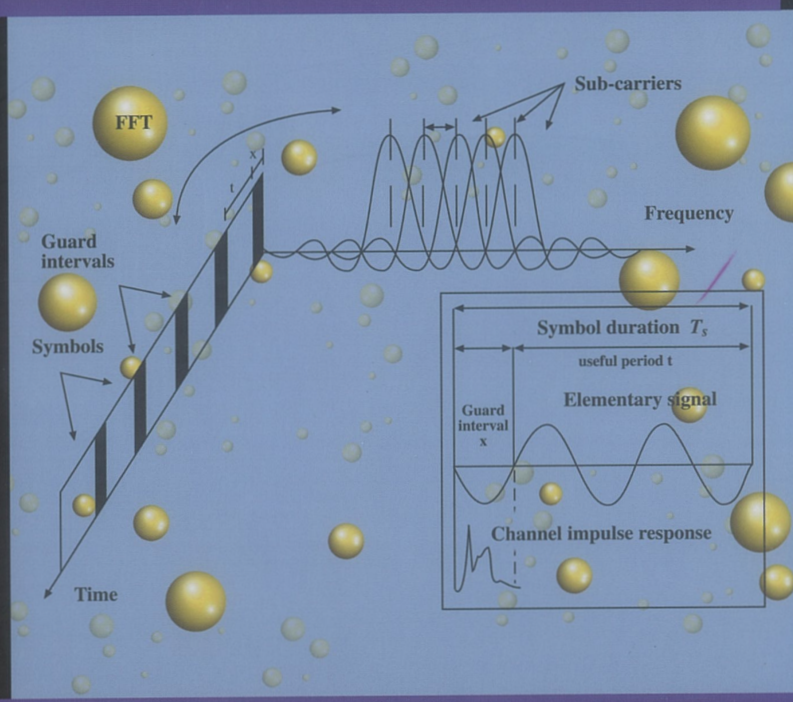
Information Technology: Transmission, Processing and Storage

Multi-Carrier Digital Communications Theory and Applications of OFDM

Second Edition



Ahmad R. S. Bahai | Burton R. Saltzberg | Mustafa Ergen

INTEL-1038
9,614,566

Multi-Carrier Digital Communications

Theory and Applications of OFDM

Second Edition

Multi-Carrier Digital Communications

Theory and Applications of OFDM

Second Edition

Ahmad R. S. Bahai
National Semiconductor
Stanford University
University of California Berkeley

Burton R. Saltzberg
Consultant on Digital Communications
Formerly Bell Laboratories

Mustafa Ergen
University of California Berkeley

Springer

Library of Congress Control Number: 2004052575

Bahai, Ahmad R. S.,
Multi-Carrier Digital Communications: Theory and Applications of OFDM/
Ahmad R. S. Bahai, Burton R. Saltzberg, Mustafa Ergen – 2nd ed.

p. cm. (Information Technology: Transmission, Processing and Storage)
1. Digital Communications. 2. Multiplexing. 3. Spread spectrum communications.
4. Orthogonalization methods. I. Saltzberg, Burton R. II. Ergen, Mustafa.
III. Title. IV. Series.

ISBN: 978-0-387-22575-3 (13) ISBN: 0-387-22575-7 (10) E- ISBN: 978-0-387-22576-0

Printed on acid-free paper

Copyright © 2004 by Springer Science +Business Media, LLC
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science +Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

9 8 7 6 5 4 3

springer.com

Contents

Preface XVII

Acknowledgements XXI

1 Introduction to Digital Communications **1**

1.1 Background 1

1.2 Evolution of OFDM 5

2 System Architecture **15**

2.1 Wireless Channel Fundamentals 15

2.1.1 Path Loss 16

2.1.2 Shadowing 17

2.1.3 Fading Parameters 18

2.1.4 Flat Fading 21

2.1.5 Frequency Selective Fading 22
2.1.6 Fast Fading 23
2.1.7 Slow Fading 23
2.1.8 Rayleigh Fading 23
2.1.9 Ricean Fading 25
2.1.10 Distortions for Wireless Systems 26
2.1.11 Diversity Techniques in a Fading Environment 26
2.2 Digital Communication System Fundamentals 27
2.2.1 Coding 28
2.2.2 Modulation 29
2.3 Multi-Carrier System Fundamentals 33
2.4 DFT 35
2.5 Partial FFT 41
2.6 Cyclic Extension 42
2.7 Channel Estimation 45
2.8 Modelling of OFDM for Time-Varying Random Channel 48
2.8.1 Randomly Time-Varying Channels 48
2.8.2 OFDM in Randomly Time-Varying Channels 51

3 Performance over Time-Invariant Channels 55
3.1 Time-Invariant Non-Flat Channel with Colored Noise 55
3.2 Error Probability 56
3.3 Bit Allocation 59

3.4 Bit and Power Allocation Algorithms for Fixed Bit Rate 66

4 Clipping in Multi-Carrier Systems 69

4.1 Introduction 69

4.2 Power Amplifier Non-Linearity 71

4.3 Error Probability Analysis 73

4.3.1 System Model 77

4.3.2 BER Due to Clipping 78

4.4 Performance in AWGN and Fading 86

4.5 Bandwidth Regrowth 94

5 Synchronization 99

5.1 Timing and Frequency Offset in OFDM 99

5.2 Synchronization & System Architecture 104

5.3 Timing and Frame Synchronization 105

5.4 Frequency Offset Estimation 107

5.5 Phase Noise 108

6 Channel Estimation and Equalization 117

6.1 Introduction 117

6.2 Channel Estimation 117

6.2.1 Coherent Detection 119

6.2.2 Block-Type Pilot Arrangement 122

6.2.3 Comb-Type Pilot Arrangement 126

6.2.4 Interpolation Techniques 127

6.2.5 Non-coherent Detection 130
6.2.6 Performance 130
6.2.7 Channel Estimation for MIMO-OFDM . . . 134
6.3 Equalization 137
6.3.1 Time Domain Equalization 138
6.3.2 Equalization in DMT 142
6.3.3 Delay Parameter 145
6.3.4 AR Approximation of ARMA Model 146
6.3.5 Frequency Domain Equalization 149
6.3.6 Echo Cancellation 152
6.3.7 Appendix - Joint Innovation Representation of ARMA Models 160

7 Channel Coding **167**
7.1 Need for Coding 167
7.2 Block Coding in OFDM 168
7.3 Convolutional Encoding 173
7.4 Concatenated Coding 179
7.5 Trellis Coding in OFDM 180
7.6 Turbo Coding in OFDM 185

8 ADSL **189**
8.1 Wired Access to High Rate Digital Services 189
8.2 Properties of the Wire-Pair Channel 190

8.3 ADSL Systems 200

9 Wireless LAN Applications 203

9.1 Introduction 203

9.1.1 Background 203

9.1.2 Pros and Cons 204

9.2 Topology 206

9.2.1 Independent BSS 206

9.2.2 Infrastructure BSS 208

9.2.3 Services 211

9.3 Architecture 213

9.4 Medium Access Control 217

9.4.1 DCF Access 221

9.4.2 Markov Model of DCF 227

9.4.3 PCF 233

9.5 Management 239

9.5.1 Synchronization 239

9.5.2 Scanning 240

9.5.3 Power Management 241

9.5.4 Security 242

9.6 IEEE 802.11 and 802.11b Physical Layer 244

9.6.1 Spread Spectrum 245

9.6.2 FHSS Physical Layer 247

9.6.3 DSSS Physical Layer 249

9.6.4 IR Physical Layer 252
9.6.5 HR/DSSS Physical Layer 252
9.6.6 RF Interference 253
9.7 IEEE 802.11a Physical Layer 254
9.7.1 OFDM Architecture 255
9.7.2 Transmitter 256
9.7.3 Receiver 269
9.7.4 Degradation Factors 274
9.8 Performance of 802.11a Transceivers 275
9.8.1 Data Rate 276
9.8.2 Phase Noise 278
9.8.3 Channel Estimation 279
9.8.4 Frequency Offset 280
9.8.5 IQ Imbalance 281
9.8.6 Quantization and Clipping Error 283
9.8.7 Power Amplifier Nonlinearity 285
9.8.8 Hard or Soft Decision Decoding 286
9.8.9 Co-channel Interference 286
9.8.10 Narrowband Interference 288
9.8.11 UWB Interference 288
9.8.12 Performance of 64QAM 289
9.9 Rate Adaptation 291
9.10 Zero IF Technology 292

9.11 IEEE 802.11e MAC Protocol 293
9.11.1 IEEE 802.11e MAC Services 294
9.11.2 IEEE 802.11e MAC Architecture 295
9.11.3 Hybrid Coordination Function (HCF) 295
9.11.4 HCF Controlled Channel Access 299
9.11.5 Admission Control 300
9.11.6 Block Acknowledgement 301
9.11.7 Multi-rate Support 301
9.11.8 Direct Link Protocol 301
9.12 HIPERLAN/2 302
9.12.1 Protocol Architecture 304
9.12.2 Data Link Control 304
9.12.3 Convergence Layer 310
9.12.4 HIPERLAN/2 vs 802.11a 310
9.13 MMAC-HiSWAN 312
9.14 Overview of IEEE 802.11 Standards 312
9.14.1 IEEE 802.11c - Bridge Operation Procedures 312
9.14.2 IEEE 802.11d - Global Harmonization 313
9.14.3 IEEE 802.11f - Inter Access Point Protocol . 313
9.14.4 IEEE 802.11g - Higher Rate Extensions in the 2.4GHz Band 313
9.14.5 IEEE 802.11h - Spectrum Managed 802.11a . 314
9.14.6 IEEE 802.11i - MAC Enhancements for Enhanced Security 315

9.14.7 IEEE 802.1x - Port Based Network Access Control 316

9.14.8 IEEE 802.1p - QoS on the MAC Level 316

10 Digital Broadcasting **317**

10.1 Broadcasting of Digital Audio Signals 317

10.2 Signal Format 320

10.3 Other Digital Broadcasting Systems 323

10.3.1 DAB in the U.S.A 323

10.4 Digital Video Broadcasting 324

11 OFDM based Multiple Access Techniques **327**

11.1 Introduction 327

11.2 OFDM-FDMA 329

11.3 OFDM-TDMA 330

11.4 Multi Carrier CDMA (OFDM-CDMA) 331

11.5 OFDMA 336

11.5.1 OFDMA Architecture 337

11.5.2 Resource Allocation Regarding QoS 341

11.5.3 Resource Allocation Regarding Capacity . . 342

11.6 Flash-OFDM 343

11.7 OFDM-SDMA 346

12 Ultra WideBand Technologies **349**

12.1 Impulse Radio 350

12.2 Multiband Approach 352
12.3 Multiband OFDM 352

13 IEEE 802.16 and WiMAX **357**
13.1 Introduction 357
13.2 WiMAX 358
13.3 WirelessMAN OFDM 359
13.4 802.16 MAC 361
13.5 Conclusion 362

14 Future Trends **363**
14.1 Comparison with Single Carrier Modulation 363
14.2 Mitigation of Clipping Effects 365
14.3 Overlapped Transforms 366
14.4 Advances in Implementation 370

Bibliography **373**

List of Figures **393**

List of Tables **405**

Index **407**

We dedicate this book
to our families...

