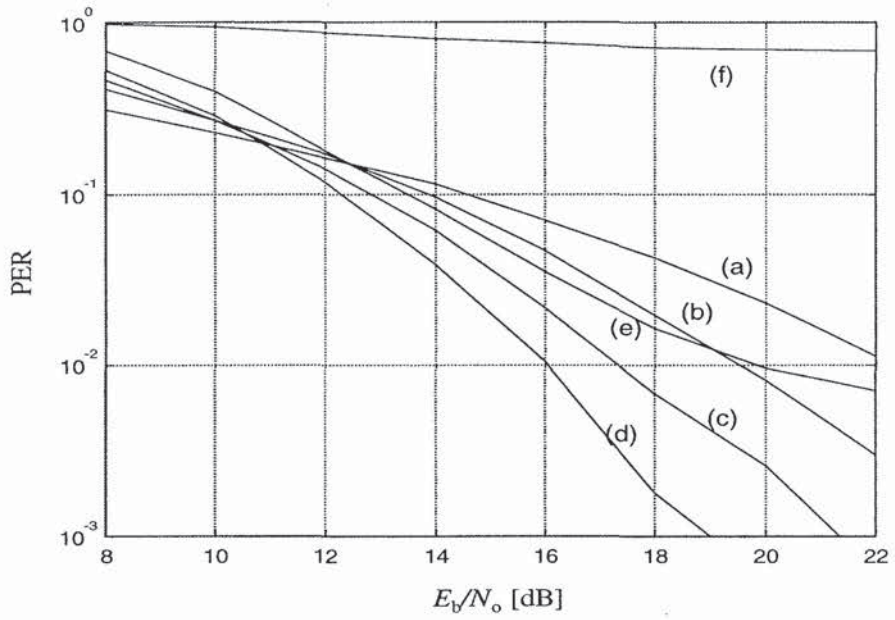


Figure 3.13 PERs versus mean E_b/N_o in a Rayleigh fading channel for QPSK, rate 3/4 convolutional coding, packet length = 256 bytes, normalized guard time $T_G N/T$ = 12, normalized delay spread $\tau_{rms} N_s/T$ = (a) 0.25, (b) 0.5, (c) 1, (d) 2, (e) 4, (f) 8.

Figure 3.14 shows simulated BERs and PERs for 256-byte packets in a Rayleigh fading channel with an exponentially decaying power delay profile. Two different combinations of coding rate and QAM type are used. Curves (a) and (c) are based on 16-QAM with rate 1/2 coding, giving a spectral efficiency of 2 bps/Hz, while curves (b) and (d) use QPSK with rate 3/4 coding, giving an efficiency of 1.5 bps/Hz. This example leads to the surprising result that in a fading channel, a higher order QAM system with a better spectral efficiency can be actually better in terms of E_b/N_o performance than a system with a lower spectral efficiency based on a lower order QAM, while the opposite is true in AWGN as demonstrated by Figure 3.11. The explanation for this effect is that in a frequency selective channel where a certain percentage of the subcarriers can be completely lost in deep fades, the ability to correct for those lost subcarriers by having a large Hamming distance is more important than a large minimum Euclidean distance for each individual subcarrier. In the example of Figure 3.11, the rate 1/2 code combined with 16-QAM can tolerate more weak subcarriers than the rate 3/4 code with QPSK, resulting in an E_b/N_o gain that is larger than the loss in Euclidean distance of 16-QAM versus QPSK.

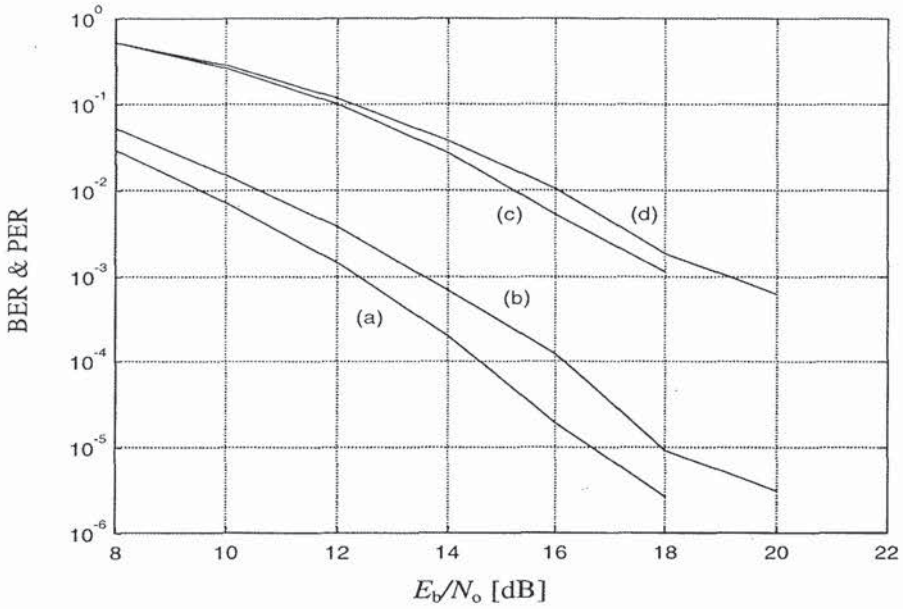


Figure 3.14 PER and BER versus mean E_b/N_o in a Rayleigh fading channel for a packet length of 256 bytes and a normalized delay spread $\tau_{rms}N_s/T = 2$. (a) BER of 16-QAM, rate 1/2 coding; (b) BER of QPSK, rate 3/4 coding; (c) PER of 16-QAM, rate 1/2 coding; (d) PER of QPSK, rate 3/4 coding.

As mentioned before, when the delay spread increases, the performance of an OFDM link increases until a limit is reached where ISI and ICI cause an unacceptably high error floor. This error floor depends on the type of modulation and coding rate. Figure 3.15 illustrates this with simulation curves of the packet error floor versus the normalized delay spread τ_{rms}/T_G. No noise was present in the simulations, so all errors are purely caused by ISI and ICI. As expected, more delay spread can be tolerated for smaller QAM constellations. There is little difference, however, in the robustness of, for instance QPSK with rate 3/4 coding and 16-QAM with rate 1/2 coding, thanks to the fact that the latter is able to tolerate more erroneous subcarriers, which partly compensates the smaller distance between constellation points.

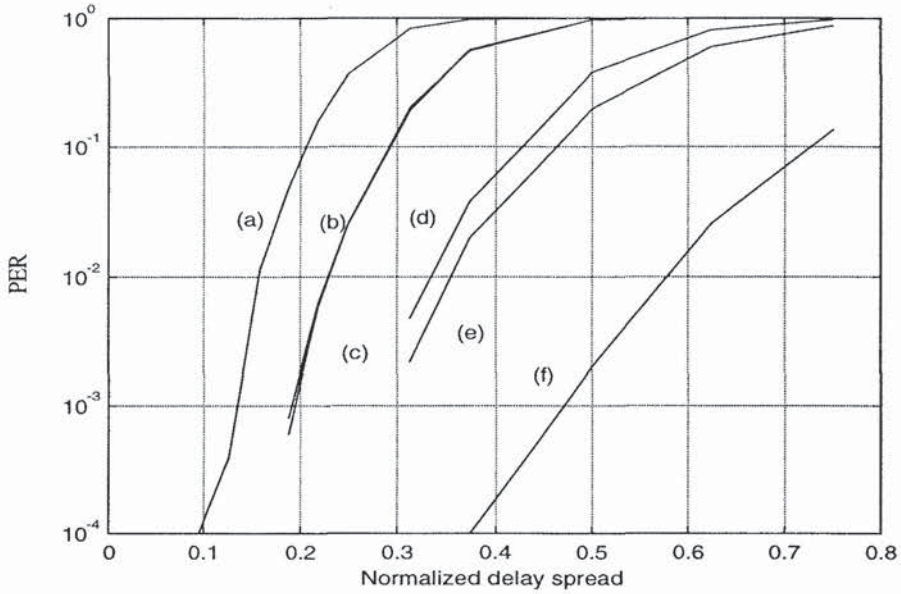


Figure 3.15 Irreducible packet error ratios versus normalized delay spread τ_{rms}/T_G for 256 byte packets and (a) 64-QAM, rate ¾; (b) 64-QAM, rate ½; (c) 16-QAM, rate ¾; (d) 16-QAM, rate ½; (e) QPSK, rate ¾; (f) QPSK, rate ½.

When an OFDM system has to be designed, Figure 3.15 can be used to derive a minimum requirement on the guard time, based on the maximum delay spread for which the system should work. For instance, for a tolerable packet-error floor of 1%, the guard time has to be about twice the delay spread for QPSK with rate 1/2 coding, but it has to be six times the delay spread for 64-QAM with rate 3/4 coding.

REFERENCES

[1] Wesel, R. D., "Joint Interleaver and Trellis Code Design," *Proceedings of IEEE Globecom*, 1997.

[2] Le Floch, B., M. Alard, and C. Berrou, "Coded Orthogonal Frequency Division Multiplex," *Proceedings of the IEEE*, Vol. 83, no. 6, June 1995.

[3] Alard, M., and R. Lasalle, "Principles of Modulation and Channel Coding for Digital Broadcasting for Mobile Receivers," *EBU Technical Review*, No. 224, pp. 168 - 190.

[4] Wang, Q., and L. Y. Onotera, "Coded QAM Using a Binary Convolutional Code," *IEEE Transactions on Communications*, Vol. 43, No. 6, June 1995.

[5] Wesel, R. D., and J. M. Cioffi, "Fundamentals of Coding for Broadcast OFDM," *Proceedings of IEEE ASILOMAR-29*, 1996.

[6] Massey, J. L., "Shift-Register Synthesis and BCH Decoding," *IEEE Transactions on Information Theory*, IT-15, pp. 122 - 127, Jan. 1979.

[7] Berlekamp, E. R., "The Technology of Error Correcting Codes," *Proceedings of the IEEE*, Vol.68, No.5, pp.564 – 593, May 1980.

[8] Charles Lee, L. H., "*Convolutional Coding: Fundamentals and Applications*," London: Artech House, 1997.

[9] Ungerboeck, G., "Channel Coding with Multilevel/Phase Signals," *IEEE Transactions on Information Theory*, Vol. IT-28, No.1, pp.55-67, Jan. 1982.

CHAPTER 4

Synchronization

Before an OFDM receiver can demodulate the subcarriers, it has to perform at least two synchronization tasks. First, it has to find out where the symbol boundaries are and what the optimal timing instants are to minimize the effects of intercarrier interference (ICI) and intersymbol interference (ISI). Second, it has to estimate and correct for the carrier frequency offset of the received signal, because any offset introduces ICI. In this chapter, we discuss the effects of timing and frequency offsets and describe several synchronization techniques that can be used to obtain symbol timing and frequency synchronization. Notice that these two synchronization tasks are not the only training required in an OFDM receiver. For coherent receivers, except for the frequency, the carrier phase also needs to be synchronized. Further, a coherent QAM receiver needs to learn the amplitudes and phases of all subcarriers to find out the decision boundaries for the QAM constellation of each subcarrier. The latter training tasks are described in the next chapter.

4.1 INTRODUCTION

In an OFDM link, the subcarriers are perfectly orthogonal only if transmitter and receiver use exactly the same frequencies. Any frequency offset immediately results in ICI. A related problem is phase noise; a practical oscillator does not produce a carrier at exactly one frequency, but rather a carrier that is phase modulated by random phase jitter. As a result, the frequency, which is the time derivative of the phase, is never perfectly constant, thereby causing ICI in an OFDM receiver. For single-carrier systems, phase noise and frequency offsets only give a degradation in the received signal-to-noise ratio (SNR), rather than introducing interference. This is the reason that the sensitivity to phase noise and frequency offset are often mentioned as disadvantages of OFDM relative to single-carrier systems. Although it is true that OFDM is more susceptible to phase noise and frequency offset than single-carrier systems, the following sections show that degradation can be kept to a minimum. They describe

ZYXEL-1015
Page 088 (of 277)

techniques to achieve symbol timing and frequency synchronization by using the cyclic prefix or special OFDM training symbols. They also demonstrate that OFDM is rather insensitive to timing offsets, although such offsets do reduce the delay spread robustness. An optimal symbol timing technique is derived that maximizes the delay spread robustness.

4.2 SENSITIVITY TO PHASE NOISE

The issue of phase noise in OFDM systems has been the subject of many studies [1–5]. In [1], the power density spectrum of an oscillator signal with phase noise is modeled by a Lorentzian spectrum, which is equal to the squared magnitude of a first order lowpass filter transfer function. The single-sided spectrum $S_s(f)$ is given by

$$S_s(f) = \frac{2/\pi f_l}{1 + f^2 / f_l^2} \tag{4.1}$$

Here, f_l is the –3 dB linewidth of the oscillator signal. In practice, only double-sided spectra are measured, which are equal to mirrored versions of the one-sided spectrum around the carrier frequency f_c. Further, because the bandwidth is doubled, the spectrum is divided by 2 in order to keep the total power normalized to 1. Hence, the double-sided phase noise spectrum is given by

$$S_d(f) = \frac{1/\pi f_l}{1 + |f - f_c|^2 / f_l^2} \tag{4.2}$$

Figure 4.1 shows an example of a Lorentzian phase noise spectrum with a single-sided –3-dB linewidth of 1 Hz. The slope of –20 dB per decade of this model agrees with measurements in [3], which shows measured phase noise spectra for two oscillators at 5 and 54 GHz.

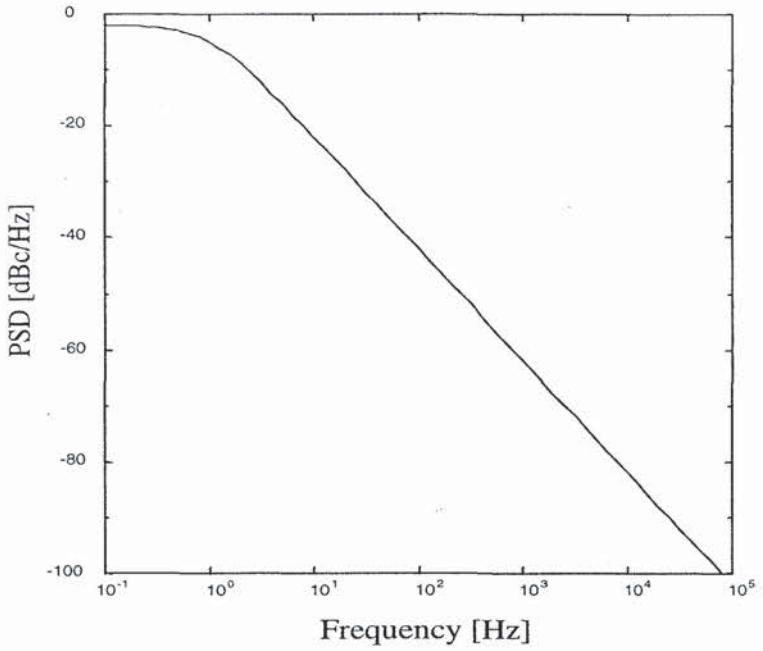


Figure 4.1 Phase noise power spectral density (PSD) with a single-sided –3-dB linewidth of 1 Hz and a –100 dBc/Hz density at 100 kHz offset.

Phase noise basically has two effects. First, it introduces a random phase variation that is common to all subcarriers. If the oscillator linewidth is much smaller than the OFDM symbol rate, which is usually the case, then the common phase error is strongly correlated from symbol to symbol, so tracking techniques or differential detection can be used to minimize the effects of this common phase error. The second and more disturbing effect of phase noise is that it introduces ICI, because the subcarriers are no longer spaced at exactly $1/T$ in the frequency domain. In [1], the amount of ICI is calculated and translated into a degradation in SNR that is given as

$$D_{phase} \cong \frac{11}{6\ln 10} 4\pi\beta T \frac{E_s}{N_o} \tag{4.3}$$

Here, β is the –3-dB one-sided bandwidth of the power density spectrum of the carrier. The phase noise degradation is proportional to βT, which is the ratio of the line-width and subcarrier spacing $1/T$. Figure 4.2 shows the SNR degradation in dB as a function of the normalized linewidth βT. Curves are shown for three different E_s/N_o values, corresponding to the required values to obtain a bit-error ratio (BER) of 10^{-6} for uncoded QPSK, 16-QAM, and 64-QAM, respectively. The main conclusion that we can draw from this figure is that for a negligible SNR degradation of less than 0.1 dB, the –3-dB phase noise bandwidth has to be about 0.1 to 0.01 percent of the subcarrier spacing, depending on the modulation. For instance, to support 64-QAM in an OFDM link with a subcarrier spacing of 300 kHz, the –3-dB linewidth should be 30 Hz at

most. According to (4.2), this means that at a distance of 1 MHz from the carrier frequency, the phase noise spectral density has to have a value of approximately –110 dBc/Hz.