Preface

Multi-carrier modulation, Orthogonal Frequency Division Multiplexing (OFDM) particularly, has been successfully applied to a wide variety of digital communications applications over the past several years. Although OFDM has been chosen as the physical layer standard for a diversity of important systems, the theory, algorithms, and implementation techniques remain subjects of current interest. This is clear from the high volume of papers appearing in technical journals and conferences.

Multi-carrier modulation continues to evolve rapidly. It is hoped that this book will remain a valuable summary of the technology, providing an understanding of new advances as well as the present core technology.

The Intended Audience

This book is intended to be a concise summary of the present state of the art of the theory and practice of OFDM technology. The authors believe that the time is ripe for such a treatment. Particularly based on one of the author's long experience in development of wireless systems (AB), and the other's in wireline systems (BS), we have attempted to present a unified presentation of OFDM performance and

implementation over a wide variety of channels.

It is hoped that this will prove valuable both to developers of such systems and to researchers and graduate students involved in analysis of digital communications.

In the interest of brevity, we have minimized treatment of more general communication issues. There exist many excellent texts on communication theory and technology. Only brief summaries of topics not specific to multi-carrier modulation are presented in this book where essential. As a background, we presume that the reader has a clear knowledge of basic fundamentals of digital communications.

Highlights of the Second Edition

During the past few years since the publication of the first edition of this text, the technology and application of OFDM has continued their rapid pace of advancement. As a result, it became clear to us that a new edition of the text would be highly desirable. The new edition provides an opportunity to make those corrections and clarifications whose need became apparent from continued discussions with many readers. However, the main purpose is to introduce new topics that have come to the forefront during the past few years, and to amplify the treatment of other subject matter.

Because of the particularly rapid development of wireless systems employing OFDM, we have introduced a section early in the text on wireless channel fundamentals. We have extended and modified our analysis of the effects of clipping, including simulation results that have been reported in a recent publication. These new results are restated here. A section on channel estimation has been added to the chapter on equalization. The chapter on local area networks has been greatly expanded to include the latest technology and applications. Three totally new chapters are added, on OFDM multiple access tech-

nology, on ultra wideband technology and on WiMAX (IEEE 802.16).

Organization of This Book

We begin with a historical overview of multi-carrier communications, wherein its advantages for transmission over highly dispersive channels have long been recognized, particularly before the development of equalization techniques. We then focus on the bandwidth efficient technology of OFDM, in particular the digital signal processing techniques that have made the modulation format practical. Several chapters describe and analyze the sub-systems of an OFDM implementation, such as clipping, synchronization, channel estimation, equalization, and coding. Analysis of performance over channels with various impairments is presented.

The book continues with descriptions of three very important and diverse applications of OFDM that have been standardized and are now being deployed. ADSL provides access to digital services at several Mbps over the ordinary wire-pair connection between customers and the local telephone company central office. Digital Broadcasting enables the radio reception of high quality digitized sound and video. A unique configuration that is enabled by OFDM is the simultaneous transmission of identical signals by geographically dispersed transmitters. And, the new development of wireless LANs for multi-Mbps communications is presented in detail. Each of these successful applications required the development of new fundamental technology.

Finally, the book concludes with describing the OFDM based multiple access techniques, ultra wideband technology and WiMAX.

Acknowledgements

The two authors of the first edition of this text are very pleased to include our colleague as an additional author, and gratefully acknowledge his extensive contributions in making this second edition possible.

The first edition of this text has been used for classes in University of California Berkeley, Stanford University, University of Cambridge, CEI-Europe and in other institutions. We are grateful to colleagues in the institutions where this book has been used. For the first edition, we acknowledge the extensive review and many valuable suggestions of Professor Kenji Kohiyama, our former colleagues at AT&T Bell Laboratories and colleagues at Algorex. Gail Bryson performed the very difficult task of editing and assembling this text. The continuing support of Kambiz Homayounfar was essential to its completion.

In preparing the second edition, we acknowledge Professor Pravin Varaiya for his valuable help, Dr. Haiyun Tangh for providing some graphs, National Semiconductor and IMEC for providing the MATLAB simulation tool, Sinem Coleri for proof-reading. We wish to express our appreciation to Fran Wilkinson for editing.

Last, but by no means least, we are thankful to our families for their support and patience.

Despite all our efforts to keep the text error free, for any that remain, any comments, corrections and suggestions received will be much appreciated for the future printings. We can be reached via e-mail at *bahai@stanford.edu, bsaltzberg@worldnet.att.net*, and *ergen@eecs.berkeley.edu*. We will post any corrections and comments at the Web site ***http://ofdm.eecs.berkeley.edu/*** in addition to the support materials that may be necessary to prepare a lecture or paper.

Chapter 1

Introduction to Digital Communications

1.1 Background

The physical layer of digital communications includes mapping of digital input information into a waveform for transmission over a communication channel and mapping of the received waveform into digital information that hopefully agrees with the original input since the communication channel may introduce various forms of distortion as well as noise [3].

The simplest form of such communication, as least conceptually, is Pulse Amplitude Modulation (PAM), shown in Figure 1.1. Here the transmitted waveform is of the form

$$s(t) = \sum_n a_n g(t - nT) \tag{1.1}$$

where the information to be transmitted is given by the sequence of a_ns, $1/T$ is the symbol rate, and $g(t)$ is the impulse response of the

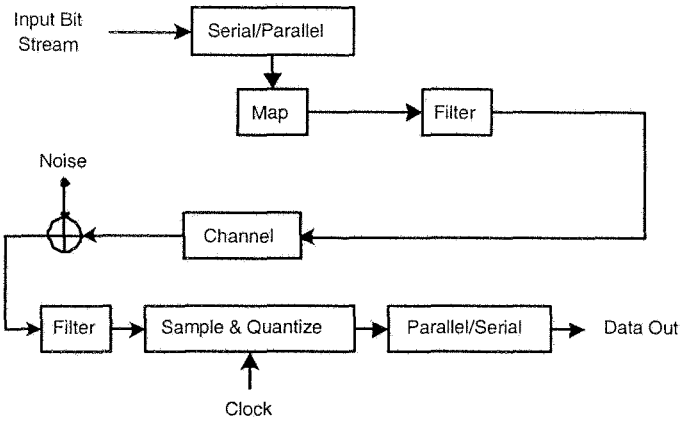


Figure 1.1: A basic PAM system

transmit filter, usually low-pass. The a_ns are chosen from an alphabet of size L, so the bit rate is $\frac{1}{T}log_2 L$. It is desirable that the alphabet be both zero mean and equally spaced. The values of a_n can be written as

$$\{-A(L-1), \ldots, -A, A, \ldots, A(L-1)\}. \tag{1.2}$$

Assuming the a_ns are equiprobable, the transmitted power is

$$A^2 \frac{L^2-1}{3T} \int_{-\infty}^{\infty} g^2(t)dt. \tag{1.3}$$

At the receiver, the signal is filtered by $r(t)$, which may include an adaptive equalizer (sampled), and the nearest permitted member of the alphabet is output. In order to avoid inter-symbol interference, it is desirable that $x(t) = g(t) * h(t) * r(t) = 0$ for all $t = kT$, k an integer $\neq 0$, where $h(t)$ is the channel impulse response. This is the Nyquist criterion, which is given in the frequency domain by:

$$\sum_m X(f + \frac{m}{T}) = const \tag{1.4}$$

The minimum bandwidth required is $1/2T$. This is met by a frequency response that is constant for $-1/2T < f < 1/2T$, whose corresponding time response is

$$x(t) = \frac{sin(\pi t/T)}{(\pi t/T)} \tag{1.5}$$

Some excess bandwidth, denoted by the roll-off factor, is desirable in order for the time response to decay more quickly. Note that $r(t)$ is not a matched filter, because it must satisfy the inter-symbol interference constraint.

If $r(t)$ has gain such that the alphabet levels of $x(0)$ are also spaced by $2A$, then errors will occur when the noise at the sampler satisfies $|n| > A$ for interior levels, or $n > A$ or $n < -A$ for the outer levels. If the noise is Gaussian with power spectral density $N(f)$ at the receiver input, then the noise variance is:

$$\sigma^2 = \int_{-\infty}^{\infty} N(f)|R(f)|^2 df \tag{1.6}$$

and the error probability per symbol is

$$P_e = \frac{2(L-1)}{L} Q(\frac{A}{\sigma}), \tag{1.7}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-y^2/2} dy \tag{1.8}$$

is the normal error integral.

PAM is only suitable over channels that exist down to, but might not necessarily include, zero frequency. If zero frequency is absent, a modulation scheme that puts the signal spectrum in the desired frequency band is required. Of particular interest, both in its own right

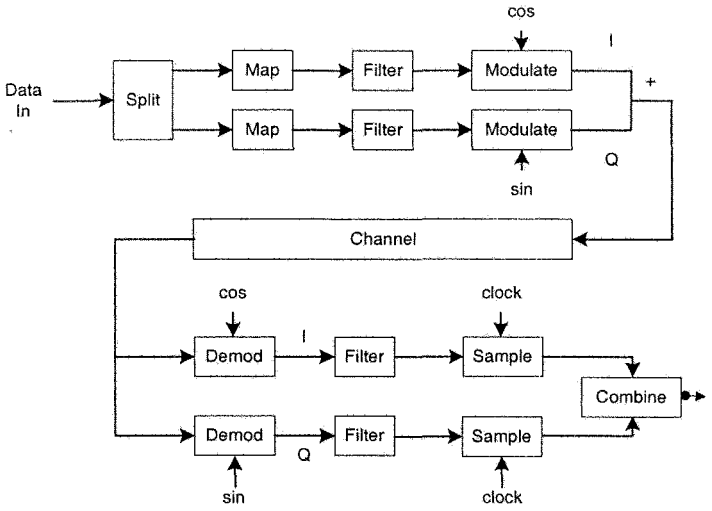


Figure 1.2: A basic QAM system

and as a component of OFDM, is Quadrature Amplitude Modulation (QAM). The simplest form of QAM, shown in Figure 1.2, may be thought of as two PAM signals, modulated by carriers at the same frequency but 90 degrees out of phase. At the receiver, demodulation by the same carriers separates the signal components. Unlike some other modulation schemes, such as FM, QAM is bandwidth efficient in that it requires the same bandwidth as a PAM signal of the same bit rate. Furthermore, the performance of QAM in noise is comparable to that of PAM. The QAM line signal is of the form

$$\sum_n a_n g(t - nT) cos\omega t - \sum_n b_n g(t - nT) sin\omega t. \qquad (1.9)$$

This line signal may also be written in the form of:

$$Re\{\sum_n c_n g(t - nT) e^{j\omega t}\}, \qquad (1.10)$$

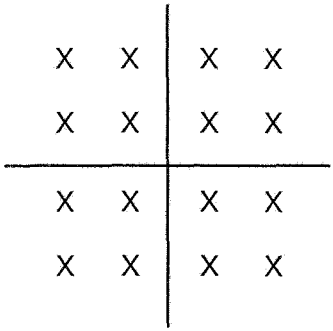


Figure 1.3: A QAM constellation

where the pair of real symbols a_n and b_n are treated as a complex symbol $c_n = a_n + jb_n$. The required bandwidth for transmitting such complex symbols is $1/T$. The complex symbol values are shown as a "constellation" in the complex plane. Figure 1.3 shows the constellation of a 16-point QAM signal, which is formed from 4-point PAM.

It is not necessary that the constellation be square. Figure 1.4 shows how input information can be mapped arbitrarily into constellation points. A constellation with a more circular boundary provides better noise performance. By grouping n successive complex symbols as a unit, we can treat such units as symbols in $2n$-dimensional space. In this case, Figure 1.4 can be extended to include a large enough serial-to-parallel converter that accommodates the total number of bits in n symbols, and a look-up table with $2n$ outputs.

1.2 Evolution of OFDM

The use of Frequency Division Multiplexing (FDM) goes back over a century, where more than one low rate signal, such as telegraph, was carried over a relatively wide bandwidth channel using a separate car-

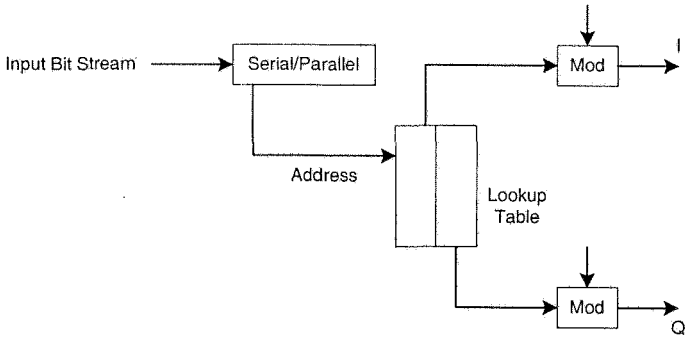


Figure 1.4: General form of QAM generation

rier frequency for each signal. To facilitate separation of the signals at the receiver, the carrier frequencies were spaced sufficiently far apart so that the signal spectra did not overlap. Empty spectral regions between the signals assured that they could be separated with readily realizable filters. The resulting spectral efficiency was therefore quite low.

Instead of carrying separate messages, the different frequency carriers can carry different bits of a single higher rate message. The source may be in such a parallel format, or a serial source can be presented to a serial-to-parallel converter whose output is fed to the multiple carriers.

Such a parallel transmission scheme can be compared with a single higher rate serial scheme using the same channel. The parallel system, if built straightforwardly as several transmitters and receivers, will certainly be more costly to implement. Each of the parallel subchannels can carry a low signalling rate, proportional to its bandwidth. The sum of these signalling rates is less than can be carried by a single serial channel of that combined bandwidth because of the unused guard space between the parallel sub-carriers. On the other hand, the

single channel will be far more susceptible to inter-symbol interference. This is because of the short duration of its signal elements and the higher distortion produced by its wider frequency band, as compared with the long duration signal elements and narrow bandwidth in sub-channels in the parallel system.

Before the development of equalization, the parallel technique was the preferred means of achieving high rates over a dispersive channel, in spite of its high cost and relative bandwidth inefficiency. An added benefit of the parallel technique is reduced susceptibility to most forms of impulse noise.

The first solution of the bandwidth efficiency problem of multitone transmission (not the complexity problem) was probably the *Kineplex* system. The Kineplex system was developed by Collins Radio Co. [4] for data transmission over an H.F. radio channel subject to severe multi-path fading. In that system, each of 20 tones is modulated by differential 4-PSK without filtering. The spectra are therefore of the $sin(kf)/f$ shape and strongly overlap. However, similar to modern OFDM, the tones are spaced at frequency intervals almost equal to the signalling rate and are capable of separation at the receiver.

The reception technique is shown in Figure 1.5. Each tone is detected by a pair of tuned circuits. Alternate symbols are gated to one of the tuned circuits, whose signal is held for the duration of the next symbol. The signals in the two tuned circuits are then processed to determine their phase difference, and therefore the transmitted information. The older of the two signals is then quenched to allow input of the next symbol. The key to the success of the technique is that the time response of each tuned circuit to all tones, other than the one to which it is tuned, goes through zero at the end of the gating interval, at which point that interval is equal to the reciprocal of the frequency separation between tones. The gating time is made somewhat shorter than the symbol period to reduce inter-symbol interference, but efficiency of 70% of the Nyquist rate is achieved. High performance over actual long H.F. channels was obtained, although at a high im-

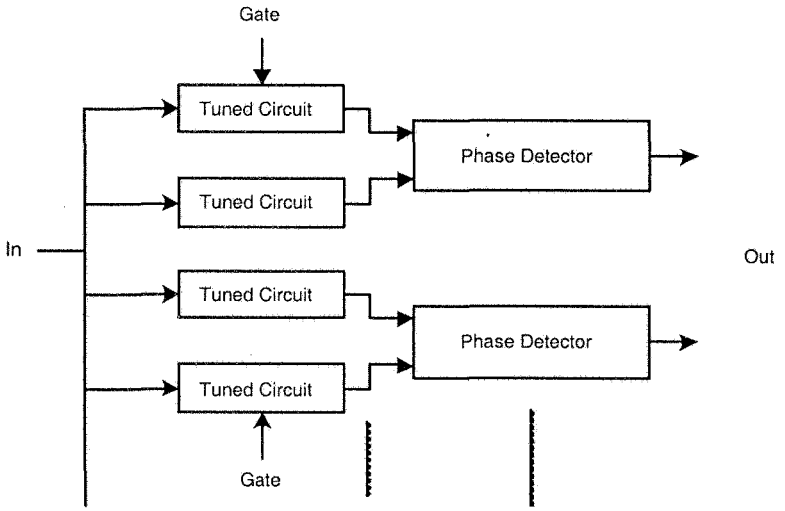


Figure 1.5: The Collins Kineplex receiver

plementation cost. Although fully transistorized, the system required two large bays of equipment.

A subsequent multi-tone system [5] was proposed using 9-point QAM constellations on each carrier, with correlation detection employed in the receiver. Carrier spacing equal to the symbol rate provides optimum spectral efficiency. Simple coding in the frequency domain is another feature of this scheme. The above techniques do provide the orthogonality needed to separate multi-tone signals spaced by the symbol rate. However the $sin(kf)/f$ spectrum of each component has some undesirable properties. Mutual overlap of a large number of sub-channel spectra is pronounced. Also, spectrum for the entire system must allow space above and below the extreme tone frequencies to accommodate the slow decay of the sub-channel spectra. For these reasons, it is desirable for each of the signal components to be bandlimited so as to overlap only the immediately adjacent sub-carriers, while remaining orthogonal to them. Criteria for meeting this objec-

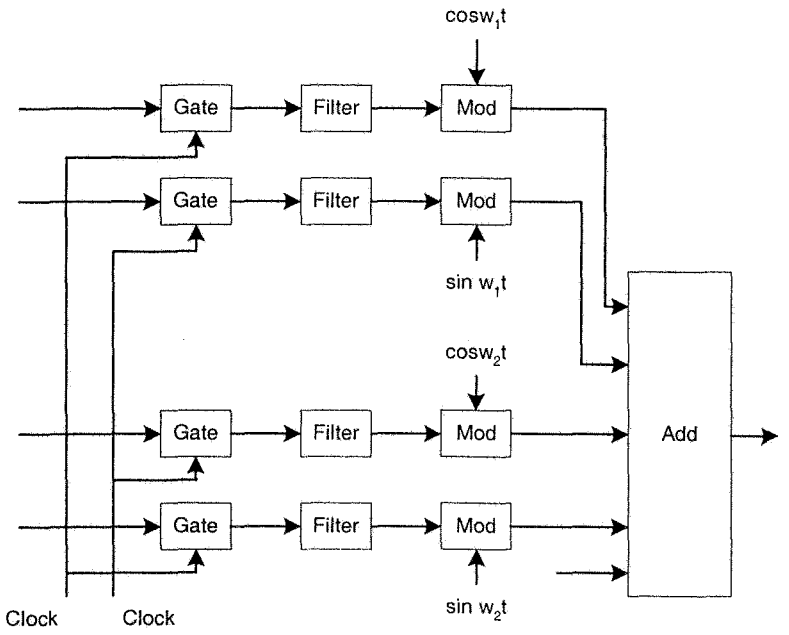


Figure 1.6: An early version of OFDM

tive is given in References [6] and [7].

In Reference [8] it was shown how bandlimited QAM can be employed in a multi-tone system with orthogonality and minimum carrier spacing (illustrated in Figure 1.6). Unlike the non-bandlimited OFDM, each carrier must carry Staggered (or Offset) QAM, that is, the input to the I and Q modulators must be offset by half a symbol period. Furthermore, adjacent carriers must be offset oppositely. It is interesting to note that Staggered QAM is identical to Vestigial Sideband (VSB) modulation. The low-pass filters $g(t)$ are such that $G^2(f)$, the combination of transmit and receive filters, is Nyquist, with the roll-off factor assumed to be less than 1.