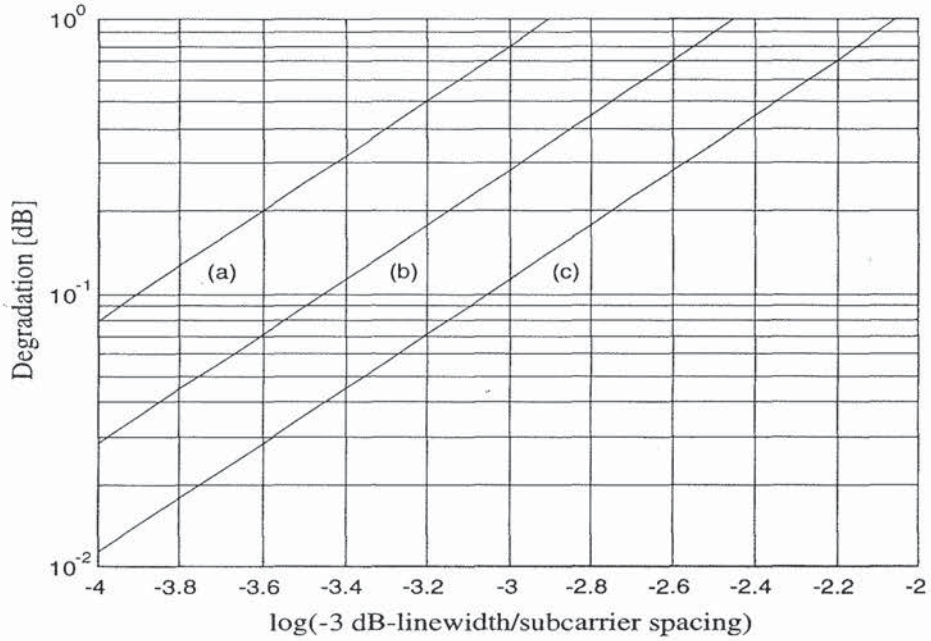


Figure 4.2 SNR degradation in dB versus the –3-dB bandwidth of the phase noise spectrum for (a) 64-QAM (E_s/N_o=19 dB), (b) 16-QAM (E_s/N_o=14.5 dB), (c) QPSK (E_s/N_o=10.5 dB).

The phase noise analysis in [1] assumed a free-running voltage-controlled oscillator (VCO). In practice, however, normally a phase-locked loop (PLL) is used to generate a carrier with a stable frequency. In a PLL, the frequency of a VCO is locked to a stable reference frequency, which is usually produced by a crystal oscillator. The PLL is able to track the phase jitter of the free-running VCO for jitter frequency components that fall within the *tracking loop bandwidth* of the loop. As a result, for frequencies below the tracking loop bandwidth the phase noise of the PLL output is determined mainly by the phase noise of the reference oscillator, which is usually smaller than the VCO phase noise, while for frequencies larger than the tracking loop bandwidth, the phase noise is dominated by the VCO phase noise. In this case, a typical phase noise spectrum will have a shape as depicted in Figure 4.3. The loop bandwidth of this example is around 100 Hz. For such phase noise spectra, the above analysis does not directly apply. We can, however, use the above results to get some requirements for a practical phase noise spectrum. For example, a heuristic approach is to require that the total power in the range of a minimum frequency offset of 10% of the subcarrier spacing to a maximum offset equal to the total bandwidth of the OFDM signal is equal to that of the Lorentzian model. For instance, suppose we have an OFDM system with a subcarrier spacing of 300 kHz and a bandwidth of 20 MHz. For the above-mentioned

example of a 30-Hz linewidth, the total power in the range of 30 kHz to 20 MHz is $(\pi/2)(\text{atan}(2\cdot10^7/30)-\text{atan}(3\cdot10^4/30)) \cong -32$ dBc. In fact, for this case, the exact value of the total bandwidth does not matter much, as the amount of phase noise power for frequency offsets larger than 20 MHz is negligible. The value of –32 dBc means that the total amount of phase noise for frequency offsets larger than 10% of the subcarrier spacing is less than 0.1% of the total carrier power. For a practical PLL, the phase noise spectrum can be measured and integrated over the same frequency interval to check whether the total phase noise power meets the requirement.

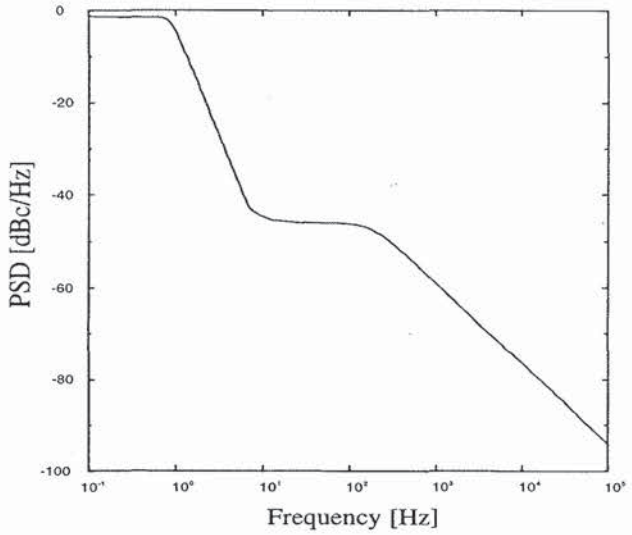


Figure 4.3 Example of a PLL phase noise spectrum.

4.3 SENSITIVITY TO FREQUENCY OFFSET

Chapter 2 explained that all OFDM subcarriers are orthogonal if they all have a different integer number of cycles within the FFT interval. If there is a frequency offset, then the number of cycles in the FFT interval is not an integer anymore, with the result that ICI occurs after the FFT. The FFT output for each subcarrier will contain interfering terms from all other subcarriers, with an interference power that is inversely proportional to the frequency spacing. The amount of ICI for subcarriers in the middle of the OFDM spectrum is approximately twice as large as that for subcarriers at the band edges, because the subcarriers in the middle have interfering subcarriers on both sides, so there are more interferers within a certain frequency distance. In [1], the degradation in SNR caused by a frequency offset that is small relative to the subcarrier spacing is approximated as

$$D_{freq} \cong \frac{10}{3\ln 10}(\pi \Delta f T)^2 \frac{E_s}{N_o} \tag{4.4}$$

This degradation is depicted in Figure 4.4 as a function of the frequency offset, normalized to the subcarrier spacing, and for three different E_s/N_o values. Note that for a negligible degradation of about 0.1 dB, the maximum tolerable frequency offset is less than 1% of the subcarrier spacing. For instance, for an OFDM system at a carrier frequency of 5 GHz and a subcarrier spacing of 300 kHz, the oscillator accuracy needs to be 3 kHz or 0.6 ppm. The initial frequency error of a low-cost oscillator will normally not meet this requirement, which means that a frequency synchronization technique has to be applied before the FFT. Examples of such synchronization techniques are described further in this chapter.

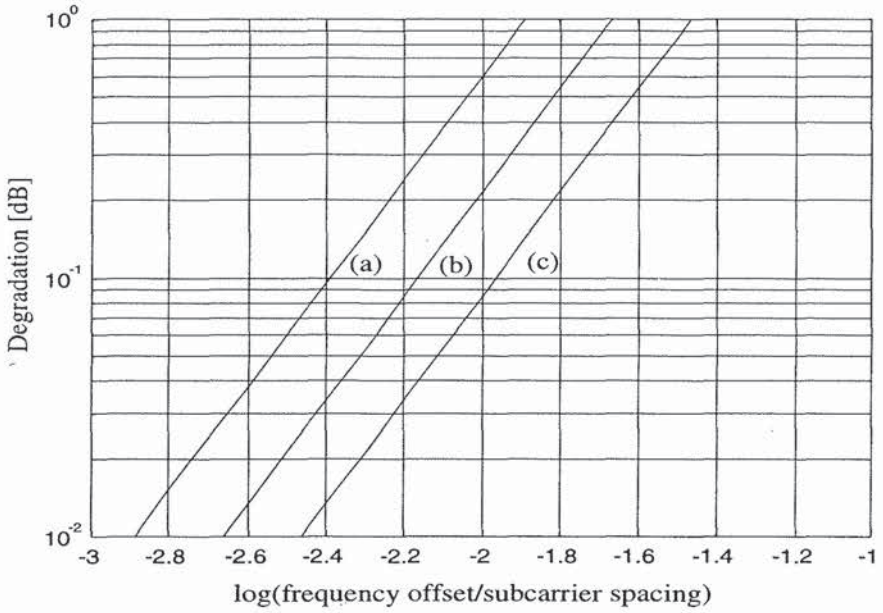


Figure 4.4 SNR degradation in dB versus the normalized frequency offset for (a) 64-QAM (E_s/N_o=19 dB), (b) 16-QAM (E_s/N_o=14.5 dB), (c) QPSK (E_s/N_o=10.5 dB).

4.4 SENSITIVITY TO TIMING ERRORS

The previous section explained that frequency offset and phase jitter introduce a certain amount of ICI. With respect to timing offsets, OFDM is relatively more robust; in fact, the symbol timing offset may vary over an interval equal to the guard time without causing ICI or ISI, as depicted in Figure 4.5. ICI and ISI occur only when the FFT interval extends over a symbol boundary or extends over the rolloff region of a symbol. Hence, OFDM demodulation is quite insensitive to timing offsets. To achieve the best possible multipath robustness, however, there exists an optimal timing instant as explained in Section 4.7. Any deviation from this timing instant means that the sensitivity to delay spread increases, so the system can handle less delay spread than

the value it was designed for. To minimize this loss of robustness, the system should be designed such that the timing error is small compared with the guard interval.

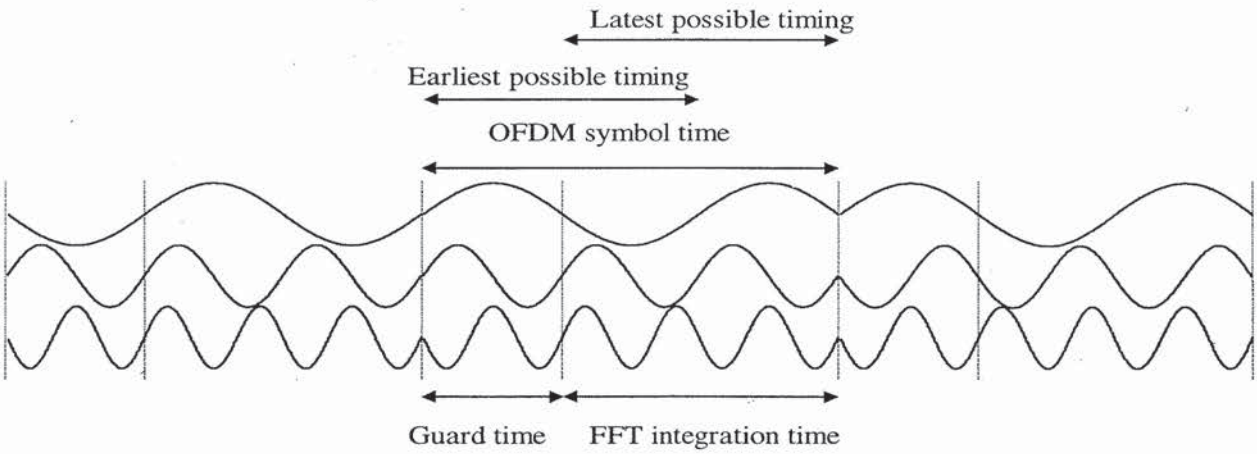
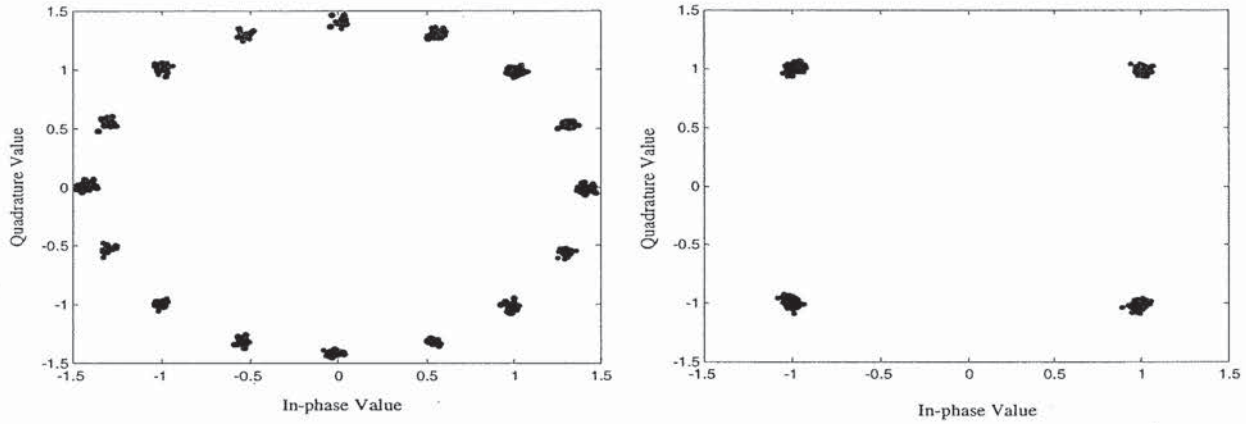


Figure 4.5 Example of an OFDM signal with three subcarriers, showing the earliest and latest possible symbol timing instants that do not cause ISI or ICI.

An interesting relationship exists between symbol timing and the demodulated subcarrier phases [6]. Looking at Figure 4.5, we can see that as the timing changes, the phases of the subcarriers change. The relation between the phase φ_i of subcarrier i and the timing offset τ is given by

$$\varphi_i = 2\pi f_i \tau \tag{4.5}$$

Here, f_i is the frequency of the ith subcarrier before sampling. For an OFDM system with N subcarriers and a subcarrier spacing of $1/T$, a timing delay of one sampling interval of T/N causes a significant phase shift of $2\pi(1\text{-}1/N)$ between the first and last subcarrier. These phase shifts add to any phase shifts that are already present because of multipath propagation. In a coherent OFDM receiver, channel estimation is performed to estimate these phase shifts for all subcarriers, which is described in the next chapter. Figure 4.6(a) shows an example of the QPSK constellation of a received OFDM signal with 48 subcarriers, an SNR of 30 dB, and a timing offset equal to 1/16 of the FFT interval. The timing offset translates into a phase offset of a multiple of $2\pi/16$ between the subcarriers. Because of this phase offset, the QPSK constellation points are rotated to 16 possible points on a circle. After estimation and correction of the phase rotations, the constellation diagram of Figure 4.6(b) is obtained. Chapter 5 describes methods to estimate the unknown phase rotations.



(a) (b)

Figure 4.6 Constellation diagram with a timing error of T/16 before (a) and after (b) phase correction.

In the above analysis, we implicitly assumed that there is an error only in the timing offset and not in the sampling frequency. An error in the sampling frequency has two effects [7]. First, it gives a time-varying timing offset, resulting in time-varying phase changes that have to be tracked by the receiver. Second, it causes ICI because an error in the sampling frequency means an error in the FFT interval duration, such that the sampled subcarriers are not orthogonal anymore. Fortunately, for practically achievable sampling offsets of 10 ppm, the amount of ICI is rather small, about 0.01 dB at an E_s/N_o of 20 dB, as shown in [7].

4.5 SYNCHRONIZATION USING THE CYCLIC EXTENSION

Because of the cyclic prefix, the first T_G seconds part of each OFDM symbol is identical to the last part. This property can be exploited for both timing and frequency synchronization by using a synchronization system like depicted in Figure 4.7. Basically, this device correlates a T_G long part of the signal with a part that is T seconds delayed [8, 9]. The correlator output can be written as

$$x(t) = \int_0^{T_G} r(t-\tau)r(t-\tau-T)d\tau \tag{4.6}$$

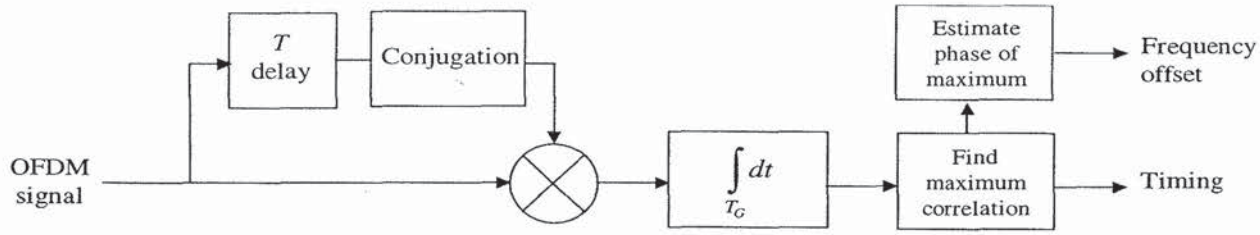


Figure 4.7 Synchronization using the cyclic prefix.