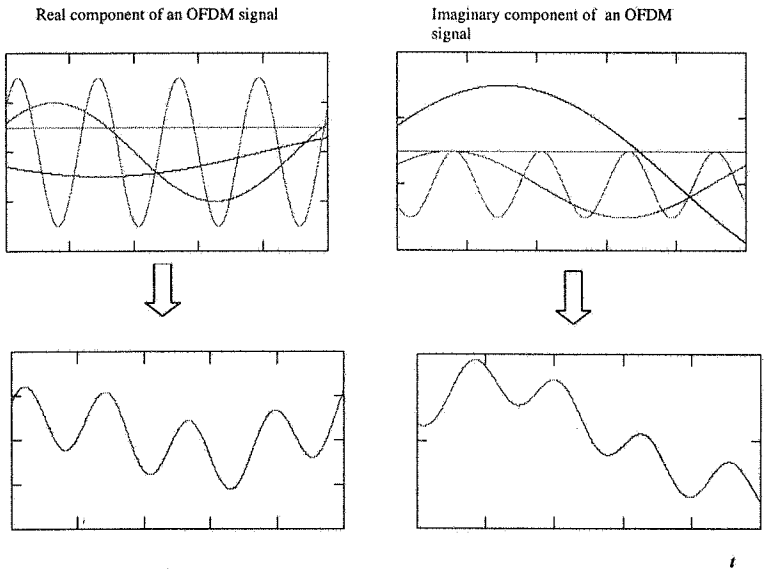


Figure 1.7: OFDM modulation concept: Real and Imaginary components of an OFDM symbol is the superposition of several harmonics modulated by data symbols

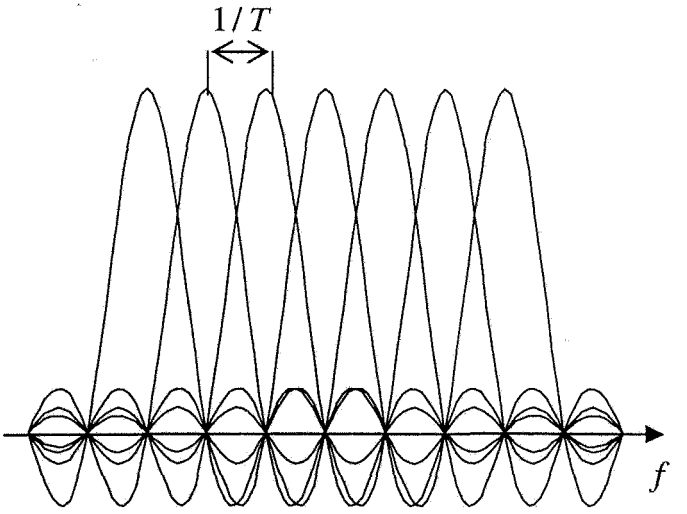


Figure 1.8: Spectrum overlap in OFDM

The major contribution to the OFDM complexity problem was the application of the Fast Fourier Transform (FFT) to the modulation and demodulation processes [9]. Fortunately, this occurred at the same time digital signal processing techniques were being introduced into the design of modems. The technique involved assembling the input information into blocks of N complex numbers, one for each sub-channel. An inverse FFT is performed on each block, and the resultant transmitted serially. At the receiver, the information is recovered by performing an FFT on the received block of signal samples. This form of OFDM is often referred to as Discrete Multi-Tone (DMT). The spectrum of the signal on the line is identical to that of N separate QAM signals, at N frequencies separated by the signalling rate. Each such QAM signal carries one of the original input complex numbers. The spectrum of each QAM signal is of the form $sin(kf)/f$, with nulls at the center of the other sub-carriers, as in the earlier OFDM systems, and as shown in Figure 1.8 and Figure 1.9.

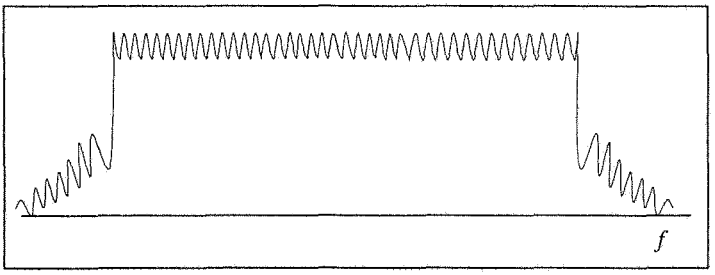


Figure 1.9: Spectrum of OFDM signal

A block diagram of a very basic DMT system is shown Figure 1.10. Several critical blocks are not shown. As described more thoroughly in Chapter 2, care must be taken to avoid overlap of consecutive transmitted blocks, a problem that is solved by the use of a cyclic prefix. Another issue is how to transmit the sequence of complex numbers from the output of the inverse FFT over the channel. The process is straightforward if the signal is to be further modulated by a modulator with I and Q inputs.

Otherwise, it is necessary to transmit real quantities. This can be accomplished by first appending the complex conjugate to the original input block. A $2N$-point inverse FFT now yields $2N$ real numbers to be transmitted per block, which is equivalent to N complex numbers.

The most significant advantage of this DMT approach is the efficiency of the FFT algorithm. An N-point FFT requires only on the order of $NlogN$ multiplications, rather than N^2 as in a straightforward computation. The efficiency is particularly good when N is a power of 2, although that is not generally necessary. Because of the use of the FFT, a DMT system typically requires fewer computations per unit time than an equivalent single channel system with equalization. An overall cost comparison between the two systems is not as clear, but the costs should be approximately equal in most cases. It should be noted that the bandlimited system of Figure 1.6 can also

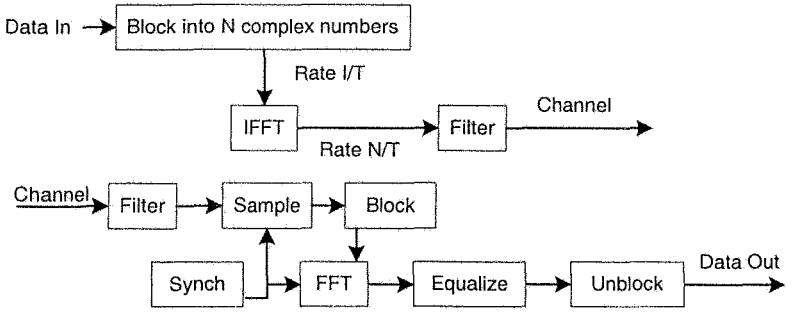


Figure 1.10: Very basic OFDM system

be implemented with FFT techniques [10], although the complexity and delay will be greater than DMT. Over the last 20 years or so, OFDM techniques and, in particular, the DMT implementation, has been used in a wide variety of applications [64]. Several OFDM voice-band modems have been introduced, but did not succeed commercially because they were not adopted by standards bodies. DMT has been adopted as the standard for the Asymmetric Digital Subscriber Line (ADSL), which provides digital communication at several Mbps from a telephone company central office to a subscriber, and a lower rate in the reverse direction, over a normal twisted pair of wires in the loop plant. OFDM has been particularly successful in numerous wireless applications, where its superior performance in multi-path environments is desirable. Wireless receivers detect signals distorted by time and frequency selective fading. OFDM in conjunction with proper coding and interleaving is a powerful technique for combating the wireless channel impairments that a typical OFDM wireless system might face, as is shown in Figure 1.11. A particularly interesting configuration, discussed in Chapter 10, is the Single Frequency Network (SFN) used for broadcasting of digital audio or video signals. Here many geographically separated transmitters broadcast identical and synchronized signals to cover a large region. The reception of such signals by a receiver is equivalent to an extreme form of multi-

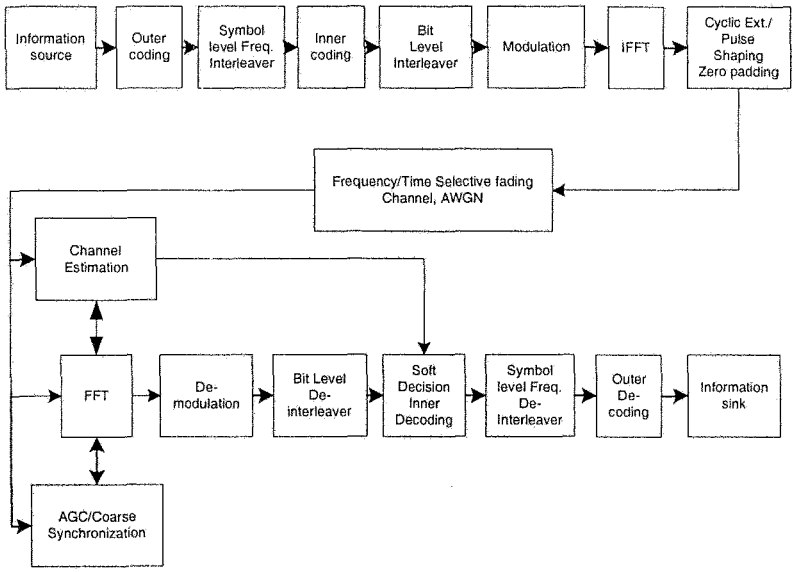


Figure 1.11: A typical wireless OFDM architecture

path. OFDM is the technology that makes this configuration viable.

Another wireless application of OFDM is in high speed local area networks (LANs). Although the absolute delay spread in this environment is low, if very high data rates, in the order of many tens of Mbps, is desired, then the delay spread may be large compared to a symbol interval. OFDM is preferable to the use of long equalizers in this application. It is expected that OFDM will be applied to many more new communications systems over the next several years.

Chapter 2

System Architecture

This chapter presents a general overview of system design for multicarrier modulation. First, a review of wireless channels are discussed. Second, fundamentals of digital communication is cited and finally, the major system blocks of OFDM is analyzed.

2.1 Wireless Channel Fundamentals

Wireless transmission uses air or space for its transmission medium. The radio propagation is not as smooth as in wire transmission since the received signal is not only coming directly from the transmitter, but the combination of reflected, diffracted, and scattered copies of the transmitted signal. It is interesting and rewarding to examine the effects of propagation to a radio signal since consequences determine data rate, range, and reliability of the wireless system.

Reflection occurs when the signal hits a surface where partial energy is reflected and the remaining is transmitted into the surface. Reflection coefficient, the coefficient that determines the ratio of reflec-

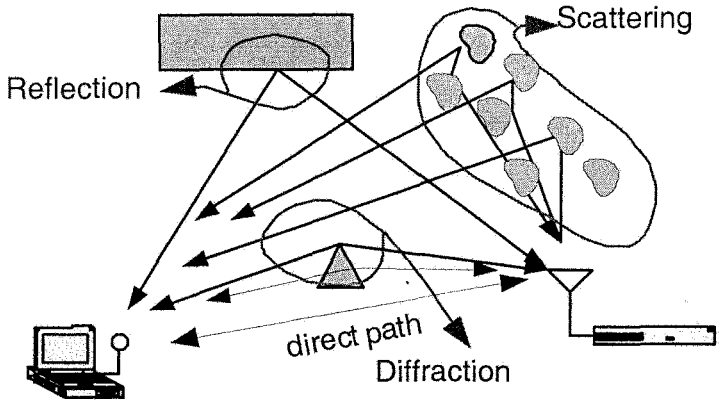


Figure 2.1: Wireless propagation

tion and transmission, depends on the material properties. Diffraction occurs when the signal is obstructed by a sharp object which derives secondary waves. Scattering occurs when the signal impinges upon rough surfaces, or small objects. Received signal is sometimes stronger than the reflected and diffracted signal since scattering spreads out the energy in all directions and consequently provides additional energy for the receiver which can receive more than one copies of the signal in multiple paths with different phases and powers. Figure 2.1 illustrates the propagation mechanisms [2].