Two examples of the correlation output are shown in Figures 4.8 and 4.9 for eight OFDM symbols with 192 and 48 subcarriers, respectively. These figures illustrate a few interesting characteristics of the cyclic extension correlation method. First, both figures clearly show eight peaks for the eight different symbols, but the peak amplitudes show a significant variation. The reason for this is that although the average power for a T-seconds interval of each OFDM symbol is constant, the power in the guard time can substantially vary from this average power level. Another effect is the level of the undesired correlation sidelobes between the main correlation peaks. These sidelobes reflect the correlation between two pieces of the OFDM signal that belong partly or totally to two different OFDM symbols. Because different OFDM symbols contain independent data values, the correlation output is a random variable, which may reach a value that is larger than the desired correlation peak. The standard deviation of the random correlation magnitude is related to the number of independent samples over which the correlation is performed. The larger the number of independent samples, the smaller the standard deviation is. In the extreme case where the correlation is performed over only one sample, the output magnitude is proportional to the signal power, and there is no distinct correlation peak in this case. In the other extreme case where the correlation is performed over a very large number of samples, the ratio of sidelobes-to-peak amplitude will go to zero. Because the number of independent samples is proportional to the number of subcarriers, the cyclic extension correlation technique is only effective when a large number of subcarriers are used, preferably more than 100. An exception to this is the case where instead of random data symbols, specially designed training symbols are used [10]. In this case, the integration can be done over the entire symbol duration instead of the guard time only. The level of undesired correlation sidelobes can be minimized by a proper selection of the training symbols.

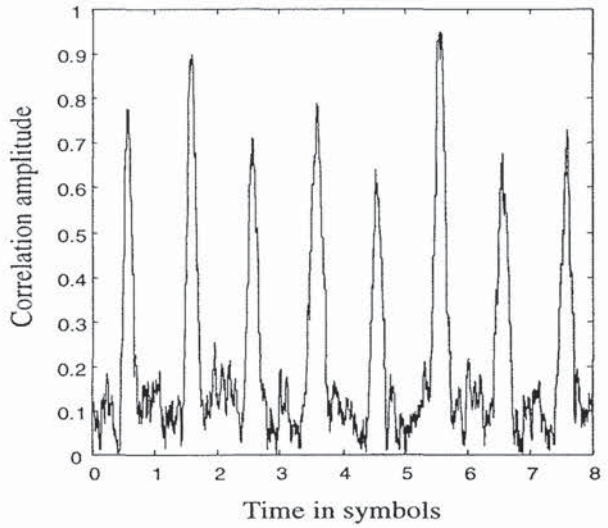


Figure 4.8 Example of correlation output amplitude for eight OFDM symbols with 192 subcarriers and a 20% guard time.

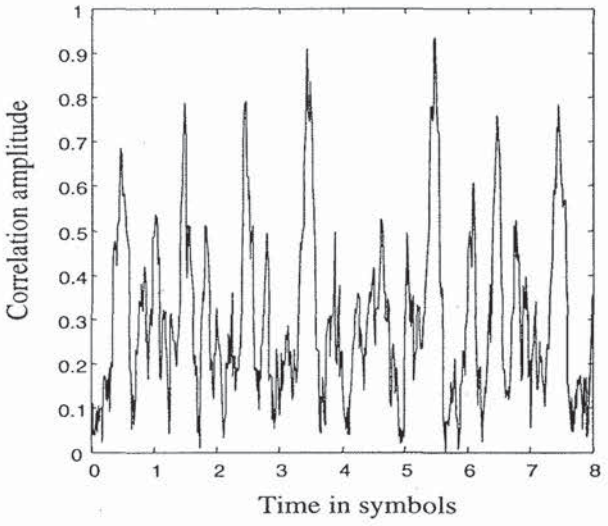


Figure 4.9 Example of correlation output amplitude for eight OFDM symbols with 48 subcarriers and a 20% guard time.

Notice that the undesired correlation sidelobes only create a problem for symbol timing. For frequency offset estimation, they do not play a role. Once symbol timing is known, the cyclic extension correlation output can be used to estimate the frequency offset. The phase of the correlation output is equal to the phase drift between samples that are T seconds apart. Hence, the frequency offset can simply be found as the correlation phase divided by $2\pi T$. This method works up to a maximum absolute

frequency offset of half the subcarrier spacing. To increase this maximum range, shorter symbols can be used, or special training symbols with different PN sequences on odd and even subcarrier frequencies to identify a frequency offset of an integer number of subcarrier spacings [13].

The noise performance of the frequency offset estimator is now determined for an input signal $r(t)$ that consists of an OFDM signal $s(t)$ with power P and additive Gaussian noise $n(t)$ with a one-sided noise power spectral density of N_0 within the bandwidth of the OFDM signal:

$$r(t) = s(t) + n(t) \tag{4.7}$$

The frequency offset estimator multiplies the signal by a delayed and conjugated version of the input to produce an intermediate signal $y(t)$ given by

$$y(t) = r(t)r^*(t-T) = \|s(t)\|^2 \exp(j\varphi) + n(t)s^*(t-T) + n^*(t-T)s(t) + n(t)n^*(t-T) \tag{4.8}$$

The first term in the right-hand side of (4.8) is the desired output component with a phase equal to the phase drift over a T-second interval and a power equal to the squared signal power. The next two terms are products of the signal and the Gaussian noise. Because the signal and noise are uncorrelated, and because noise samples separated by T seconds are uncorrelated, the power of the two terms is equal to twice the product of signal power and noise power. Finally, the power of the last term of (4.8) is equal to the squared noise power. If the input SNR is much larger than one, the power of the squared noise component becomes negligible compared with the power of the other two noise terms. For practical OFDM systems, the minimum input SNR is about 6 dB, so the signal power is four times the noise power. In this case, the power of the squared noise component is eight times smaller than the power of the two signal-noise product terms.

The frequency offset is estimated by averaging $y(t)$ over an interval equal to the guard time T_G and then estimating the phase of $y(t)$. Because the desired output component of (4.8) is a constant vector, averaging reduces the noise that is added to this vector. Assuming that the squared noise component may be neglected, the output SNR is approximated as

$$SNR_o \cong \frac{P^2}{2PN_o / T_G} = \frac{PT_G}{2N_o} \tag{4.9}$$

Figure 4.10 shows a vector representation of the phase estimation, where the noise is divided into in-phase and quadrature components, both having a noise power of N_o/T_G.

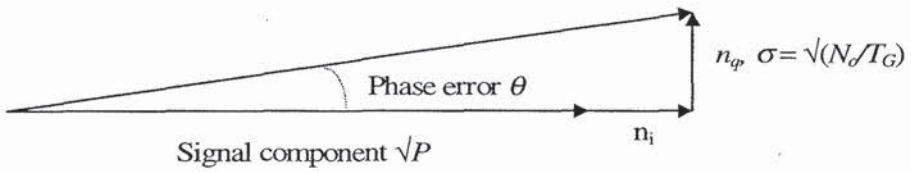


Figure 4.10 Vector representation of phase drift estimation.

The phase error θ is given by (4.10), where the approximation has been made that n_i and n_q are small compared with the signal amplitude $\sqrt{P}$.

$$\theta = \tan^{-1}\left(\frac{n_q}{\sqrt{P} + n_i}\right) \cong \frac{n_q}{\sqrt{P}} \tag{4.10}$$

Because the frequency offset estimation error is equal to the phase error θ divided by $2\pi T$, the standard deviation of the frequency error is given by

$$\sigma_f \cong \frac{1}{2\pi T}\sqrt{\frac{N_o}{PT_G}} = \frac{1}{2\pi T}\sqrt{\frac{1}{E_s / N_o}\frac{T_G}{T_s}} \tag{4.11}$$

Here, T_s is the symbol interval and E_s/N_o is the symbol-to-noise energy ratio, defined as

$$\frac{E_s}{N_o} = \frac{PT_s}{N_o} \tag{4.12}$$

E_s/N_o is equal to the bit energy-to-noise density E_b/N_o multiplied by the number of bits per symbol. Because OFDM typically has a large number of bits per symbol and E_b/N_o is larger than 1 for successful communications, typical E_s/N_o values are much larger than 1. For instance, with 48 subcarriers using 16-QAM and rate ½ coding, there are 96 bits per OFDM symbol. In this case, E_s/N_o is about 20 dB larger than E_b/N_o. So for typical E_b/N_o values around 10 dB, typical E_s/N_o values are around 30 dB.

Figure 4.11 shows the frequency estimation error versus E_s/N_o for three different T_G/T_s ratios. The frequency error is normalized to the subcarrier spacing $1/T$, so a value of 0.01 means 1% of the subcarrier spacing. The solid lines represent calculated values according to equation (4.11), while the dotted lines are derived from simulations. The difference between the two set of curves show the effect of the simplifications made in the derivation of (4.11). For E_s/N_o values of 30 dB or more, the difference is negligible, but around 20 dB, the simulated errors are about 50% larger than the calculated values.

Section 4.3 explained that the frequency error preferably had to be less than 1% of the subcarrier spacing to have a negligible performance degradation. From Figure 4.11, we can learn that such an error level can be achieved at an E_s/N_o value of 26, 31, and 34 dB for a T_G/T_s ratio of 1, 0.2, and 0.1, respectively. A lower T_G/T_s ratio means that a smaller fraction of an OFDM symbol is used for synchronization, hence more SNR is required to attain the same performance as for a larger T_G/T_s value.

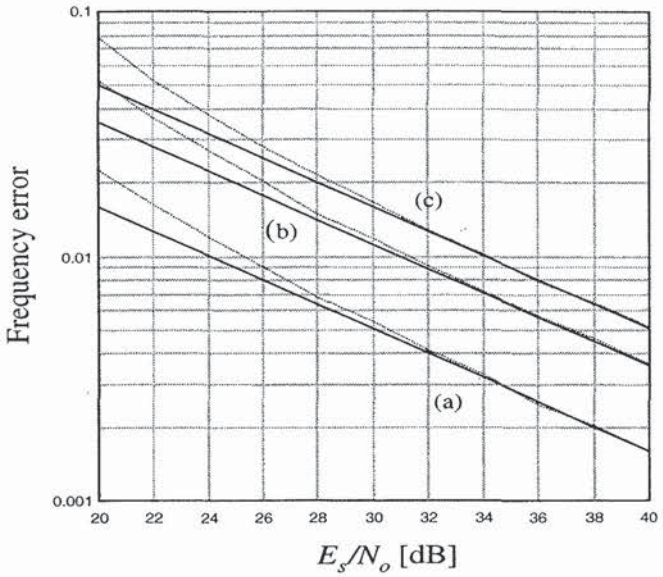


Figure 4.11 Frequency estimation error normalized to the subcarrier spacing. Solid lines are calculated; dotted lines are simulated. (a) T_G/T_s=1, (b) T_G/T_s=0.2, (c) T_G/T_s=0.1.

If the required E_s/N_o value for an acceptable frequency error level is too large, then averaging the vector $y(t)$ in (4.8) over multiple OFDM symbols can be used to increase the effective signal-to-noise ratio. For averaging over K symbols, the frequency error standard deviation becomes

$$\sigma_f \cong \frac{1}{2\pi T}\sqrt{\frac{1}{KE_s / N_o}\frac{T_G}{T_s}} \tag{4.13}$$

Averaging over K symbols has the effect that the curves of Figure 4.11 shift to the left by 10log$K$ dB. For instance, when averaging over four OFDM symbols, a 1% frequency error is achieved at an E_s/N_o value of 28 dB for a T_G/T_s ratio of 0.1 instead of 34 dB without averaging.

Notice that a T_G/T_s ratio of one – curve (a) in Figure 4.11 – is a special case where the guard time is equal to the symbol period. For normal OFDM data symbols, this is not possible, as it would mean that the FFT interval is zero. It does, however,

correspond to the interesting case where two identical OFDM symbols are used to estimate the frequency offset. In this case, all samples of a symbol can be used to estimate the phase difference with the corresponding samples of the other symbol. Hence, (4.11) applies with T_G/T_s set to 1, although T_G is not really a guard time in this case.

4.6 SYNCHRONIZATION USING SPECIAL TRAINING SYMBOLS

The synchronization technique based on the cyclic extension is particularly suited to tracking or to blind synchronization in a circuit-switched connection, where no special training signals are available. For packet transmission, however, there is a drawback because an accurate synchronization needs an averaging over a large (>10) number of OFDM symbols to attain a distinct correlation peak and a reasonable SNR. For high-rate packet transmission, the synchronization time needs to be as short as possible, preferably a few OFDM symbols only. To achieve this, special OFDM training symbols can be used for which the data content is known to the receiver [11-13]. In this way, the entire received training signal can be used to achieve synchronization, whereas the cyclic extension method only uses a fraction of each symbol.

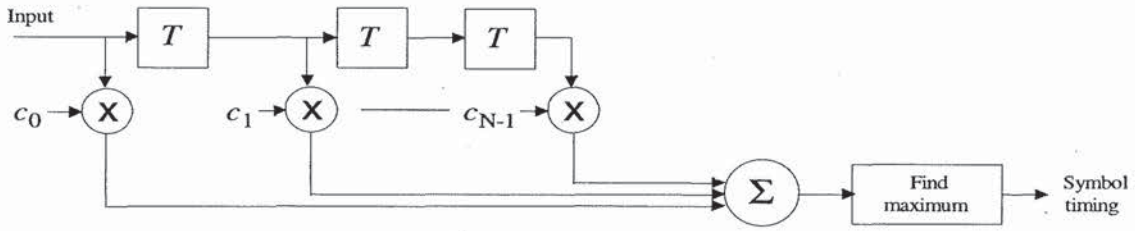


Figure 4.12 Matched filter that is matched to a special OFDM training symbol.

Figure 4.12 shows a block diagram of a matched filter that can be used to correlate the input signal with the known OFDM training signal. Here, T is the sampling interval and c_i are the matched filter coefficients, which are the complex conjugates of the known training signal. From the correlation peaks in the matched filter output signal, both symbol timing and frequency offset can be estimated, as will be explained in this section. Notice that the matched filter correlates with the OFDM time signal, *before* performing an FFT in the receiver. Hence, this technique is very similar to synchronization in a direct-sequence spread-spectrum receiver, where the input signal is correlated with a known spreading signal. In fact, the latter approach of using a single-carrier training signal can also be combined with OFDM, as proposed in [14], but here we will assume that the training signal consists of normal OFDM data symbols.

Figure 4.13 shows an example of the matched filter output for an OFDM training symbol with 48 subcarriers. The training signal for this case consisted of five

identical OFDM symbols without a guard time. Alternatively, it could be stated that there is only one OFDM symbol with a guard time equal to four IFFT intervals, because the IFFT output is repeated four times. The reason for having the training symbol interval equal to the IFFT interval is that this gives the best possible cyclic autocorrelation properties in terms of low undesired sidelobes. This can be seen in Figure 4.13(a), which shows the undesired sidelobes to be at least 20 dB lower than the main correlation peaks. An exception to this occurs at the beginning of the correlation. The reason for this is that at this point, an aperiodic correlation is performed instead of a cyclic correlation, because the matched filter is partly filled with zero values until a full OFDM symbol has been received. A similar effect happens at the end of the training—not shown in the picture—when the matched filter will partly correlate with samples from the following OFDM data symbol that is different from the training symbol. Hence, to avoid undesired partial correlations, the matched filter outputs during the first and last symbol intervals should be skipped. The values in between can be used to detect the main correlation peak, which gives the desired symbol timing information.

The correlation function of Figure 4.13(a) was made for the case of a zero fractional timing offset between the input signal and the known training signal. This means that the matched filter tap values, which are equal to the conjugated training signal samples, are exactly equal to the conjugated sample values of the incoming OFDM signal. This ideal situation does not occur when there is a timing offset of some fraction of a sample interval between the input signal and the known training signal. To see the effect of a fractional timing offset, Figure 4.13(b) shows the correlation output for the worst case timing offset of half a sampling interval. In this case, instead of one main peak per symbol interval, there are two equally strong peaks with a slightly smaller amplitude than the single peak in the case of no timing offset. However, the relative level of undesired correlation sidelobes is still 20 dB below the main peaks.

The plots of Figure 4.13 assumed unquantized input signals and tap values. In practice, it is desirable to have a low number of quantization bits to keep the implementation simple. Figure 4.14 shows the correlation output where the matched filter tap values are quantized to {-1, 0, 1} values for both the real and imaginary parts. This reduces the complexity of the multiplications in the matched filter to additions, having a relatively low hardware complexity. As we can see from Figure 4.14, the correlation output looks different from the unquantized case in Figure 4.13, but the undesired sidelobe level is still about 20 dB below the main peak. Such good correlation properties cannot be achieved with any arbitrary quantized OFDM signal. To minimize the effects of the quantization, the best results are obtained with OFDM signals that have minimum amplitude fluctuations. In the case of Figure 4.14, the OFDM symbol consists of the IFFT of a length 48 complementary code, which results in a signal with peak amplitude fluctuations that are no more than 3 dB larger than the root mean square value. More details about these complementary codes can be found in Chapter 6, which deals with the OFDM peak-to-average power issue.