Models play an important role in the designs of wireless communication systems. There are key modelling parameters that help to characterize a wireless channel: path loss, delay spread, coherence bandwidth, doppler spread, and coherence time.

2.1.1 Path Loss

The average received power diminishes with the distance. In free space when there is a direct path between transmitter and receiver and when

there are no secondary waves from the medium objects, the received power is inversely proportional to square of the carrier frequency and square of the distance. If P_R and P_T are the received and transmitted powers respectively, the following relation describes free space propagation [1],

$$P_R \propto \frac{G * P_T}{f^2 * d^\alpha}, \tag{2.1}$$

where f is the carrier frequency, d is the distance between transmitter and receiver, G is the power gain from the transmit and receive antennas, $\alpha = 2$ is the path loss component and *path loss* is defined as P_T/P_R. One can infer from the equation that the larger the carrier frequency the smaller the operating range.

When there is clutter in the medium, it is not easy to determine the received signal since now the direct path component is added with the multipath components at the receiver. Adopted method is to use a model that is derived from analytical and empirical methods. A good model is a model that obeys the theoretical findings and test results.

2.1.2 Shadowing

Signal power attenuates randomly with distance when the medium includes obstructions. Equation 2.1 is not enough to describe the real environment since Equation 2.1 gives the same path loss to the different receivers having the equal distance apart from the transmitter. However, if the two locations have different surroundings then the variations in the signal should be different. This behavior is called *shadowing*. Measurements indicate that the path loss is random and distributed log-normally. Modified version of Equation 2.1 for indoor path loss is called log normal shadowing and follows the formula,

$$PL(dB) = PL(d_o) + 10\alpha \log_{10}(d/d_o) + X_\sigma, \tag{2.2}$$

where d_o is the reference distance and X_σ is a zero mean Gaussian random variable with a standard deviation σ. σ and α are determined

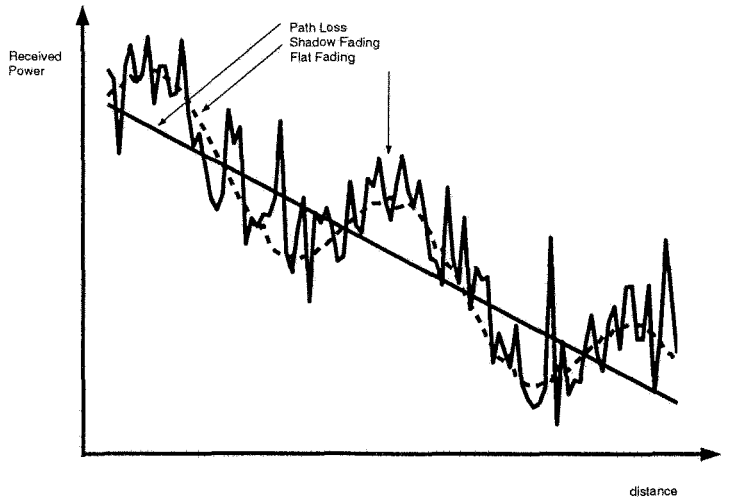


Figure 2.2: Received signal versus distance © [1]

from measured data.

A typical situation is depicted in Figure 2.2. The figure is useful to understand the difference between path loss, shadowing and flat fading, which is described in the following section.

2.1.3 Fading Parameters

One of the most intriguing aspect in wireless communication is *fading* which is present when there are multipath components. Multipath components arrive at the receiver at slightly different times. If there is movement in the system then there is also phase difference between the received components which leads to shift in the frequency. These multiple received copies apply constructive or destructive interference to the signal and create a standing wave pattern that shows rapid signal strength changes, frequency shifts or echoes.

Multipath propagation, movement and bandwidth are the factors that influence the fading. Multipath delay nature of the channel is quantified by *delay spread* and *coherence bandwidth*. The time-varying nature of the channel caused by movement is quantified by *doppler spread* and *coherence time*.

√ **Delay Spread**:

The reflected parts arrive later than the original signal to the receiver. RMS delay spread (σ_τ) is the metric to characterize this delay in terms of second order moment of the channel power profile. Typical values are on the order of microseconds in outdoor and on the order of nanoseconds in indoor radio channels [2]. Depending on the symbol duration (T_s), σ_τ plays an important role to find out how the signal is treated by the channel.

√ **Coherence Bandwidth**:

The channel power profile is coupled with channel frequency response through Fourier transform. Coherence bandwidth (B_c) is used to quantify the channel frequency response and it is inversely proportional to RMS delay spread. Coherence bandwidth is used to measure how flat the channel bandwidth is. Flatness is described as the close correlation between the two frequency components. This is important since a signal having a larger bandwidth (B_s) than B_c is severely distorted. For a 0.9 correlation $B_c \approx 1/50\sigma_\tau$ [2].

√ **Doppler Shift**:

Movement causes shift in the signal frequency. When the stations are moving, the received signal frequency is different than the original signal frequency. Doppler shift is defined as the change in the frequency.

Suppose a mobile is moving at velocity ν as in Figure 2.3. The difference in path length of received signals is $l = d\cos\theta$ where

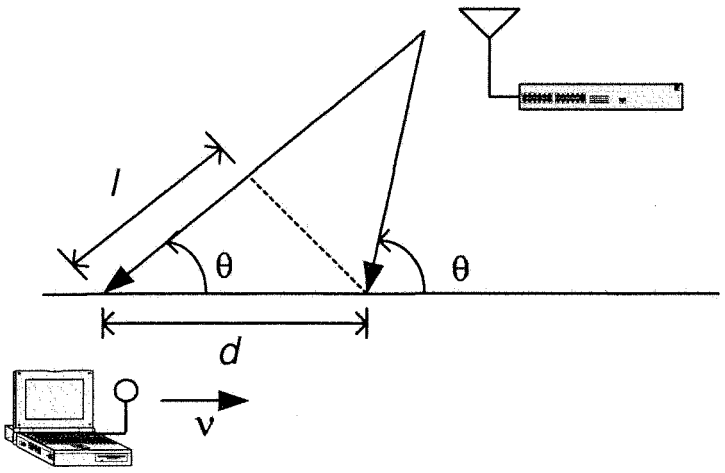


Figure 2.3: Doppler effect

$d = \nu \Delta t$. The phase change is

$$\Delta\phi = \frac{2\pi l}{\lambda} \tag{2.3}$$

and Doppler shift is given by

$$f_d = \frac{1}{2\pi} . \frac{\Delta\phi}{\Delta t} = \frac{\nu}{\lambda} . \cos\theta \tag{2.4}$$

From the Equation 2.4, one can see that if the movement is toward the signal generator, the doppler shift is positive otherwise it is negative [2].

√ **Doppler Spread**:

A channel shows a time varying nature when there is a movement in either the source or destination or even objects in the middle. Doppler spread (B_D) is the measure of maximum broadening of the spectrum due to Doppler shift. Thereby B_D is $f_m = \nu/\lambda$ where f_m is the maximum Doppler shift.

√ **Coherence Time**:

The time domain dual of Doppler spread is coherence time (T_c) where $T_c \approx 1/f_m$. Coherence time identifies the time period wherein two received signal have high amplitude correlation.

√ **Coherence Distance**:

Coherence distance gives a measure of a minimum distance between points in space for which the signal are uncorrelated. This distance is important for multiple antenna systems. It is 0.5 wavelength for wide beamwidth receive antennas and about 10 and 20 wavelengths for low-medium and high BTS antenna heights, respectively [141].

2.1.4 Flat Fading

Delay spread and coherence bandwidth are independent from Doppler spread and coherence time. The signal goes to *flat fading* if the channel bandwidth is greater than the signal bandwidth ($B_c > B_s$). Spectral shape of the signal remains but the gain changes. Narrow band signals fall into this category since their bandwidth is small as compared to the channel bandwidth. Flat fading channels bring challenges such as variation in the gain and in the frequency spectrum. Distortion in the gain may cause deep fades thereby requiring significant increase in the power in some frequencies. Destructive interferences may cause deep nulls in the signal power spectrum which is a particular problem for narrow band signals since any null in a frequency may cause lost of the signal.

There are various ways to reduce the fading distortion [1]. Diversity is one way which leverages multiple independent channels. Since the channels are independent they have lower probability to experience fades at the same time. Space diversity is one of the most efficient diversity technique compared to time or frequency. Multiple slightly

separated antennas in space are used to create independent paths to achieve uncorrelated channels.

Coding is another way which spreads the signal power among the bandwidth and lessens the power allocated in the null frequencies. As a result, loss occurs in the smaller portion of the signal power.

Another method is chopping the bandwidth into small bandwidths as in orthogonal frequency division multiplexing (OFDM). In each subcarrier small portion of the code is sent. Thus, if there is a loss in a subcarrier then it is recoverable since it corresponds to the small part of the code.

2.1.5 Frequency Selective Fading

If channel bandwidth is smaller than the signal bandwidth then the channel creates frequency selective fading. Spectral response of a radio signal will show dips due to the multipath. Channel bandwidth limits the signal bandwidth which leads to overlapping of the successive symbols in the time domain due to the convolution ($\sigma_\tau > T_s$). Overall symbol duration becomes more than the actual symbol duration. This phenomenon is called *intersymbol interference* (ISI).

There are several ways to combat ISI [1]. Equalization is one way that tries to invert the effects of the channel. In OFDM, each subcarrier undergoes flat fading since its bandwidth is less than the coherence bandwidth. Therefore ISI is eliminated within an OFDM symbol but OFDM symbols may overlap in time. A portion of the OFDM symbol, larger than the coherence bandwidth, is appended to the symbol to overcome the ISI between symbols. Equalization is simple in OFDM since each subcarrier only needs a one tap equalizer.

2.1.6 Fast Fading

Channel bandwidth may also change due to movement. Fast fading occurs if the coherence time is smaller than the symbol duration of the signal ($T_s > T_c$). The channel becomes time varying and within a symbol duration rapid changes occur in impulse response of the channel. Low data rates is vulnerable to fast fading since the symbol duration is wide in low data rates.

2.1.7 Slow Fading

If the signal bandwidth is much higher than the Doppler spread ($B_c > B_d$) then the shift in the frequency due to Doppler shift is insignificant. This is called slow fading and the channel impulse response changes slowly.