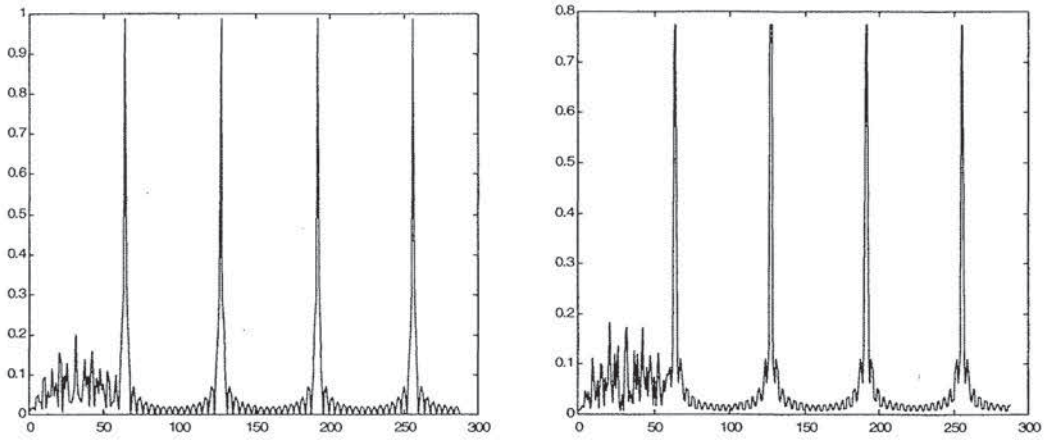


Figure 4.13 Matched filter output versus sample number for 4 training symbols, using 48 subcarriers and 64 samples per symbol. (a) Zero fractional timing offset between input signals and matched filter coefficients, (b) worst case fractional offset of half a sample between input signal and reference pulse.

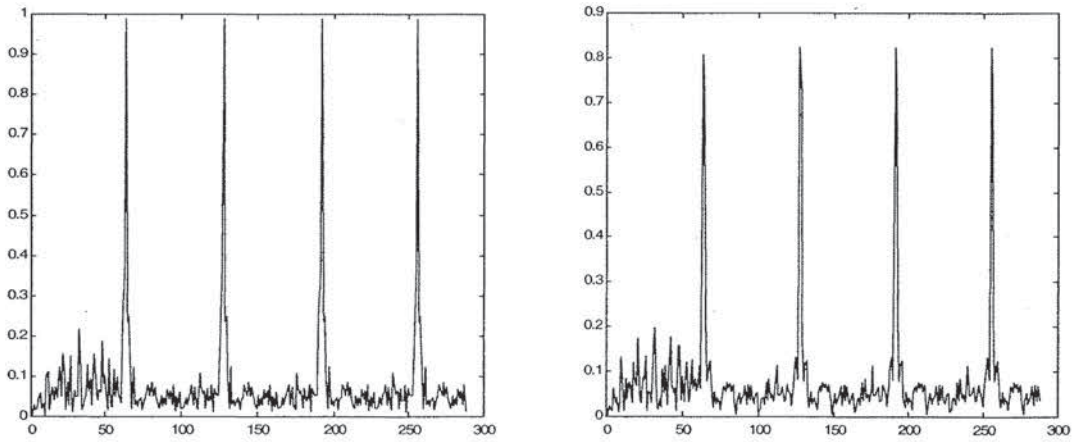


Figure 4.14 Matched filter output versus sample number with {1,–1,0} values for in-phase and quadrature coefficients. (a) Zero fractional timing offset between input signals and matched filter coefficients, (b) worst case fractional offset of half a sample between input signal and reference pulse.

4.7 OPTIMAL TIMING IN THE PRESENCE OF MULTIPATH

The task of OFDM symbol timing is to minimize the amount of ISI and ICI. This type of interference is absent when the FFT is taken over the flat part of the signaling

window, which is shown in Figure 4.15. This window is the envelope of the transmitted OFDM symbols. Within the flat part of the window, all subchannels maintain perfect orthogonality. In the presence of multipath, however, orthogonality is lost if the multipath delays exceed the effective guard time, which is equal to the duration of the flat window part minus the FFT period.

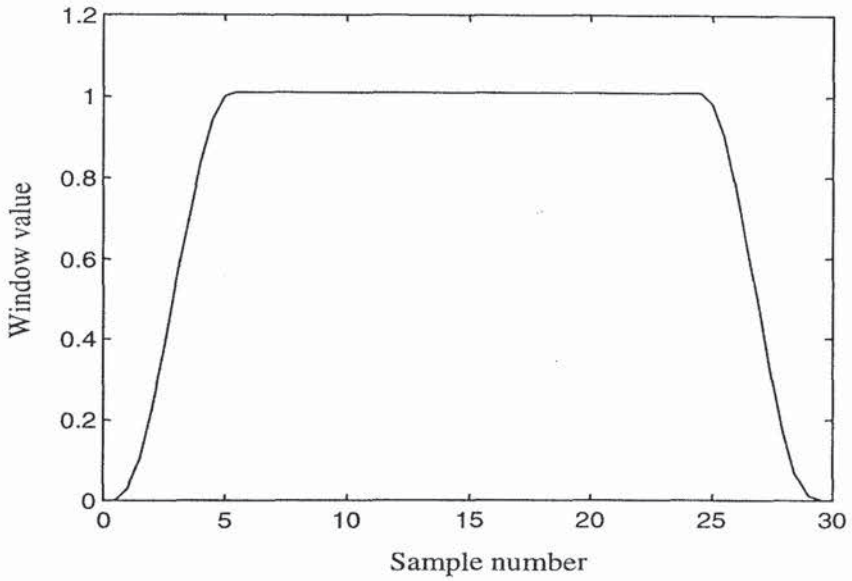


Figure 4.15 Raised cosine window.

The effect of multipath propagation on ISI and ICI is illustrated in Figure 4.16. It shows the windowing envelopes of three OFDM symbols. The radio channel consists of two paths with a relative delay of almost half a symbol and a relative amplitude of 0.5. The receiver selects the FFT timing such that the FFT is taken over the flat envelope part of the strongest path. Because the multipath delay is larger than the guard time, however, the FFT period cannot at the same time cover a totally flat envelope part of the weaker signal. As a result, the nonflat part of the symbol envelope causes ICI. At the same time, the partial overlap of the previous OFDM symbol in the FFT period causes ISI.

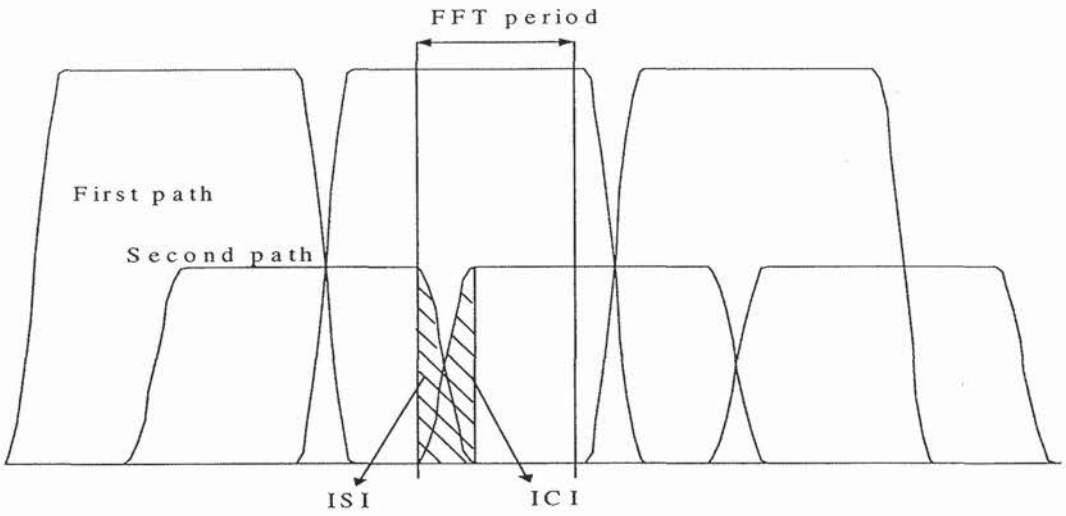


Figure 4.16 ISI/ICI caused by multipath signals.

The solution to the timing problem is to find the delay window—with a width equal to the guard time—that contains maximum signal power. The optimal FFT starting time, then, is equal to the starting delay of the found delay window, plus the delay that occurs between a matched filter peak output from a single OFDM pulse and the delay of the last sample on the flat part of the OFDM signal envelope, minus the length of the FFT interval.

As an example of the timing procedure, Figure 4.17 depicts a simulated matched filter output. By performing a moving average over four samples (which is assumed here to be the length of the guard interval) of the matched filter output power, it is detected that samples 21 to 24 contain the most power. Hence, the starting sample for the FFT is 21 minus some constant number which is mentioned above.

Figure 4.17 clearly shows the advantage of looking for maximum power in the whole guard interval, rather than looking for the maximum correlation output only. If the latter is applied to the example of Figure 4.17, then sample 23 would be chosen instead of sample 21. As a result, the multipath power at samples 21 and 22 would cause extra ISI and ICI, while the useful signal power of samples 24 to 27 would be less than the power of the samples 21 to 24. Hence, the signal-to-interference ratio can easily be degraded by several dB if the suboptimal maximum peak detection is used.

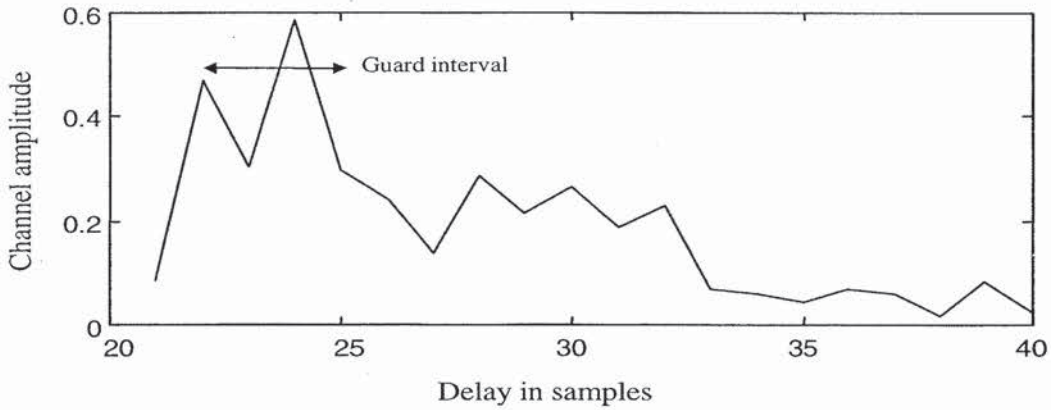


Figure 4.17 Example of channel impulse response.

We now prove that maximizing power in a certain delay window actually maximizes the signal-to-interference ratio. Figure 4.18 shows the OFDM symbol structure, where T is the time needed by the FFT. If a multipath signal is introduced with a relative[1] delay exceeding T_{g1}, it will cause ISI and ICI. Similarly, multipath signals with relative delays less than $-T_{g2}$ cause ISI and ICI. The timing problem is now to choose T_{g1} and T_{g2} such that the amount of ICI and ISI after the FFT is minimized.

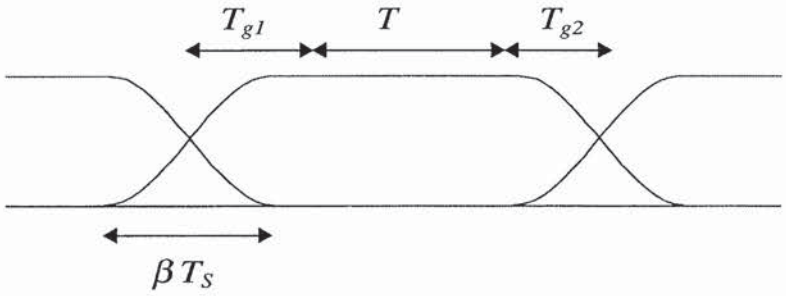


Figure 4.18 OFDM symbol structure.

From the above, it is clear that ISI and ICI are caused by all multipath signals, which delays fall outside a window of $T_g = T_{g1} + T_{g2}$. All multipath signals within this delay window contribute to the effectively used signal power. Hence, the optimal timing circuit maximizes the signal-to-(ISI+ICI) ratio (SIR), given by

[1] Relative to the delay of the shown reference OFDM signal.

$$SIR = \frac{S_u}{S_t - S_u},\ S_u = \int_{T_o}^{T_o+T_g} \|h(\tau)\|^2 d\tau,\ S_t = \int_{-\infty}^{\infty} \|h(\tau)\|^2 d\tau \tag{4.14}$$

Here, $T_o = -T_{g2}$ is the timing offset of the guard time window T_g. S_t denotes the total received signal power and S_u is the useful signal power. Because only S_u depends on the timing offset T_o, the SIR is maximized by maximizing S_u; that is, choosing the T_o value that contains the largest power of $h(\tau)$ in the interval $\{T_o, T_o+T_G\}$.

REFERENCES

[1] Pollet, T., M. van Bladel and M. Moeneclaey, "BER Sensitivity of OFDM Systems to Carrier Frequency Offset and Wiener Phase Noise," *IEEE Trans. on Comm.*, Vol. 43, No. 2/3/4, pp. 191–193, Feb.-Apr. 1995.

[2] Pollet, T., and M. Moeneclaey, "Synchronizability of OFDM Signals," *Proceedings of Globecom '95*, Vol. 3, Singapore, pp. 2054–2058, Nov. 1995.

[3] Kivinen, J., and P. Vainikainen, "Phase Noise in a Direct Sequence Based Channel Sounder," *Proceedings of IEEE PIMRC '97*, Helsinki, pp. 1115–1119, Sep. 1–4, 1997.

[4] Robertson, P., and S. Kaiser, "Analysis of the Effects of Phase-Noise in Orthogonal Frequency Division Multiplex Systems," *Proceedings of IEEE VTC '95*, pp. 1652–1657.

[5] Tomba, L., "On the Effects of Wiener Phase Noise in OFDM systems," *IEEE Trans. on Comm.*, Vol. 46, No. 5, pp. 580–583, May 1998.

[6] Zogakis, T. N., and J. M. Cioffi, "The Effect of Timing Jitter on the Performance of a Discrete Multitone System," *IEEE Trans. on Comm.*, Vol. 44, No. 7, pp. 799–808, July 1996.

[7] Pollet, T., P. Spruyt and M. Moeneclaey, "The BER Performance of OFDM Systems Using Non-Synchronized Sampling," *Proceedings of Globecom '94*, pp. 253–257, Nov. 1994.

[8] Van de Beek, J. J., M. Sandell, M. Isaksson, and P. O. Börjesson, "Low-Complex Frame Synchronization in OFDM Systems," *Proceedings of International Conference on Universal Personal Communications ICUPC '95*, Nov. 1995.

[9] Sandell, M., J. J. van de Beek, and P. O. Börjesson, "Timing and Frequency Synchronization in OFDM Systems Using the Cyclic Prefix," *Proceedings of Int. Symp. On Synchronization*, Saalbau, Essen, Germany, 1995, pp. 16–19, Dec. 14–15, 1995

[10] Böhnke, R., and T. Dölle, "Preamble Structures for HiperLAN Type 2 System," ETSI BRAN Document No. HL13SON1A, Apr. 7, 1999.

[11] Moose, P. H., "A Technique for Orthogonal Frequency Division Multiplexing Frequency Offset Correction," *IEEE Trans. on Comm.*, Vol. 42, No. 10, pp. 2908–2914, Oct. 1994.

[12] Warner, W. D., and C. Leung, "OFDM/FM Frame Synchronization for Mobile Radio Data Communization," *IEEE Trans. on Vehicular Tech.*, Vol. 42, No. 3, pp. 302–313, Aug. 1993.

[13] Schmidl, T. M., and D. C. Cox, "Robust Frequency and Timing Synchronization for OFDM," *IEEE Trans. on Comm.*, Vol. 45, No. 12, pp. 1613–1621, Dec. 1997.

[14] Lambrette, U., M. Speth, and H. Meyr, "OFDM Burst Frequency Synchronization by Single Carrier Training Data," *IEEE Communications Letters*, Vol. 1, No. 2, pp. 46–48, Mar. 1997.

CHAPTER 5

Coherent and Differential Detection

5.1 INTRODUCTION

In an OFDM link, the data bits are modulated on the subcarriers by some form of phase shift keying (PSK) or quadrature amplitude modulation (QAM). To estimate the bits at the receiver, knowledge is required about the reference phase and amplitude of the constellation on each subcarrier. In general, the constellation of each subcarrier shows a random phase shift and amplitude change, caused by carrier frequency offset, timing offset, and frequency selective fading. To cope with these unknown phase and amplitude variations, two different approaches exist. The first one is coherent detection, which uses estimates of the reference amplitudes and phases to determine the best possible decision boundaries for the constellation of each subcarrier. The main issue with coherent detection is how to find the reference values without introducing too much training overhead. To achieve this, several channel estimation techniques exist that will be described in the next section. The second approach is differential detection, which does not use absolute reference values, but only looks at the phase and/or amplitude differences between two QAM values. Differential detection can be done both in the time domain or in the frequency domain; in the first case, each subcarrier is compared with the subcarrier of the previous OFDM symbol. In the case of differential detection in the frequency domain, each subcarrier is compared with the adjacent subcarrier within the same OFDM symbol.