Figure 2.4 shows a table that summarizes the relationship between fading parameters and signal. As it can be seen from the figure there are two independent effects that classify the wireless channel: multipath and movement which are quantified by Delay spread and Doppler spread respectively.

2.1.8 Rayleigh Fading

Constructive and destructive nature of multipath components in flat fading channels can be approximated by Rayleigh distribution if there is no line of sight which means when there is no direct path between transmitter and receiver. The Rayleigh distribution is basically the magnitude of the sum of two equal independent orthogonal Gaussian random variables and the probability density function (pdf) is given by

$$p(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \; 0 \leq r \leq \infty \tag{2.5}$$

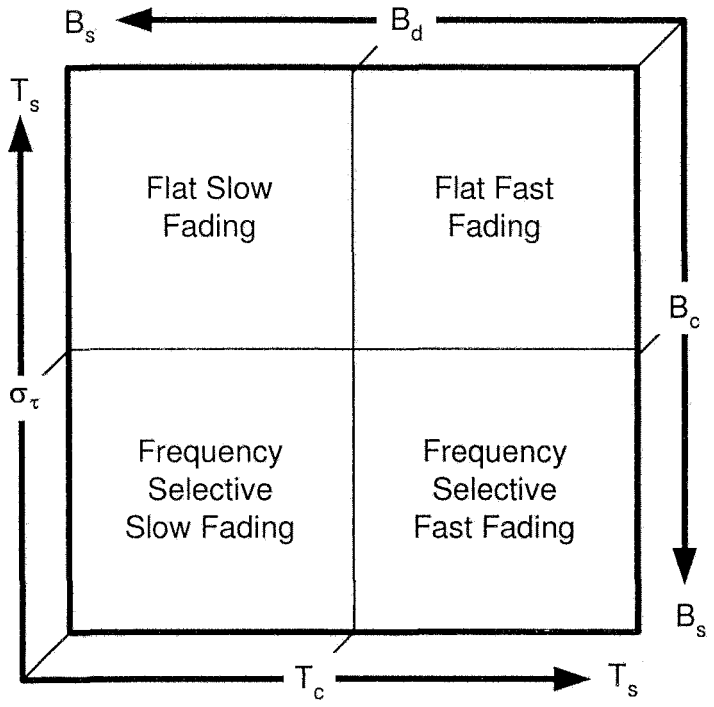


Figure 2.4: Fading illustration © [2]

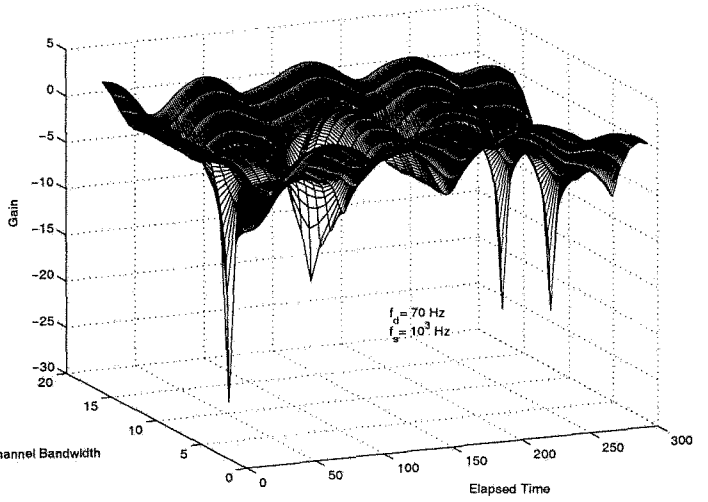


Figure 2.5: A flat fading channel where f_s is sampling frequency

where σ^2 is the time-average power of the received signal [2].

Figure 2.5 depicts a flat fading channel where the amplitude of each multipath component is created by a Rayleigh distribution.

2.1.9 Ricean Fading

When there is line of sight, direct path is normally the strongest component goes into deeper fade compared to the multipath components. This kind of signal is approximated by Ricean distribution. As the dominating component run into more fade the signal characteristic goes from Ricean to Rayleigh distribution.

The Ricean distribution is given by

$$p(r) = \frac{r}{\sigma^2} e^{-\frac{(r^2+A^2)}{2\sigma^2}} I_o(\frac{Ar}{\sigma^2}) \; for(A \geq 0, r \geq 0) \qquad (2.6)$$

where A denotes the peak amplitude of the dominant signal and $I_o(.)$ is the modified Bessel function of the first kind and zero-order [2].

2.1.10 Distortions for Wireless Systems

Wireless systems are heavily dependent on the wireless environment. Radio receivers should be able to combat the detrimental effects of the multipath channel. Typical distortions that wireless systems face are listed below.

- AWGN (receiver thermal noise)
- Receiver carrier frequency and phase offset
- Receiver timing offset
- Delay spread
- Fading
- Co-channel and adjacent channel interference
- Non-linear distortion, impulse noise

2.1.11 Diversity Techniques in a Fading Environment

Diversity techniques are important to cope with fading environment as well as coding. It tries to provide the receiver with multiple fade replicas of the same information bearing signal. If there are L independent replicas and if the p denote the outage probability then overall probability is p^L which is less than p. Diversity gain is at the highest when the diversity branches are uncorrelated. Some of the techniques are listed below and will be explained in detail in the subsequent chapters.

- √ **Frequency Diversity**: Multiple channels are separated by more than the coherence bandwidth. Using RAKE receiver, OFDM, equalization are possible techniques for frequency diversity. It can not be used over frequency-flat channel.
- √ **Time Diversity**: Channel coding, interleaving are possible techniques but it is not effective over slow fading channels.
- √ **Space Diversity**: Using multiple antennas for space diversity is efficient, coherent distance plays an important role for the selection of places of the antennas.
- √ **Angle/Direction Diversity**: Antennas with different polarizations/directions can be used for reception and transmission.

Notes

Wireless channels are studied in many books and articles, for a brief tutorial see [1] and detailed information can be found in [2].

2.2 Digital Communication System Fundamentals

The digital communication systems are composed of a transmitter, a receiver and the channel. The transmitter accepts input bits. If the initial source is analog, it is inverted to bit stream by an analog-to-digital (A/D) converter. The A/D converter samples and quantizes the continuous signal and assigns a binary number within a set. The sampling interval is determined by the *Nyquist rate* which states that in order to recover the signal, the sampling rate should be bigger than equal $2W$ where W is the bandwidth of the signal.

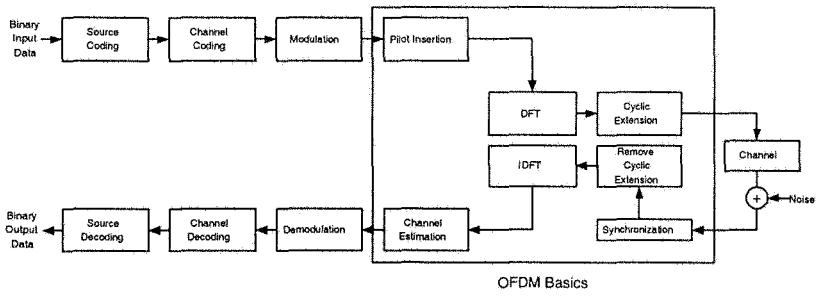


Figure 2.6: Components of a multi-carrier system

2.2.1 Coding

Source coding is used to reduce the redundancy in binary data to obtain a compressed version. An information theoretic channel is shown in Figure 2.7.

The concept of entropy defines the average information contained in a signal. The measure of entropy is the average uncertainty in the random variable. If $\log p(x_i)$ is the uncertainty $I(x_i)$ for a given discrete alphabet, then the entropy $H(X)$ is defined for an alphabet Ω as

$$H(X) = E\{I(x_i)\} = -\sum_j p(x_i) \log_2 p(x_i) \tag{2.7}$$

if M is the cardinality of the discrete alphabet set, then H is bounded by $0 \leq H(X) \leq \log_2 M$ since the maximum uncertainty is achieved if the random variable is uniformly distributed ($p(x_j) = 1/M$). The entropy of a random variable is the lower bound that can be achieved in data compression.

The expected length $L = \sum_i p(x_i) l(x_i)$ of the signal is constrained by $H \leq L$ where $l(x_i)$ is the length of code word of x_i. The basic idea of data compression is assigning short code words to the most frequent outcomes.

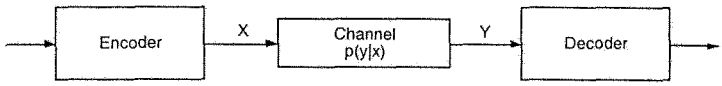


Figure 2.7: A communication channel in information theoretic view

In a communication systems, the transmitted symbols X are subject to change due to the channel impairments. As a result, the received symbols Y are not exactly the same as X. The channel capacity which is defined as the maximum achievable data rate is closely connected to the mutual information between X and Y.

Mutual information $I(X,Y)$ is a measure of the correlation between two random variables [16].

$$I(X,Y) = H(X) - H(X|Y) = \sum_{i,j} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)} \quad (2.8)$$

This measure is necessary since it is the mean of information capacity C in communication. Capacity is defined as $C = max_{p(x_i)} = I(X;Y)$ and Shannon states that all rates up to C are reliable.

Channel coding is essential to introduce code words that have data rate R less than equal to C. Unlike data compression, the channel coding introduces controllable redundancy to the transmitted data. Block, convolutional, and trellis coding are described in detail and issues about concatenated and turbo coding are cited intensely in Chapter 7.

The combination of source coding ($H \leq R = L$) and channel coding ($R \leq C$) define the reliable communication system [16].

2.2.2 Modulation

Modulation of a signal changes binary bits into an analog waveform. Modulation can be done by changing the amplitude, phase, and fre-

quency of a sinusoidal carrier. There are several digital modulation techniques used for data transmission. We will review the ones that are basic or frequently used and appropriate for OFDM. The nature of OFDM only allows the signal to modulate in amplitude and phase. There can be coherent or non-coherent modulation techniques. Unlike non-coherent modulation, coherent modulation uses a reference phase between the transmitter and the receiver which brings accurate demodulation together with receiver complexity.

Each of these classes can be represented as M-ary modulation which is the general representation of all modulation schemes. One of M possible signals can be sent in M-ary signalling schemes for a duration T. When $M = 2$ then the modulation scheme reduces to the binary modulation of its kind [18].

Coherent Modulation

Amplitude Shift Keying (ASK), Phase Shift Keying (PSK), Quadrature-Amplitude Modulation (QAM) are appropriate coherent modulation techniques for OFDM.

1. **M-ary ASK**

   Amplitude Shift Keying (ASK) carry information in amplitude. In a given set of $\{A_1, \cdots, A_M\}$, the signal s_i is

   $$s_i(t) = A_i \cos \omega t \tag{2.9}$$

   The 2-ASK also known as BPSK is a famous modulation technique which is being widely used. If E is the signal energy per symbol, then the constellation points reside at $\sqrt{2E/T}$ and $-\sqrt{2E/T}$.

2. **M-ary PSK**

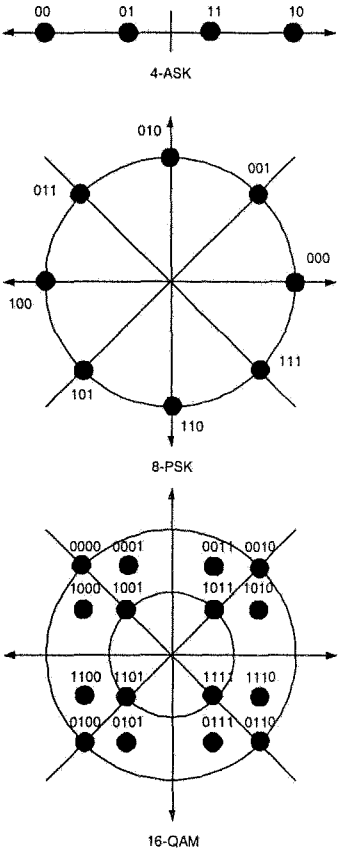


Figure 2.8: Constellation diagram

M-ary PSK shifts the phase for different constellations given a constant amplitude. There are M possible signals which has $\theta_i = 2(i-1)/M$ phases, each signal is placed on a circle and represented as

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t + \frac{2\pi}{M}(i-1)),\ i = 1, 2, \cdots, M \quad (2.10)$$

Notice that 2-PSK is again BPSK.

3. **M-ary QAM**

   M-ary QAM is a hybrid scheme which has amplitude modulation as well as a phase modulation. The M signals are represented as

$$s_i(t) = \sqrt{\frac{2E_0}{T}} a_i \cos(2\pi f_c t) + \sqrt{\frac{2E_0}{T}} b_i \cos(2\pi f_c t),\ 0 \leq t \leq T \quad (2.11)$$

   $\{a_i, b_i\}$ is an element of the L-by-L matrix:

$$\begin{bmatrix} (-L+1, L-1) & (-L+3, L-1) & \cdots & (L-1, L-1) \\ (-L+1, L-3) & (-L+3, L-3) & \cdots & (L-1, L-3) \\ \vdots & \vdots & & \vdots \\ (-L+1, -L+1) & (-L+3, -L+1) & \cdots & (L-1, -L+1) \end{bmatrix} \quad (2.12)$$

   where $L = \sqrt{M}$ [18].

   The case for $M = 4$ is same as 4-PSK which is also known as Quadrature-PSK (QPSK).

Non-coherent Modulation

IEEE 802.11 and HIPERLAN/2 uses coherent modulation, but low data rate systems such as Digital Audio Broadcasting (DAB) prefers

non-coherent modulation. Non-coherent modulation is useful when there is no knowledge of the carrier phase. It simplifies the receiver structure since there is no need for channel estimation, carrier or phase tracking.

Differential techniques are very common in non-coherent modulation. The data is encoded in phase or amplitude which is the difference between phases or amplitudes of two consecutive symbols.

Differential-PSK which is used in DAB carry the data in phase difference of two M-PSK modulated signals,

$$s_i(t) = A\cos(\omega t + \Delta\theta) \tag{2.13}$$

where $\Delta_i\theta = \theta_i - \theta_{i-1}$.

Differential detection uses two consecutive symbols and calculates $\theta_{i-1} + \Delta\theta = \theta_i$ which gives a signal that needs to be PSK demodulated.

2.3 Multi-Carrier System Fundamentals

Let $[D_0, D_1, ..., D_{N-1}]$ denote data symbols which is where $D_t = s_i(t)$ for $i \in \{1, \cdots, M\}$. Digital signal processing techniques, rather than frequency synthesizers, can be deployed to generate orthogonal sub-carriers. The DFT as a linear transformation maps the complex data symbols $[D_0, D_1, ..., D_{N-1}]$ to OFDM symbols $[d_0, d_1, ..., d_{N-1}]$ such that

$$d_k = \sum_{n=0}^{N-1} D_n e^{j2\pi n\frac{k}{N}} \tag{2.14}$$

The linear mapping can be represented in matrix form as:

$$\bar{d} = \bar{W}\bar{D}, \tag{2.15}$$

where:

$$\bar{W} = \begin{bmatrix} 1 & \cdots & 1 & 1 \\ 1 & W & \cdots & W^{N-1} \\ 1 & W^2 & \cdots & W^{2(N-1)} \\ \vdots & & & \vdots \\ 1 & W^{N-1} & & W^{N(N-1)} \end{bmatrix} \tag{2.16}$$

and,

$$W = e^{j\frac{2\pi}{N}} \tag{2.17}$$

$\bar{W}$ is a symmetric and orthogonal matrix. After FFT, a cyclic pre/postfix of lengths k_1 and k_2 will be added to each block (OFDM symbol) followed by a pulse shaping block. Proper pulse shaping has an important effect in improving the performance of OFDM systems in the presence of some channel impairments, and will be discussed in Chapter 5. The output of this block is fed into a D/A at the rate of f_s and low-pass filtered. A basic representation of the equivalent complex baseband transmitted signal is

$$x(t) = \sum_{n=0}^{N-1} \{D_n e^{j2\pi \frac{n}{N} f_s t}\} \tag{2.18}$$

for

$$-\frac{k_1}{f_s} < t < \frac{N + k_2}{f_s} \tag{2.19}$$

A more accurate representation of OFDM signal including windowing effect is

$$x(t) = \sum_{l=-\infty}^{\infty} \sum_{k=-k_1}^{N+k_2} \sum_{n=0}^{N-1} \{D_{nl} e^{j2\pi \frac{n}{N} k}\} w(t - \frac{k}{f_s} - lT) \tag{2.20}$$

D_{nl} represents the nth data symbol transmitted during the lth OFDM block, $T = (N + k_1 + k_2)/f_s$ is the OFDM block duration, and $w(t)$ is the window or pulse shaping function. The extension of the OFDM

block is equivalent to adding a cyclic pre/postfix in the discrete domain. The received signal for a time-varying random channel is

$$r(t) = \int_0^\infty x(t-\tau)h(t,\tau)d\tau + n(t). \tag{2.21}$$

The received signal is sampled at $t = \frac{k}{f_s}$ for $k = \{-k_1, ..., N+k_2-1\}$. With no inter-block interference, and assuming that the windowing function satisfies $w(n-1) = \delta_{nl}$, the output of the FFT block at the receiver is

$$\tilde{D}_m = \frac{1}{N}\sum_{k=0}^{N-1} r_k e^{-j2\pi m\frac{k}{2N}}, \tag{2.22}$$

where

$$r_k = \sum_{n=0}^{N-1} H_n D_n e^{j2\pi\frac{n}{2N}k} + n(k). \tag{2.23}$$

A complex number H_n is the frequency response of the time-invariant channel $h(t-\tau)$ at frequency n/T. So,

$$\tilde{D}_m = \begin{cases} H_n D_n + N(n), & n = m \\ N(n), & n \neq m \end{cases} \tag{2.24}$$

$n(t)$ is white Gaussian noise with a diagonal covariance matrix of $E(n(k)n(l)) = \delta I$. Therefore, the noise components for different subcarriers are not correlated,

$$E(\bar{n}(k)\bar{n}^*(l)) = W\delta I W^T = \delta I. \tag{2.25}$$

where $\bar{n}(k)$ is the vector of noise samples $\{ n(k), \cdots, n(k-N)\}$.

A detailed mathematical analysis of OFDM in multi-path Rayleigh fading is presented in the last section.

2.4 DFT

The key components of an OFDM system are the Inverse DFT in the transmitter and the DFT in the receiver. These components must im-

plement

$$d_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} D_k W_N^{kn}, \; n = 0, \cdots, N-1 \tag{2.26}$$

and,

$$D_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} d_k W_N^{-kn}, \; n = 0, \cdots, N-1 \tag{2.27}$$

respectively.

These operations perform reversible linear mappings between N complex data symbols $\{D\}$ and N complex line symbols $\{d\}$. It can be seen that the two operations are essentially identical.

The scale factor $1/\sqrt{N}$ provides symmetry between the operations and also preservation of power. Frequently, a scale factor of $1/N$ is used in one direction and unity in the other instead. In an actual implementation, this is immaterial because scaling is chosen to satisfy considerations of overflow and underflow rather than any mathematical definition.

A general N-to-N point linear transformation requires N^2 multiplications and additions. This would be true of the DFT and IDFT if each output symbol were calculated separately. However, by calculating the outputs simultaneously and taking advantage of the cyclic properties of the multipliers $e^{\pm j2\pi nk/N}$, Fast Fourier Transform (FFT) techniques reduce the number of computations to the order of $N \log N$. The FFT is most efficient when N is a power of two. Several variations of the FFT exist, with different ordering of the inputs and outputs, and different use of temporary memory. One variation, decimation in time, is shown in Figure 2.9.

Figure 2.10 shows the architecture of an OFDM system capable of using a further stage of modulation, employing both in-phase and quadrature modulators. This configuration is common in wireless communication systems for modulating baseband signals to the required IF or RF frequency band. It should be noted that the basic con-

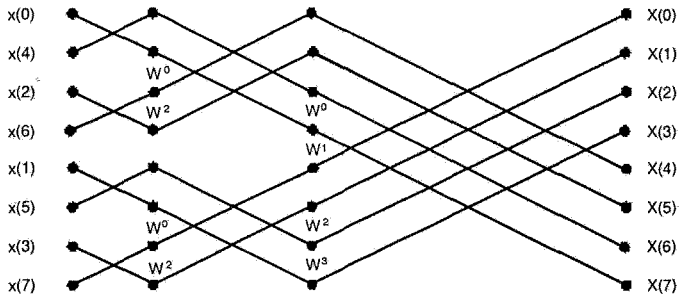


Figure 2.9: An FFT implementation (decimation in time)

figuration illustrated does not account for channel dispersion, which is almost always present. The channel dispersion problem is solved by using the cyclic prefix which will be described later.

Small sets of input bits are first assembled and mapped into complex numbers which determine the constellation points of each sub-carrier. In most wireless systems, smaller constellations are formed for each sub-carrier. In wireline systems, where the signal-to-noise ratio is higher and variable across the frequency range, the number of bits assigned to each sub-carrier may be variable. Optimization of this bit assignment is the subject of bit allocation, to be discussed in Chapter 3 and 11.