5.2 COHERENT DETECTION

Figure 5.1 shows a block diagram of a coherent OFDM receiver. After downconversion and analog-to-digital conversion, the fast Fourier transform (FFT) is used to demodulate the N subcarriers of the OFDM signal. For each symbol, the FFT output contains N QAM values. However, these values contain random phase shifts and amplitude variations caused by the channel response, local oscillator drift, and timing

ZYXEL-1015
Page 110 (of 277)

offset. It is the task of the channel estimation block to learn the reference phases and amplitudes for all subcarriers, such that the QAM symbols can be converted to binary soft decisions as explained in Chapter 3. The next subsections present several techniques to obtain the channel estimates that are required for coherent detection.

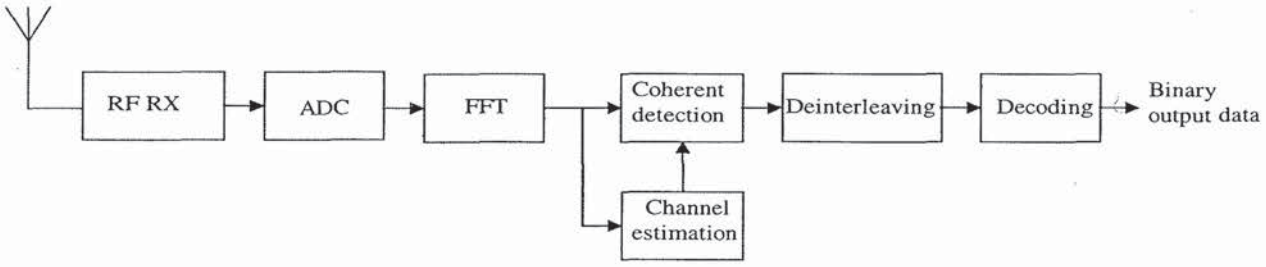


Figure 5.1 Block diagram of an OFDM receiver with coherent detection.

5.2.1 Two-Dimensional Channel Estimators

In general, radio channels are fading both in time and in frequency. Hence, a channel estimator has to estimate time-varying amplitudes and phases of all subcarriers. One way to do this is to use a two-dimensional channel estimator that estimates the reference values based on a few known pilot values. This concept is demonstrated in Figure 5.2, which shows a block of 9 OFDM symbols with 16 subcarriers. The gray subcarrier values are known pilots. Based on these pilots, all other reference values can be estimated by performing a two-dimensional interpolation [1–4].

To be able to interpolate the channel estimates both in time and frequency from the available pilots, the pilot spacing has to fulfill the Nyquist sampling theorem, which states that the sampling interval must be smaller than the inverse of the double-sided bandwidth of the sampled signal. For the case of OFDM, this means that there exist both a minimum subcarrier spacing and a minimum symbol spacing between pilots. By choosing the pilot spacing much smaller than these minimum requirements, a good channel estimation can be made with a relatively easy algorithm. The more pilots are used, however, the smaller the effective Signal-to-Noise ratio, SNR, becomes that is available for data symbols. Hence, the pilot density is a tradeoff between channel estimation performance and SNR loss.

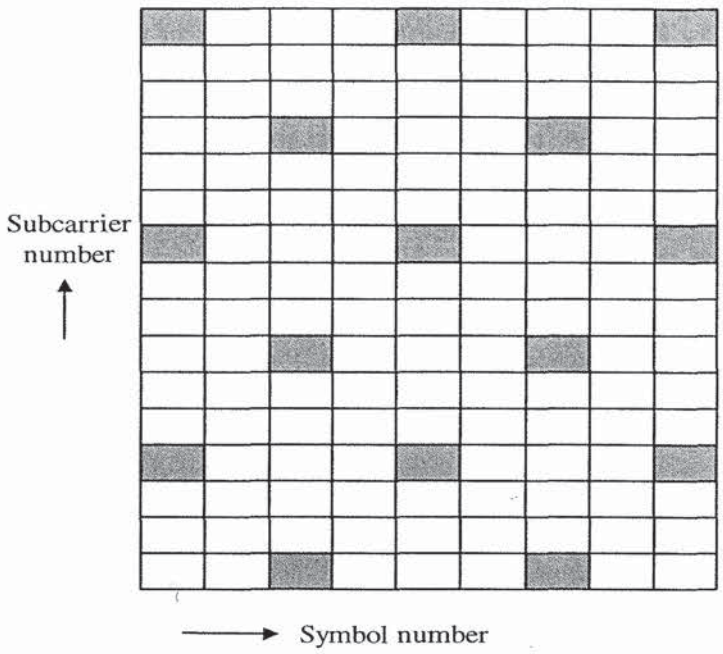


Figure 5.2 Example of pilots (marked gray) in a block of 9 OFDM symbols with 16 subcarriers.

To determine the minimum pilot spacing in time and frequency, we need to find the bandwidth of the channel variation in time and frequency. These bandwidths are equal to the Doppler spread B_d in the time domain and the maximum delay spread τ_{max} in the frequency domain [5]. Hence, the requirements for the pilot spacings in time and frequency s_t and s_f are

$$s_t < \frac{1}{B_d} \tag{5.1}$$

$$s_f < \frac{1}{\tau_{\max}} \tag{5.2}$$

Assume now the available pilot values are arranged in a vector $\hat{\mathbf{p}}$ and the channel values that have to be estimated from $\hat{\mathbf{p}}$ are in a vector $\mathbf{h}$. Notice that we use bold letters to distinguish vectors from scalar variables. It is assumed that any known modulation of the pilots is removed before the estimation; for instance, if the transmitter applies some phase shifts to the pilots, then the receiver has to back-rotate those known phases. The channel estimation problem is now to find the channel estimates $\hat{\mathbf{h}}$ as a linear combination of the pilot estimates $\hat{\mathbf{p}}$. According to [6], the minimum mean square error estimate for this problem is given by

$$\hat{\mathbf{h}} = \mathbf{R}_{\mathbf{h}\hat{\mathbf{p}}}\mathbf{R}_{\hat{\mathbf{p}}\hat{\mathbf{p}}}^{-1}\hat{\mathbf{p}} \tag{5.3}$$

$\mathbf{R}_{\mathbf{h}\hat{\mathbf{p}}}$ is the cross-covariance matrix between $\mathbf{h}$ and the noisy pilot estimates $\hat{\mathbf{p}}$, given by

$$\mathbf{R}_{\mathbf{h}\hat{\mathbf{p}}} = \mathrm{E}\left\{ \mathbf{h}\hat{\mathbf{p}}^{\mathbf{H}} \right\} \tag{5.4}$$

$\mathbf{R}_{\hat{\mathbf{p}}\hat{\mathbf{p}}}$ is the auto-covariance matrix of the pilot estimates:

$$\mathbf{R}_{\hat{\mathbf{p}}\hat{\mathbf{p}}} = \mathrm{E}\left\{ \hat{\mathbf{p}}\hat{\mathbf{p}}^{\mathbf{H}} \right\} = \mathbf{R}_{\mathbf{pp}} + \sigma_n^2(\mathbf{pp}^{\mathbf{H}})^{-1} \tag{5.5}$$

Assuming the pilots all have the same power, which is the case if all pilots are for instance known QPSK symbols, then the pilots' auto-covariance matrix can be rewritten as

$$\mathbf{R}_{\hat{\mathbf{p}}\hat{\mathbf{p}}} = \mathbf{R}_{\mathbf{pp}} + \frac{1}{\mathrm{SNR}}\mathbf{I} \tag{5.6}$$

where SNR is the signal-to-noise ratio per pilot and $\mathbf{R}_{\mathbf{pp}}$ is the autocovariance matrix of the noiseless pilots. With this, the channel estimates can be written as

$$\hat{\mathbf{h}} = \mathbf{R}_{\mathbf{h}\hat{\mathbf{p}}}(\mathbf{R}_{\mathbf{pp}} + \frac{1}{\mathrm{SNR}}\mathbf{I})^{-1}\hat{\mathbf{p}} \tag{5.7}$$

(5.7) basically gives the desired channel estimates as the multiplication of an interpolation matrix with the pilot estimates. Notice that the interpolation matrix does not depend on the received symbols; it only depends on the position of the pilots and the number of pilots and channel estimates. Hence, the interpolation matrix can be designed as a constant matrix, avoiding the need to do matrix inversions in the OFDM receiver.

The elements covariance matrices $\mathbf{R}_{\mathbf{h}\hat{\mathbf{p}}}$ and $\mathbf{R}_{\hat{\mathbf{p}}\hat{\mathbf{p}}}$ can be calculated as follows. Both matrices contain correlation values between subcarrier values for different time and frequency spacings. If k and l are the subcarrier number and OFDM symbol number, respectively, the correlation values are given by

$$E\{h_{k,l}\hat{p}^*_{k',l'}\} = E\{p_{k,l}\hat{p}^*_{k',l'}\} = r_f(k-k')r_t(l-l') \tag{5.8}$$

Here, $r_t(l)$ and $r_f(k)$ are the correlation functions in time and frequency, respectively. For an exponentially decaying multipath power delay profile, $r_f(k)$ is given by

$$r_f(k) = \frac{1}{1 + j2\pi\tau_{rms}k/T} \tag{5.9}$$

Here, $1/T$ is the subcarrier spacing, which is the inverse of the FFT interval T. For a time-fading signal with a maximum Doppler frequency f_{max} and a Jakes spectrum, the time correlation function $r_t(l)$ is given as

$$r_t(l) = J_o(2\pi f_{\max} l T_s) \tag{5.10}$$

where $J_o(x)$ is the zero*th* order Bessel function of the first kind and T_s is the OFDM symbol duration, which is the FFT interval T plus the guard time. To illustrate the channel estimation technique described above, an example will be given for the case of five pilots in a block of five OFDM symbols with five subcarriers. Four pilots are located at the corners of the block and one in the middle, being the third subcarrier of the third OFDM symbol. Using the Jakes fading channel model, an OFDM signal was generated that experienced fading both in time and frequency. Table 5.1 lists the 25 reference channel values for the five subcarriers in each of the five symbols for this example. Of course, an OFDM receiver does not know these reference values; it only has knowledge about the five pilot values, which are located at row and column numbers {1,1}, {1,5}, {3,3}, {5,1}, and {5,5}. In this particular example, no noise is added to the pilot values. Each column represents one symbol with five subcarrier values. We can see from the table that there is fading both in frequency—across the rows—and in time.

Table 5.1
Example channel values for a block of five OFDM symbols and five subcarriers.

1.0386–0.2468i	1.1333–0.2441i	1.1777–0.2491i	1.1693–0.2617i	1.1048–0.2761i
0.8938–0.4782i	0.9798–0.4821i	1.0172–0.4842i	1.0040–0.4847i	0.9386–0.4778i
0.6726–0.6302i	0.7479–0.6398i	0.7802–0.6402i	0.7675–0.6316Ii	0.7099–0.6082i
0.4173–0.6794i	0.4809–0.6913i	0.5093–0.6897i	0.5007–0.6750i	0.4567–0.6421i
0.1794–0.6258i	0.2321–0.6349i	0.2578–0.6303i	0.2548–0.6125i	0.2258–0.5774i

Table 5.2 gives the elements of the pilot auto-covariance matrix. This matrix is independent of the received signal, so it can be precalculated using (5.8), (5.9), and (5.10). The first row of $\mathbf{R}_{pp}$ consists of the correlations between the first pilot and all five pilots. The first value is 1, as this is the correlation of the first pilot with itself. The second value of the first row is the correlation between the first pilot at the first subcarrier with the pilot of the first subcarrier of the fifth symbol. Because these pilots are both on the same subcarrier, the frequency correlation component $r_f(k)$ is equal to 1, and the correlation value is purely determined by the time fading component $r_t(k)$. This explains why the imaginary component of the second correlation value is zero, because $r_t(k)$ (5.10) is a strictly real function. Notice that from the matrix $\mathbf{R}_{pp}$ the matrix $\mathbf{R}_{\hat{p}\hat{p}}$ is formed by adding an identity matrix multiplied by the inverse of the SNR (5.6). In practice, the SNR is not known *a priori*, so an expected value is used. In our example, we will use an SNR of 10 dB. Using a large SNR value gives a relatively large weight on $\mathbf{R}_{pp}$ in (5.7).

Table 5.2
Example of covariance matrix $\mathbf{R}_{pp}$.

1.0000	0.4720	0.7967–0.2589i	0.7568–0.5499i	0.3572–0.2595i
0.4720	1.0000	0.7967–0.2589i	0.3572–0.2595i	0.7568–0.5499i
0.7967–0.2589i	0.7967–0.2589i	1.0000	0.7967–0.2589i	0.7967–0.2589i
0.7568–0.5499i	0.3572–0.2595i	0.7967–0.2589i	1.0000	0.4720
0.3572–0.2595i	0.7568–0.5499i	0.7967–0.2589i	0.4720	1.0000

Table 5.3 gives the example matrix $\mathbf{R}_{h\hat{p}}$. This matrix contains the correlation values between all 25 channel values—in the block of five symbols by five subcarriers—with each of the five pilots. Hence, the matrix has 25 rows and 5 columns. Because the pilots are also part of the 25 channel values, the rows of $\mathbf{R}_{pp}$ are all part of $\mathbf{R}_{h\hat{p}}$; for instance, the first row of $\mathbf{R}_{h\hat{p}}$ is the same as the first row of $\mathbf{R}_{pp}$.

Table 5.3

Example of covariance matrix $\mathbf{R}_{\mathbf{h}\hat{\mathbf{p}}}$.

1.0000	0.4720	0.7967–0.2589i	0.7568–0.5499i	0.3572–0.2595i
0.9836–0.1558i	0.4643–0.0735i	0.8377–0.1327i	0.8584–0.4374i	0.4052–0.2064i
0.9355–0.3040i	0.4416–0.1435i	0.8516	0.9355–0.3040i	0.4416–0.1435i
0.8584–0.4374i	0.4052–0.2064i	0.8377–0.1327i	0.9836–0.1558i	0.4643–0.0735i
0.7568–0.5499i	0.3572–0.2595i	0.7967–0.2589i	1.0000	0.4720
0.9618	0.6820	0.8998–0.2924i	0.7279–0.5289i	0.5161–0.3750i
0.9461–0.1498i	0.6708–0.1062i	0.9461–0.1498i	0.8256–0.4207i	0.5854–0.2983i
0.8998–0.2924i	0.6380–0.2073i	0.9618	0.8998–0.2924i	0.6380–0.2073i
0.8256–0.4207i	0.5854–0.2983i	0.9461–0.1498i	0.9461–0.1498i	0.6708–0.1062i
0.7279–0.5289i	0.5161–0.3750i	0.8998–0.2924i	0.9618	0.6820
0.8516	0.8516	0.9355–0.3040i	0.6445–0.4683i	0.6445–0.4683i
0.8377–0.1327i	0.8377–0.1327i	0.9836–0.1558i	0.7310–0.3725i	0.7310–0.3725i
0.7967–0.2589i	0.7967–0.2589i	1.0000	0.7967–0.2589i	0.7967–0.2589i
0.7310–0.3725i	0.7310–0.3725i	0.9836–0.1558i	0.8377–0.1327i	0.8377–0.1327i
0.6445–0.4683i	0.6445–0.4683i	0.9355–0.3040i	0.8516	0.8516
0.6820	0.9618	0.8998–0.2924i	0.5161–0.3750i	0.7279–0.5289i
0.6708–0.1062i	0.9461–0.1498i	0.9461–0.1498i	0.5854–0.2983i	0.8256–0.4207i
0.6380–0.2073i	0.8998–0.2924i	0.9618	0.6380–0.2073i	0.8998–0.2924i
0.5854–0.2983i	0.8256–0.4207i	0.9461–0.1498i	0.6708–0.1062i	0.9461–0.1498i
0.5161–0.3750i	0.7279–0.5289i	0.8998–0.2924i	0.6820	0.9618
0.4720	1.0000	0.7967–0.2589i	0.3572–0.2595i	0.7568–0.5499i
0.4643–0.0735i	0.9836–0.1558i	0.8377–0.1327i	0.4052–0.2064i	0.8584–0.4374i
0.4416–0.1435i	0.9355–0.3040i	0.8516	0.4416–0.1435i	0.9355–0.3040i
0.4052–0.2064i	0.8584–0.4374i	0.8377–0.1327i	0.4643–0.0735i	0.9836–0.1558i
0.3572–0.2595i	0.7568–0.5499i	0.7967–0.2589i	0.4720	1.0000

The above matrices can be combined to the final interpolation matrix listed in Table 5.4. The matrix has 25 rows and 5 columns; when multiplied by the row vector containing five measured pilots, 25 channel estimates are obtained. Figure 5.3 shows the estimation errors for the channel example of Table 5.1. We can see that the relative estimation error is between 3% and 14%. The largest errors are located at the edges of the block, which is a typical phenomenon of this kind of interpolation. It suggests that to minimize the interpolation error, pilots should be used that surround the channel positions that have to be estimated.

Table 5.4

Example of interpolation matrix $\mathbf{R}_{h\hat{p}}(\mathbf{R}_{pp}+\frac{1}{\mathrm{SNR}}\mathbf{I})^{-1}$

0.9044+0.0902i	0.0331–0.0034i	0.0333–0.0749i	0.0145–0.0630i	–0.0241 + 0.0562i
0.6009+0.0022i	–0.0908–0.0346i	0.4006+0.0861i	0.1446–0.0753i	–0.1200–0.0040i
0.2760–0.0836i	–0.2100–0.0647i	0.7484+0.2459i	0.2760–0.0836i	–0.2100–0.0647i
0.1446–0.0753i	–0.1200–0.0040i	0.4006+0.0861i	0.6009+0.0022i	–0.0908–0.0346i
0.0145–0.0630i	–0.0241+0.0562i	0.0333–0.0749i	0.9044+0.0902i	0.0331–0.0034i
0.7534+0.0819i	0.2917+0.0323i	0.0409–0.1125i	0.0040–0.0267i	–0.0165+0.0364i
0.4658+0.0002i	0.0993–0.0193i	0.4555+0.0678i	0.0816–0.0547i	–0.0587–0.0169i
0.1620–0.0797i	–0.0955–0.0697i	0.8483+0.2478i	0.1620–0.0797i	–0.0955–0.0697i
0.0816–0.0547i	–0.0587–0.0169i	0.4555+0.0678i	0.4658+0.0002i	0.0993–0.0193i
0.0040–0.0267i	–0.0165+0.0364i	0.0409–0.1125i	0.7534+0.0819i	0.2917+0.0323i
0.5412+0.0619i	0.5412+0.0619i	0.0435–0.1257i	–0.0067+0.0077i	–0.0067+0.0077i
0.2920–0.0072i	0.2920–0.0072i	0.4746+0.0614i	0.0111–0.0344i	0.0111–0.0344i
0.0333–0.0749i	0.0333–0.0749i	0.8830+0.2483i	0.0333–0.0749i	0.0333–0.0749i
0.0111–0.0344i	0.0111–0.0344i	0.4746+0.0614i	0.2920–0.0072i	0.2920–0.0072i
–0.0067+0.0077i	–0.0067+0.0077i	0.0435–0.1257i	0.5412+0.0619i	0.5412+0.0619i
0.2917+0.0323i	0.7534+0.0819i	0.0409–0.1125i	–0.0165+0.0364i	0.0040–0.0267i
0.0993–0.0193i	0.4658+0.0002i	0.4555+0.0678i	–0.0587–0.0169i	0.0816–0.0547i
–0.0955–0.0697i	0.1620–0.0797i	0.8483+0.2478i	–0.0955–0.0697i	0.1620–0.0797i
–0.0587–0.0169i	0.0816–0.0547i	0.4555+0.0678i	0.0993–0.0193i	0.4658+0.0002i
–0.0165+0.0364i	0.0040–0.0267i	0.0409–0.1125i	0.2917+0.0323i	0.7534+0.0819i
0.0331–0.0034i	0.9044+0.0902i	0.0333–0.0749i	–0.0241+0.0562i	0.0145–0.0630i
–0.0908–0.0346i	0.6009+0.0022i	0.4006+0.0861i	–0.1200–0.0040i	0.1446–0.0753i
–0.2100–0.0647i	0.2760–0.0836i	0.7484+0.2459i	–0.2100–0.0647i	0.2760–0.0836i
–0.1200–0.0040i	0.1446–0.0753i	0.4006+0.0861i	–0.0908–0.0346i	0.6009+0.0022i
–0.0241+0.0562i	0.0145–0.0630i	0.0333–0.0749i	0.0331–0.0034i	0.9044+0.0902i

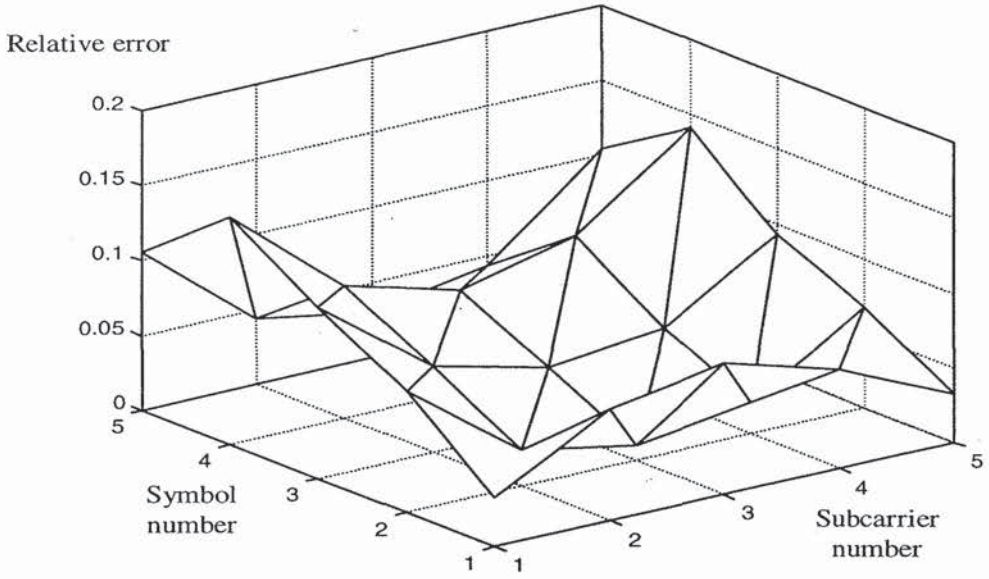


Figure 5.3 Example of relative channel estimation errors versus subcarrier number and symbol number.

5.2.2 One-Dimensional Channel Estimators

The channel estimation technique described in the previous section basically performed a two-dimensional interpolation to estimate points on a time-freqency grid based on several pilots. Instead of directly calculating the two-dimensional solution, it is also possible to separate the interpolation into two one-dimensional interpolations, as illustrated by Figure 5.4 [7]. With this technique, first an interpolation in the frequency domain is performed for all symbols containing pilots. Then, for each subcarrier an interpolation in the time domain is performed to estimate the remaining channel values.

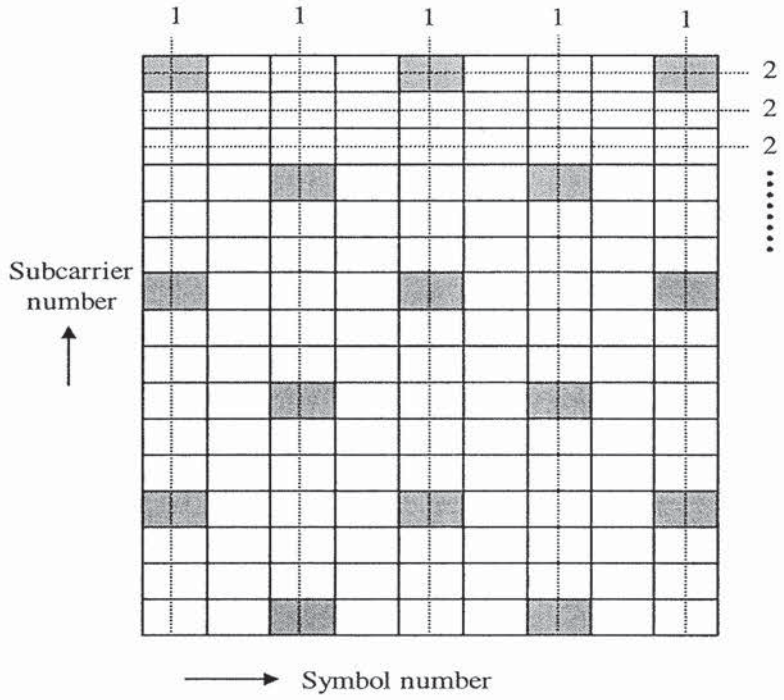


Figure 5.4 Channel estimation with separable filters in frequency (1) and time (2) direction.

5.2.3 Special Training Symbols

The channel estimation techniques from the previous sections were designed to estimate a channel that varied both in time and frequency. These techniques are especially suitable for continuous transmission systems such as Digital Audio Broadcasting or Digital Video Broadcasting, which are both described in Chapter 10. They are not very suited, however, for packet-type communications for two reasons. First, in many packet transmission systems, such as wireless LAN, the packet length is short enough to assume a constant channel during the length of the packet. This means there is no need to estimate time fading, which greatly simplifies the channel estimation problem. Second, using pilots scattered over several OFDM data symbols introduces a delay of several symbols before the first channel estimates can be calculated. Such a delay is undesirable in packet transmission like in an IEEE 802.11 wireless LAN, which requires an acknowledgment to be sent after each packet transmission. Any delay in the reception of a packet will also delay the acknowledgment and hence decrease the effective throughput of the system. An additional disadvantage is the fact that the receiver needs to buffer several OFDM symbols, thereby requiring extra hardware.

For the specific problem of channel estimation in packet transmission systems, the most appropriate approach seems to be the use of a preamble consisting of one or more known OFDM symbols. This approach is sketched in Figure 5.5. The figure shows the time-frequency grid with subcarriers on the vertical axis and symbols on the horizontal axis. All gray subcarriers are pilots. The packet starts with two OFDM symbols for which all data values are known. These training symbols can be used to obtain channel estimates, as well as a frequency offset estimate, as was explained in Chapter 4. After the first two training symbols, Figure 5.5 shows two pilot subcarriers within the data symbols. These pilots are not meant for channel estimation, but for tracking the remaining frequency offset after the initial training. Because this frequency offset affects all subcarriers in a similar way, there is no need to have many pilots with a small frequency spacing as in the case of channel estimation. This type of pilot structure was first mentioned in a proposal for the IEEE 802.11 OFDM standard [8], which is described in detail in Chapter 9.

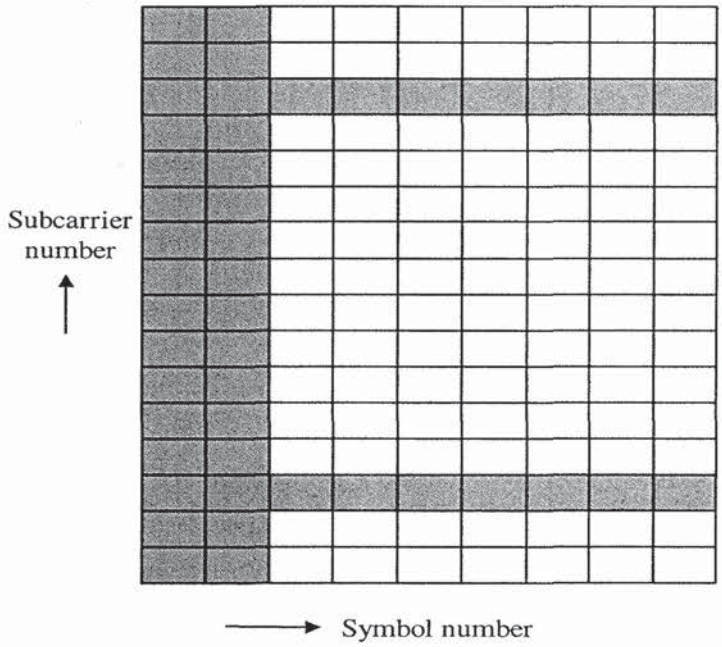


Figure 5.5 Example of a packet with two training symbols for channel estimation and two pilot subcarriers used for frequency synchronization.

The choice of the number of training symbols is a tradeoff between a short training time and a good channel estimation performance. Using two training symbols is a reasonable choice, because it gives a 3-dB-lower noise level in the channel estimates by a simple averaging of the two training symbols, and a minimum of two training symbols is convenient anyway to estimate the frequency offset by comparing the phase shift of the two identical symbols.

If multiple training symbols are used, there is actually no need to repeat an entire OFDM symbol including guard time. A more efficient way is to repeat the IFFT interval and to keep a single guard time, as depicted in Figure 5.6. This can also be viewed as extending the guard time to twice the original guard time plus an extra time equal to the IFFT interval. The advantage of this approach is that it makes the training extra robust to multipath; a receiver can perform a channel estimation by taking the FFT of the averaged IFFT intervals of the long training symbol. This channel estimate will be free from ISI and ICI as long as the relative multipath delays are smaller than the guard time of the training symbol, which is now doubled relative to the guard time of the OFDM data symbols.

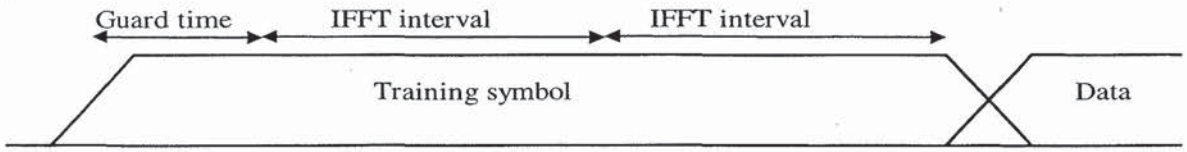


Figure 5.6 Extended training symbol for channel estimation with a single guard time and multiple IFFT intervals.

One of the main assumptions when using pilot symbols only at the start of a packet is that channel variations during the rest of the packet are negligible. Whether this is a valid assumption depends on the packet duration and the Doppler bandwidth. This issue is discussed further in Section 5.3.1. This section describes a related differential detection technique that relies on the same assumption of a channel that is nearly constant in time.

5.2.4 Decision-Directed Channel Estimation

The previously described coherent detection techniques are all based on pilots to estimate the channel. A disadvantage of those pilots is that they cost a certain percentage of the transmitted power. To avoid this loss, decision-directed channel estimation can be used. Here, instead of pilots, data estimates are used to remove the data modulation from the received subcarriers, after which all subcarriers can be used to estimate the channel. Of course, it is not possible to make reliable data decisions before a good channel estimate is available. Therefore, only decisions from previous symbols are used to predict the channel in the current symbol [9]. This is in contrast to the pilot methods, where the channel for a certain symbol is estimated from pilots within, before, and after that particular symbol. If the channel is relatively slowly varying in time, however, such that there is a large correlation between adjacent symbols, then there is a negligible impact on performance if only earlier symbols are used to estimate the channel for a particular OFDM symbol.

To start the decision-directed channel estimation, at least one known OFDM symbol must be transmitted. This enables the receiver to attain channel estimates for all subcarriers, which are then used to detect the data in the following OFDM symbol. Once data estimates are available for a symbol, these estimates are used to remove the data modulation from the subcarriers, after which those subcarrier values can be used as pilots in exactly the same way as described in Sections 5.2.1 and 5.2.2.

5.3 DIFFERENTIAL DETECTION

The key idea behind all coherent detection techniques discussed in the previous subsections is that they somehow estimate the channel to obtain an absolute reference phase and amplitude for each subcarrier in each OFDM symbol. In contrast to this, differential detection does not perform any channel estimation, thereby saving both complexity and pilots at the cost of a somewhat reduced SNR performance. A general block diagram of an OFDM receiver using differential detection is shown in Figure 5.7. Instead of using an absolute reference, differential detection compares each subsymbol with another subsymbol, which can be a previous subcarrier in the same OFDM symbol, or the same subcarrier of a previous OFDM symbol. The next sections explain both variations of differential detection and show how differential detection can even be applied to multi-amplitude modulation.

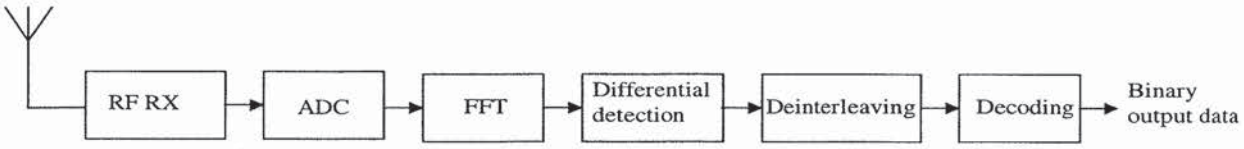


Figure 5.7 Block diagram of an OFDM receiver with differential detection.

5.3.1 Differential Detection in the Time Domain

If differential detection is applied in the time domain, then each subsymbol is compared to the subsymbol on the same subcarrier of the previous OFDM symbol, as depicted in Figure 5.8.

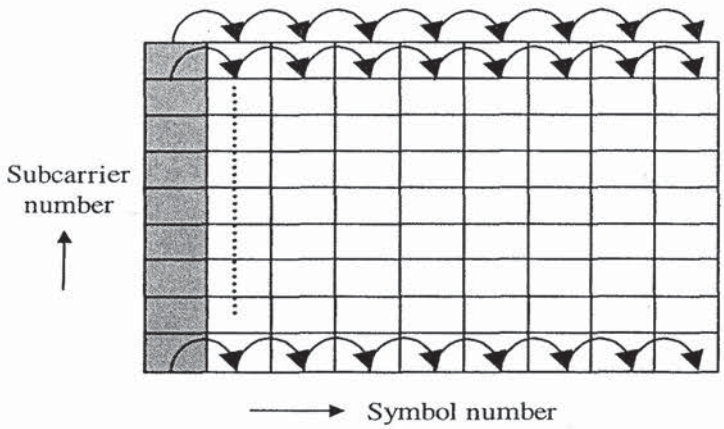


Figure 5.8 Differential detection in the time domain. Gray subcarriers are pilots that are needed as initial phase references.

To make differential detection possible, the transmitter has to apply differential encoding. For a PSK signal with input phases φ_{kj}, the differentially encoded phases θ_{ij} are

$$\theta_{ij} = \sum_{k=0}^{i} \varphi_{kj} \mod(2\pi) \tag{5.11}$$

where i and k are the symbol number and j is the subcarrier number. Differential detection is essentially applied to each subcarrier separately. Because at the start of a transmission no previous symbol values are yet available, the subcarrier values of the first symbol are chosen to be some arbitrary values.

At the receiver, the FFT output of symbol i and subcarrier j can be written as

$$x_{ij} = a_{ij} e^{\theta_{ij}+\beta_{ij}} + n_{ij} \tag{5.12}$$

where a_{ij}, β_{ij}, and n_{ij} are the channel amplitude, channel phase, and additive noise component of symbol i and subcarrier j. A differential phase detection in the time domain is performed by multiplying each FFT output with the conjugated FFT output of the same subcarrier from the previous OFDM symbol.

$$y_{ij} = x_{ij} x^*_{i-1,j} = a_{ij} a_{i-1,j} e^{\varphi_{ij}+\beta_{ij}-\beta_{i-1,j}} + n_{ij} a_{i-1,j} e^{\theta_{i-1,j}+\beta_{i-1,j}} + n_{i-1,j} a_{ij} e^{\theta_{ij}+\beta_{ij}} + n_{ij} n^*_{i-1,j} \tag{5.13}$$

The first term of (5.13) has the desired phase φ_{kj}, but it also has an undesired phase disturbance $\beta_{ij} - \beta_{i-1,j}$, which is the channel phase shift on subcarrier j from symbol $i-1$ to i. The latter disturbance depends only on the time fading, so to have a negligible impact on the phase detection, the OFDM symbol duration has to be small relative to the channel coherence time.

Figure 5.9 shows the correlation between signal samples as a function of the normalized time difference $f_{max}T_s$, where f_{max} is the maximum Doppler spread and T_s is the OFDM symbol duration. The correlation is calculated according to (5.10), with the number of OFDM symbols l set to one. Figure 5.9 can be used to determine the maximum tolerable Doppler spread, depending on the allowable phase error and the OFDM symbol duration. The maximum tolerable level of channel estimation errors can be related to the correlation between two OFDM symbols by writing the channel value y_{ij} as a function of the channel value $y_{i-1,j}$ of the previous symbol

$$y_{ij} = r_t\, y_{i-1,j} + \sqrt{(1-r_t^2)}\, q_{ij} \tag{5.14}$$

Here, r_t is the correlation between channel values of two OFDM symbols on the same subcarrier, as given by (5.10). q_{ij} is a randomly distributed component with unity power. The difference between the channel values $y_{i-1,j}$ and y_{ij} is given by

$$y_{i-1,j} - y_{ij} = y_{i-1,j}(1-r_t) + \sqrt{(1-r_t^2)}\, q_{ij} \tag{5.15}$$

This difference in channel values between two OFDM symbols determines the loss in SNR performance. For a negligible loss of performance, the signal-to-distortion ratio (SDR) should be much larger than the SNR that is needed to achieve a certain maximum BER or PER.

Assuming $y_{i-1,j}$ and q_{ij} are uncorrelated, the distortion power is the sum of the powers of both components in (5.15). Because both $y_{i-1,j}$ and q_{ij} have unity power, the SDR can be written as

$$SDR = \frac{1}{2(1-r_t)} \tag{5.16}$$

The required SNR values depend on coding rate and type of modulation. Some practical values can be deduced from Chapter 3. For instance, an SNR of 4 dB is required to get a BER of 10^{-5} using QPSK and rate 1/2 convolutional coding. For a loss in SNR performance of less than 1 dB, the distortion power should be at least 6 dB lower than the noise power, so the SDR needs to be 10 dB or more. Using (5.16), an

SDR of 10 dB requires a correlation value r_t of 0.95. From Figure 5.10—which is a zoom-in on the first part of Figure 5.9—we can see that a correlation value of 0.95 corresponds to a normalized time distance $f_{max}T_s$ of approximately 0.07. Hence, the maximum allowable Doppler frequency in this case is $0.07/T_s$. The Doppler frequency can be related to the maximum allowable velocity v as

$$f_{\max} = f_c \frac{v}{c} \tag{5.17}$$

where f_c is the carrier frequency and c the speed of light. For example, for a carrier frequency of 5 GHz and a symbol duration of 4 μs, a maximum Doppler frequency of $0.07/T_s = 17.5$ kHz leads to a maximum allowable user velocity of 1,050 m/s. Because this speed is orders of magnitudes above practical values, we can conclude that for normal speeds, the channel change between two OFDM symbols is negligible for the parameters of the above example.

An interesting analogy exists between differential detection in the time domain and the coherent detection method using pilot symbols at the beginning of a packet, which was described in Section 5.2.3. Both methods rely on the fact that the channel is relatively constant in time. When the first symbols of a packet are used as a coherent reference for the rest of the packet, a nonzero Doppler bandwidth will cause the channel estimation errors to grow with the packet length. Hence, the maximum possible packet length is determined by the allowable level of channel estimation errors. To calculate the errors in the channel estimation, the same equations can be used as for differential detection in the time domain, with the only difference that T_s is replaced by the packet duration T_p. For instance, for the parameters of the previous example, it is required that $f_{max}T_p$ is approximately 0.07. Because f_{max} is equal to $f_c v/c$, the maximum possible packet duration T_p becomes $0.07c/f_c v \cong 2.8$ ms at a walking speed of v=1.5 m/s. At a vehicle speed of 30 m/s, however, the maximum packet duration is limited to about 0.1 ms.

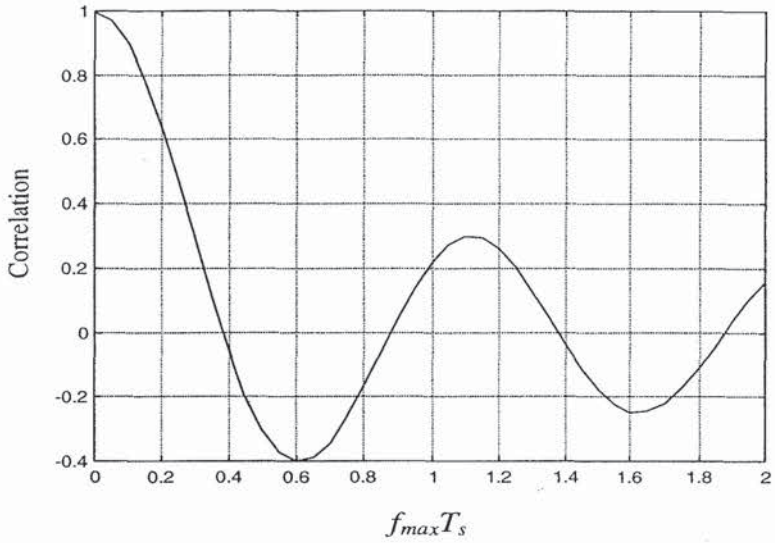


Figure 5.9 Correlation between symbols versus the normalized time distance $f_{max}T_s$.

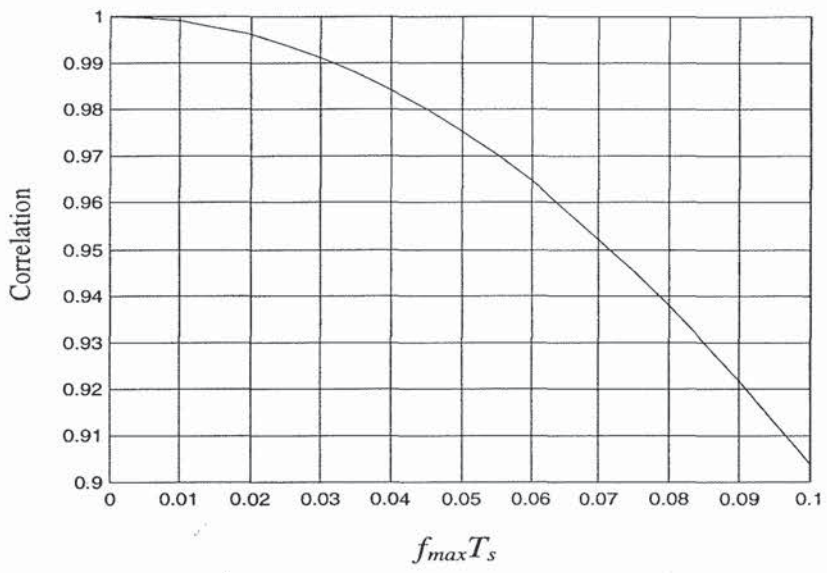


Figure 5.10 Correlation between symbols versus the normalized time distance $f_{max}T_s$, zoom-in of Figure 5.9.

The remaining three terms of (5.13) are noise components. If the difference between the amplitudes $a_{i-1,j}$ and a_j is neglected, the power of the second and third term is equal to $2P_sP_n$, where P_s and P_n are signal and noise power, respectively. The power of the last term in (5.13) is equal to the squared noise power P_n^2. For an SNR that is much larger than one, the cross products dominate the squared noise component, and the output SNR can be written as

$$SNR_y = \frac{P_s^2}{P_n^2 + 2P_sP_n} \cong \frac{P_s}{2P_n} = \frac{SNR_x}{2} \tag{5.18}$$

Hence, the SNR after differential detection is approximately 3 dB worse than the input SNR. This 3 dB is the worst case SNR loss of differential detection relative to coherent detection. In practice, coherent detection also has an SNR loss because of imperfect channel estimates and because a part of the signal power is spent on pilots. This typically reduces the difference between differential and coherent detection from 3 to about 1 to 2 dB.

5.3.2 Differential Detection in the Frequency Domain

Differential detection can also be applied across subcarriers instead of symbols. In this case, for a PSK signal with input phases φ_{kj}, the differentially encoded phases θ_{ij} are

$$\theta_{ij} = \sum_{k=0}^{j} \varphi_{ik} \mod(2\pi) \tag{5.19}$$

where i is the symbol number and j and k are subcarrier numbers. Differential detection is now applied to each symbol separately, as depicted in Figure 5.11. The first subcarrier of each symbol is a known pilot value that is needed to provide an initial value to start the differential detection process.

A differential phase detection in the frequency domain is performed by multiplying each FFT output with the conjugated FFT output of the same symbol from the previous subcarrier:

$$y_{ij} = x_{ij}x_{i,j-1}^{*} = a_{ij}a_{i,j-1}e^{\varphi_{ij}+\beta_{ij}-\beta_{i,j-1}} + n_{ij}a_{i,j-1}e^{\theta_{i,j-1}+\beta_{i,j-1}} + n_{i,j-1}a_{ij}e^{\theta_{ij}+\beta_{ij}} + n_{ij}n_{i,j-1}^{*} \tag{5.20}$$

Here, x_{ij} is the FFT output at the receiver as defined by (5.12). Equation (5.20) has exactly the same structure as (5.13), which described differential detection in the time domain. Because of this, a similar signal-to-noise analysis can be made, showing that differential detection in the frequency domain also has an SNR loss of 3 dB compared with ideal coherent detection. The main difference between differential detection in frequency and time is the phase disturbance component in the first term of (5.20). This first term contains the desired phase φ_{kj}, but also an undesired phase $\beta_{ij}-\beta_{i,j-1}$, which is the channel phase shift on symbol i from subcarrier $j-1$ to j.

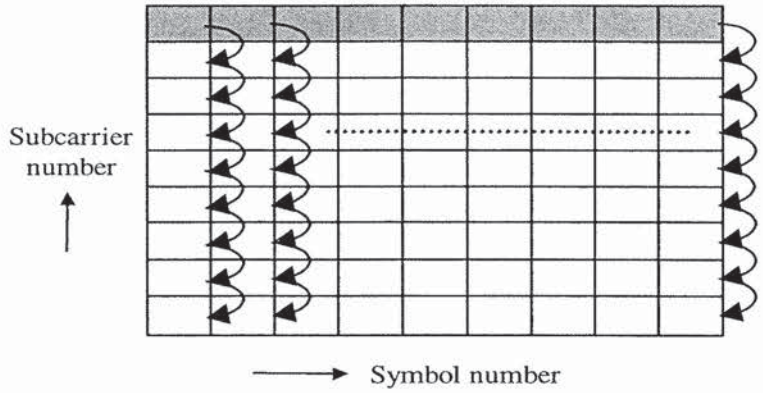


Figure 5.11 Differential detection in the frequency domain. Gray subcarriers are pilots.

The influence of the phase disturbance can be analyzed by looking at the correlation between adjacent subcarriers as a function of the normalized frequency difference τ_{max}/T. This correlation can be calculated from (5.9) by setting the number of subcarriers k to one. A plot of the correlation versus the normalized frequency difference is drawn in Figure 5.12.

Similar to the case of correlation in the time domain as described in the previous section, we can deduce that the SDR for a correlation value r_f is given by

$$SDR = \frac{1}{2(1-r_f)} \tag{5.21}$$

If we use the same example as in the previous section, which required an SDR of 10 dB, then the correlation r_f has to be 0.95. From Figure 5.12 we can deduce that the normalized subcarrier spacing τ_{rms}/T has to be approximately 0.03. This means that the subcarrier spacing has to be $0.03/\tau_{rms}$ at most. Equivalently, it can be stated that the maximum tolerable delay spread is 3% of the FFT period T. The latter means that the delay spread robustness of differential detection in the frequency domain is generally significantly worse than other detection techniques, where the delay spread robustness is related only to the guard time and not to the FFT period.

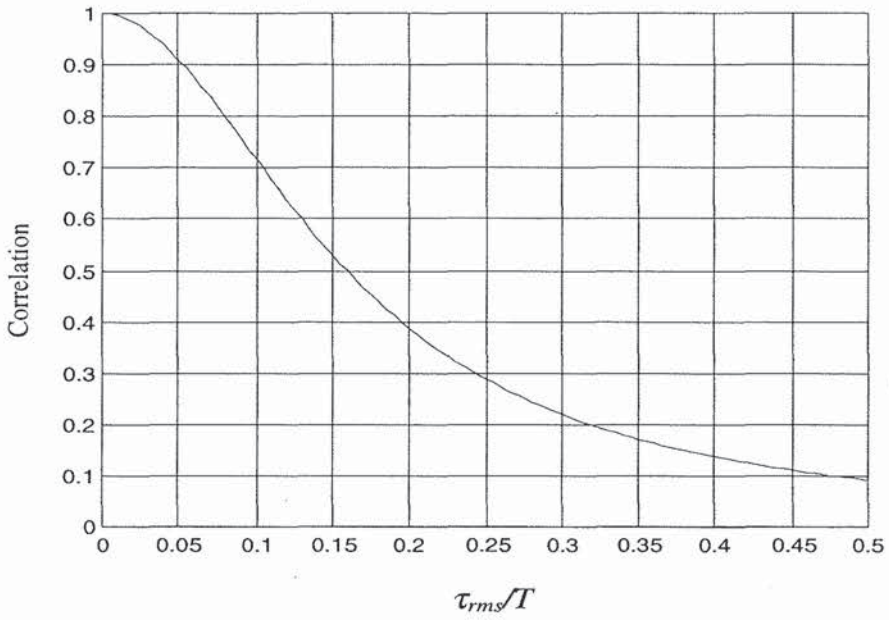


Figure 5.12 Correlation between subcarriers versus the normalized subcarrier spacing τ_{rms}/T.

Figure 5.13 presents some simulation results which compare coherent demodulation and differential detection for an OFDM system for wireless LAN applications. Table 5.5 lists the main system parameters. The data rates for this system are variable, dependent on the coding rate and the modulation type. Some possible data rates for this system are 32 Mbps with 16-QAM and rate ½ coding, or 8-PSK with rate 2/3 coding, 24 Mbps with QPSK and rate 3/4 coding, or 8-PSK with rate ½ coding, and 16 Mbps with QPSK and rate 1/2 coding.

Table 5.5
Main parameters of the simulated OFDM system.

Number of subcarriers	48
OFDM symbol duration	3 μs
Guard interval	600 ns
T_{prefix}: Pre-guard interval	600 ns
$T_{postfix}$: Post-guard interval	75 ns
Subcarrier spacing	416.666 kHz
Roll-off factor β	0.025
Channel spacing	25 MHz
Occupied -3 dB bandwidth	20 MHz

Figure 5.13 shows the irreducible packet-error probabilities versus rms delay spread for an exponentially decaying multipath delay profile. No noise was present in the simulation, so all errors are caused by ISI and ICI because of multipath components with relative delays extending the guard time of the OFDM symbols. To obtain channel estimates for coherent detection, a training symbol was present at the start of each packet. All subcarrier values of this training symbol are known to the receiver, so this is the channel estimation method discussed in Section 5.2.3.

Clearly, coherent demodulation [curves (a) and (b)] performs much better than differential detection in the frequency domain. For the same data rate and packet-error probability, coherent demodulation can tolerate about three times as much delay spread as differential detection. The reason for the relatively poor performance of differential in frequency detection is the significant phase fluctuation between subcarriers. Differential detection in the frequency domain assumes that there is a negligible phase difference between two adjacent subcarriers. For delay spreads around 50 ns, however, a significant percentage of channels show several phase changes exceeding $\pi/8$ within the 48 OFDM subcarriers per symbol. Differential 8-PSK will generate two erroneous subcarriers if the phase changes more than $\pi/8$, and that explains why the error curves for 8-PSK quickly converge to 1 for delay spreads exceeding 50 ns. Differential QPSK is more robust, but still worse than coherent 16-QAM, which operates at twice the data rate. Coherent demodulation is not affected by phase changes across the subcarriers, because it uses training symbols to estimate reference phases and amplitudes of all subcarriers. The same holds for differential detection in the time domain, which will have approximately the same delay spread robustness as coherent detection.

5.3.3 Differential Amplitude and Phase Shift Keying

Traditionally, differential detection is applied to phase-modulated systems only, as it is not obvious how differential detection can be applied to amplitude-modulated systems. It is possible to use differential amplitude and phase shift keying (DAPSK), however, by using a star constellation like depicted in Figure 5.14 [10–12].

The main advantage of DAPSK is that it does not require pilot symbols to estimate a time-varying channel or a remaining local oscillator offset. It only requires a single pilot symbol to initialize the differential detection, as depicted in the Figures 5.8 and 5.11. The disadvantage, however, is a loss in SNR performance, because the minimum distance for the DAPSK constellation is clearly worse than for the corresponding square QAM constellation. Added to this is the loss of doing differential versus coherent detection, although this loss is partly compensated by the fact that DAPSK uses fewer pilot subcarriers.

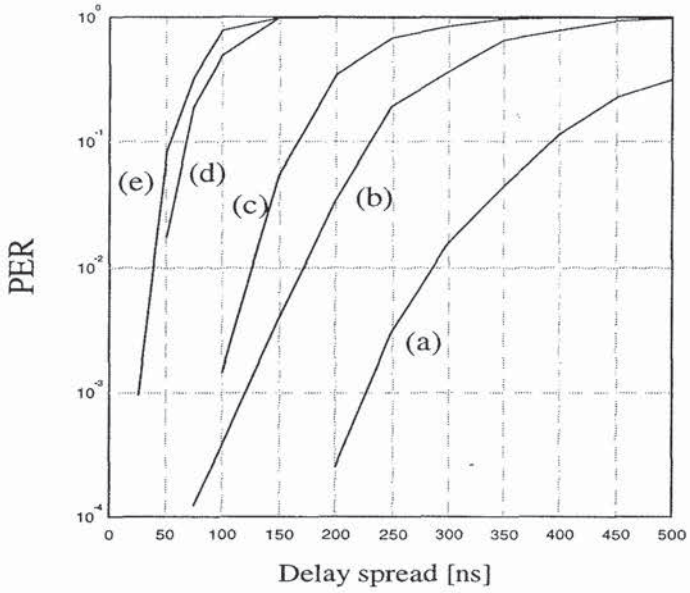


Figure 5.13 Irreducible packet error ratio versus rms delay spread, simulated for an exponentially decaying power delay profile with Rayleigh fading paths. (a) 16 Mbps with coherent QPSK and rate ½ coding, (b) 32 Mbps with coherent 16-QAM and rate ½ coding, (c) 16 Mbps with differential QPSK (in frequency domain) and rate ½ coding, (d) 24 Mbps with differential 8-PSK (in frequency domain) and rate ½ coding, (e) 32 Mbps with differential 8-PSK (in frequency domain) and rate 2/3 coding.

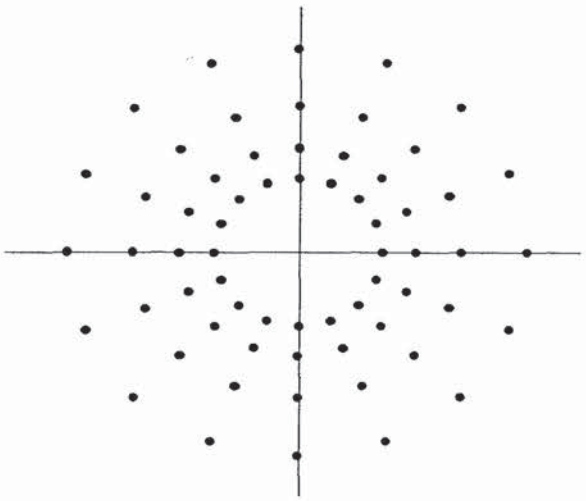


Figure 5.14 64-DAPSK constellation.

REFERENCES

[1] Sandell, M., and O. Edfors, "A Comparative Study of Pilot-Based Channel Estimators for Wireless OFDM," Research Report TULEA 1996:19, Division of Signal Processing, Luleå University of Technology, Sep. 1996.

[2] Edfors, O., M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Börjesson, "OFDM Channel Estimation by Singular Value Decomposition," *Proceedings of the 46th IEEE Vehicular Technology Conference*, pp. 923–927, Apr. 28–May 1, 1996.

[3] Sandell, M., S. K. Wilson, and P. O. Börjesson, "Performance Analysis of Coded OFDM on Fading Channels with Non-Ideal Interleaving and Channel Knowledge," Research Report TULEA 1996:20, Division of Signal Processing, Luleå University of Technology, Sep. 1996.

[4] van de Beek, J. J., O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On Channel Estimation in OFDM Systems," *Proceedings of the 45th IEEE Vehicular Technology Conference*, Rosemont, IL, pp. 715–719, July 1995.

[5] Proakis, J. G., *Digital Communications*, Prentice-Hall, 3rd ed., 1995.

[6] Scharf, L. L., *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*, Addison-Wesley, 1991.

[7] Höher, P., "TCM on Frequency-Selective Land-Mobile Fading Channels," *Proceedings of the Tirrenia International Workshop on Digital Communications*, Tirrenia, Italy, Sep. 1991.

[8] Takanashi, H., and R. van Nee, "Merged Physical Layer Specification for the 5 GHz Band," *IEEE P802.11-98/72-r1*, March 1998.

[9] Frenger, P., and A. Svensson, "Decision directed Coherent Detection in Multicarrier Systems on Rayleigh Fading Channels," *IEEE Trans. on Veh. Technology*, Vol. 48, No. 2, pp. 490–498, 1999.

[10] Engels, V., and H. Rohling, "Multilevel Differential Modulation Techniques (64-DAPSK) for Multicarrier Transmission Techniques," *European Transactions on Telecommunication Related Technologies*, Vol. 6, No. 6, pp. 633–640, Nov.–Dec. 1995.

[11] Engels, V., and H. Rohling, "Differential Modulation Techniques for a 34 Mbit/s Radio Channel Using OFDM," *Wireless Personal Communications*, Vol. 2, pp. 29–44, 1995.

[12] May, T., H. Rohling, and V. Engels, "Performance Analysis of Viterbi Decoding for 64-DAPSK and 64-QAM Modulated Signals," *IEEE Trans. on Comm.*, Vol. 46, No. 2, pp. 182-190, February 1998.

CHAPTER 6

The Peak Power Problem

6.1 INTRODUCTION

An OFDM signal consists of a number of independently modulated subcarriers, which can give a large peak-to-average power (PAP) ratio when added up coherently. When N signals are added with the same phase, they produce a peak power that is N times the average power. This effect is illustrated in Figure 6.1. For this example, the peak power is 16 times the average value. The peak power is defined as the power of a sine wave with an amplitude equal to the maximum envelope value. Hence, an unmodulated carrier has a PAP ratio of 0 dB. An alternative measure of the envelope variation of a signal is the *Crest factor*, which is defined as the maximum signal value divided by the rms signal value. For an unmodulated carrier, the Crest factor is 3 dB. This 3 dB difference between PAP ratio and Crest factor also holds for other signals, provided that the center frequency is large in comparison with the signal bandwidth.

A large PAP ratio brings disadvantages like an increased complexity of the analog-to-digital and digital-to-analog converters and a reduced efficiency of the RF power amplifier. To reduce the PAP ratio, several techniques have been proposed, which basically can be divided in three categories. First, there are signal distortion techniques, which reduce the peak amplitudes simply by nonlinearly distorting the OFDM signal at or around the peaks. Examples of distortion techniques are clipping, peak windowing and peak cancellation. The second category is coding techniques that use a special forward-error correcting code set that excludes OFDM symbols with a large PAP ratio. The third technique is based on scrambling each OFDM symbol with different scrambling sequences and selecting that sequence that gives the smallest PAP ratio. This chapter discusses all of these techniques, but first makes an analysis of the PAP ratio distribution function. This will give a better insight in the PAP problem and will explain why PAP reduction techniques can be quite effective.

ZYXEL-1015
Page 134 (of 277)

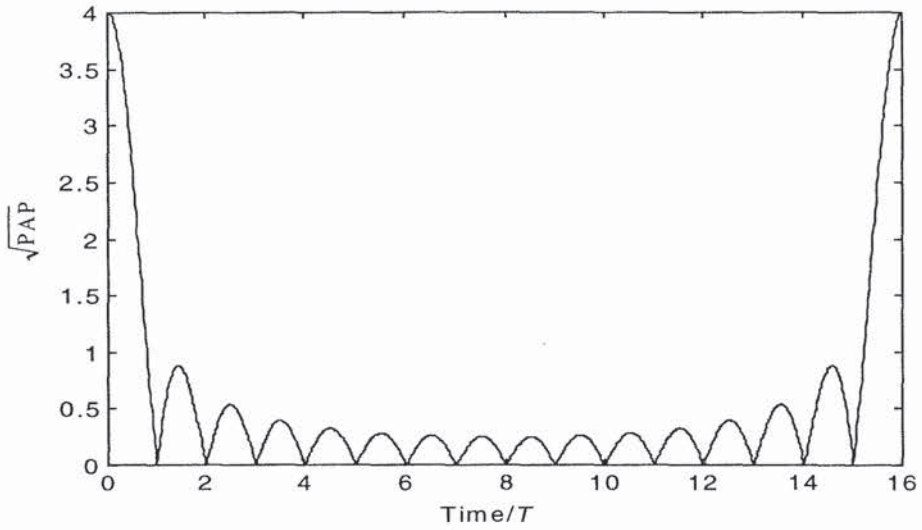


Figure 6.1 Square root of peak-to-average power ratio for a 16-channel OFDM signal, modulated with the same initial phase for all subchannels.

6.2 DISTRIBUTION OF THE PEAK-TO-AVERAGE POWER RATIO

For one OFDM symbol with N subcarriers, the complex baseband signal can be written as

$$x(t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} a_n \exp(j\omega_n t) \tag{6.1}$$

Here, a_n are the modulating symbols. For QPSK, for instance, $a_n \in \{-1,1,j,-j\}$. From the central limit theorem it follows that for large values of N, the real and imaginary values of $x(t)$ become Gaussian distributed, each with a mean of zero and a variance of 1/2. The amplitude of the OFDM signal therefore has a Rayleigh distribution, while the power distribution becomes a central chi-square distribution with two degrees of freedom and zero mean, with a cumulative distribution given by

$$F(z) = 1 - e^{-z} \tag{6.2}$$

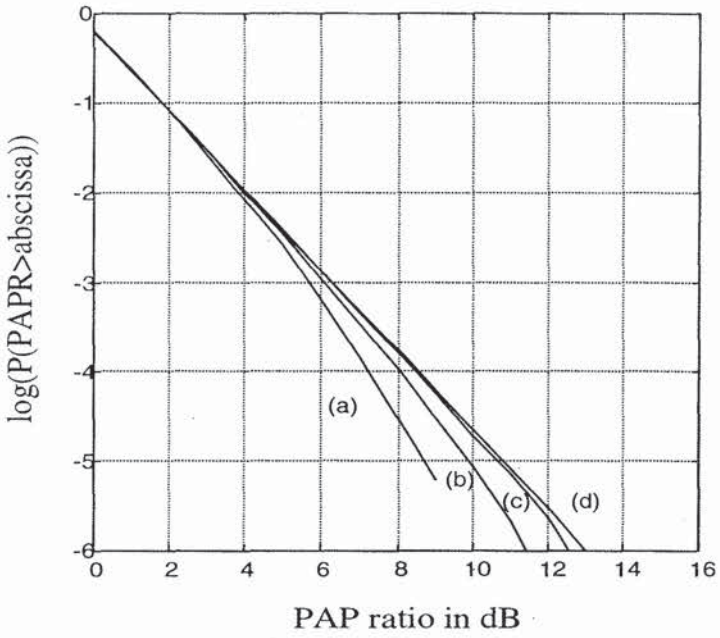


Figure 6.2 PAP distribution of an OFDM signal with (a) 12, (b) 24, (c) 48 and (d) an infinite number of subcarriers (pure Gaussian noise). Four times oversampling used in simulation, total number of simulated samples = 12 million.

Figure 6.2 shows the probability that the PAP ratio exceeds a certain value. We can see that the curves for various numbers of subcarriers are close to a Gaussian distribution (d) until the PAP value comes within a few dB from the maximum PAP level of $10\log N$, where N is the number of subcarriers.

What we want to derive now is the cumulative distribution function for the peak power per OFDM symbol. Assuming the samples are mutually uncorrelated—which is true for non-oversampling—the probability that the PAP ratio is below some threshold level can be written as

$$P(PAPR \le z) = F(z)^N = \left(1 - \exp(-z)\right)^N \tag{6.3}$$

This theoretical derivation is plotted against simulated values in Figure 6.3 for different values of N.

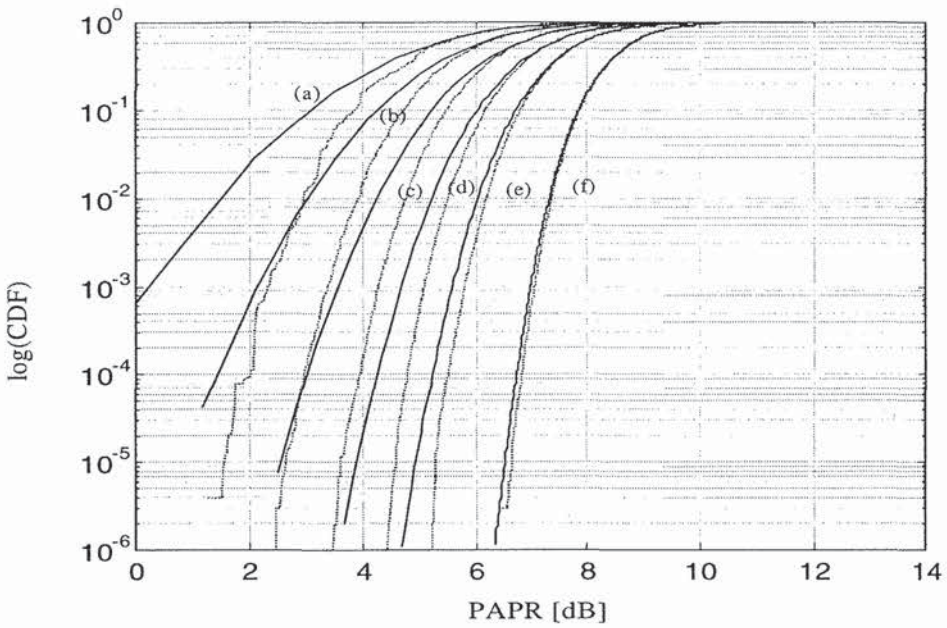


Figure 6.3 PAP distribution without oversampling for a number of subcarriers of (a) 16, (b) 32, (c) 64, (d) 128, (e) 256, and (f) 1024 (dotted lines are simulated).

The assumption made in deriving (6.3) that the samples should be mutually uncorrelated is not true anymore when oversampling is applied. Because it seems quite difficult to come up with an exact solution for the peak power distribution, we propose an approximation by assuming that the distribution for N subcarriers and oversampling can be approximated by the distribution for αN subcarriers without oversampling, with α larger than one. Hence, the effect of oversampling is approximated by adding a certain number of extra independent samples. The distribution of the PAP ratio is then given by

$$P(PAPR \leq z) = \left(1 - \exp(-z)\right)^{\alpha \cdot N} \tag{6.4}$$

In Figure 6.4, the PAP distribution for different amounts of carriers is given for $\alpha = 2.8$. The dotted lines are simulated curves. We see in Figure 6.4 that Equation (6.4) is quite accurate for N>64. For large values of the cumulative distribution function close to one (>0.5), however, (6.3) is actually more accurate.

From Figure 6.4, we can deduce that coding techniques to reduce the PAP ratio may be a viable option, as reasonable coding rates are possible for a PAP ratio around 4 dB. For 64 subcarriers, for instance, about 10^{-6} of all possible QPSK symbols have a PAP ratio of less than 4.2 dB. This means that only 20 out of a total of 128 bits would be lost if only the symbols with a low PAP ratio would be transmitted. However, the main problem with this approach is to find a coding scheme with a reasonable coding rate ($\geq$ 1/2) that produces only these low PAP ratio symbols and that also has