If the number of sub-carriers is not a power of two, then it is common to add symbols of zero value to the input block so that the advantage of using such a block length in the FFT is achieved. The quantity N which determines the output symbol rate is then that of the padded input block rather than the number of sub-carriers.

The analog filters at the transmitter output and the receiver input should bandlimit the respective signals to a bandwidth of $1/T$. Low pass filters could be used instead of the band-pass filters shown, placed on the other side of the modulator or demodulator. The transmit filter

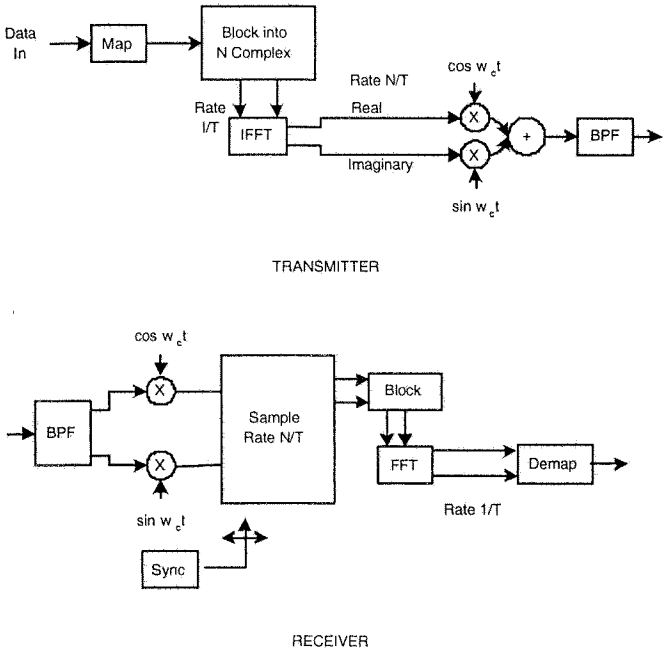


Figure 2.10: System with complex transmission

eliminates out-of-band power which may interfere with other signals. The receive filter is essential to avoid aliasing effects. The transmitted signal has a continuous spectrum whose samples at frequencies spaced $1/T$ apart agree with the mapped input data. In particular, the spectrum of each sub-carrier is the form $sinc(1/T)$, whose central value is that input value, and whose nulls occur at the central frequencies of all other sub-carriers.

The receiver operations are essentially the reverse of those in the transmitter. A critical set of functions, however, are synchronization of carrier frequency, sampling rate, and block framing. There is a minimum delay of $2T$ through the system because of the block assembly functions in the transmitter and receiver.

The FFT functions may be performed either by a general purpose DSP or special circuitry, depending primarily on the information rate to be carried. Some simplification compared with full multiplication may be possible at the transmitter, by taking advantage of small constellation sizes.

The number of operations per block of duration T is $KN \log_2 N$, where K is a small quantity. To compare this with a single carrier system, the number of operations per line symbol interval T/N is $K \log_2 N$, which is substantially below the requirement of an equalizer in a typical single carrier implementation for wireline applications. In most wireline systems it is desirable to transmit the transformed symbols d_n without any further modulation stages. In this case, it is only possible to transmit real line symbols, and not the above complex quantities. The problem is solved by augmenting the original sequence D_n by appending its complex conjugate to it, as shown in Figure 2.11. The $2N$-point IFFT of this augmented sequence is then a sequence of $2N$ real numbers, which is equivalent in bandwidth to N complex numbers.

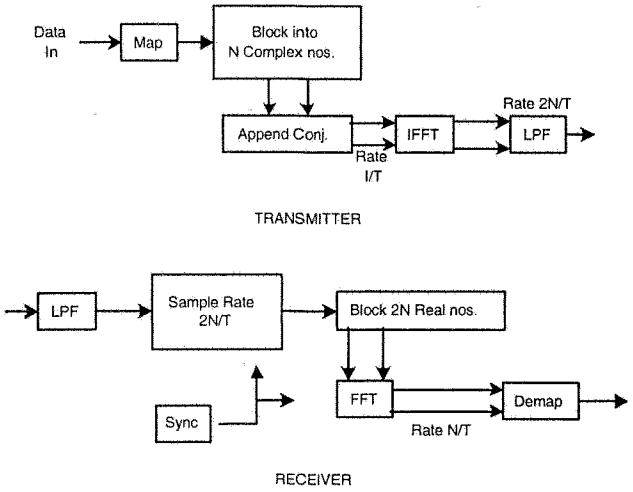


Figure 2.11: System with real transmission

The augmented sequence is formed from the original sequence as

$$D_n' = \begin{cases} D_n, & n = 1, \cdots, N-1 \\ D_{2N-n}^*, & n = N+1, \cdots, 2N-1 \end{cases} \tag{2.28}$$

In order to maintain conjugate symmetry, it is essential that D_0' and D_N' be real. If the original D_0 is zero, as is common, then D_0' and D_N' are set to zero. Otherwise, D_0' may be set to $Re(D_0)$ and D_N' to $Im(D_0)$.

For the simple case of $D_0 = 0$, the output of the IFFT is:

$$\begin{aligned} d_m &= \sum_{n=0}^{2N-1} D_n' e^{\frac{j\pi mn}{N}} \\ &= 2Re \sum_{n=0}^{N-1} D_n e^{\frac{j\pi mn}{N}} \\ &= 2 \sum_{n=0}^{N-1} [A_n cos\frac{\pi mn}{N} - B_n sin\frac{\pi mn}{N}], \quad m = 0, \cdots, 2N-1, \end{aligned} \tag{2.29}$$

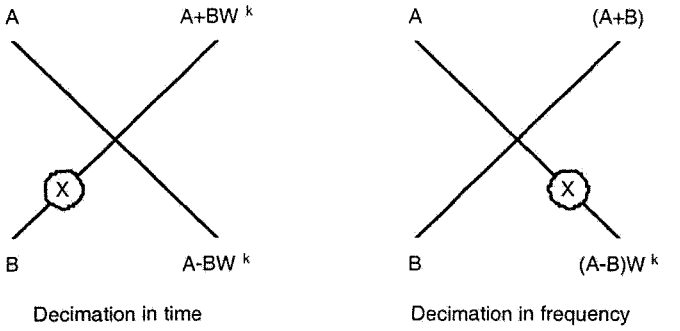


Figure 2.12: Two different techniques for FFT butterfly

where $D_n = A_n + jB_n$. Of course the scaling by a factor of two is immaterial and can be dropped. This real orthogonal transformation is fully equivalent to the complex one, and all subsequent analyzes are applicable.

2.5 Partial FFT

In some applications, the receiver makes use of a subset of transmitted carriers. For example, in a digital broadcasting system receiver [12] that decodes a group of sub-carriers (channels) or in some multirate applications, the receiver has several fall-back modes so using a repetitive structure is advantageous. One of the benefits of an OFDM system with an FFT structure is the fact that it lends itself to a repetitive structure very well. This structure is preferred compared to the required filtering complexity in other wideband systems. Two common structures are shown in Figure 2.12.

An example of partial FFT is shown in Figure 2.13. In order to detect (receive) the marked point at the output, we can restrict the

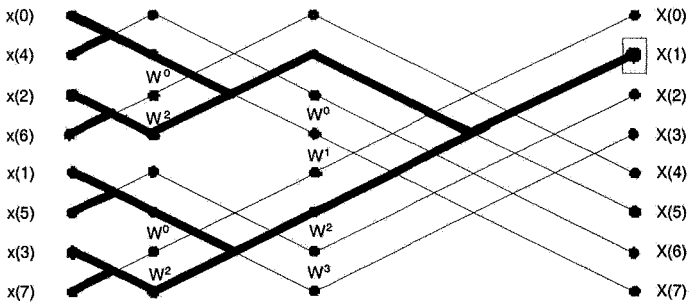


Figure 2.13: Partial FFT (DIT)

FFT calculation to the marked lines. Therefore, a significant amount of processing will be saved.

Two main differences between decimation in time (DIT) and decimation in frequency (DIF) are noted [11]. First, for DIT, the input is bit-reversed and output is in natural order, while in DIF , the reverse is true. Secondly, for DIT complex multiplication is performed before the add-subtract operation, while in DIF the order is reversed. While complexity of the two structures is similar in typical DFT, this is not the case for partial FFT [12]. The reason is that in the DIT version of partial FFT, a sign change (multiplication by 1 and -1) occurs at the first stages, but in the DIF version it occurs in later stages.

2.6 Cyclic Extension

Transmission of data in the frequency domain using an FFT, as a computationally efficient orthogonal linear transformation, results in robustness against ISI in the time domain. Unlike the Fourier Transform (FT), the DFT (or FFT) of the circular convolution of two signals is

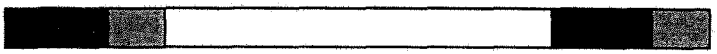


Figure 2.14: Prefix and postfix cyclic extension

equal to the product of their DFT's (FFT).

$$\begin{aligned} FT\{d_n * h_n\} &= FT\{d_n\} \times FT\{h_n\} \\ DFT\{d_n \otimes h_n\} &= DFT\{d_n\} \times DFT\{h_n\}, \end{aligned} \tag{2.30}$$

where $*$ and $\otimes$ denote linear and circular convolution respectively.

The signal and channel, however, are linearly convolved. After adding prefix and postfix extensions to each block, the linear convolution is equivalent to a circular convolution as shown in Figure 2.14. Instead of adding a prefix and a postfix, some systems use only prefix, then by adjusting the window position at the receiver proper cyclic effect will be achieved.

Using this technique, a signal, otherwise aliased, appears infinitely periodic to the channel. Let's assume the channel response is spread over M samples, and the data block has N samples then:

$$y(n) = \sum_{m=0}^{N-1} d(m)h(n-m) \times R_N(n), \; n = 0, 1, \cdots, N+M-1 \tag{2.31}$$

where $R_N(n)$ is a rectangular window of length N. To describe the effect of distortion, we proceed with the Fourier Transform noting that convolution is linear

$$\begin{aligned} Y(k) &= Y(e^{jw}) \Big|_{\omega = k\omega_s} \\ &= \{D(e^{jw}) \times H(e^{jw})\} * \frac{sin\omega\frac{N}{2}}{sin\frac{\omega}{2}} e^{j\omega\frac{N-1}{2}} \Big|_{\omega = k\omega_s} \end{aligned} \tag{2.32}$$

After linear convolution of the signal and channel impulse response, the received sequence is of length $N + M - 1$. The sequence is truncated to N samples and transformed to the frequency domain, which is equivalent to convolution and truncation. However, in the case of cyclic pre/postfix extension, the linear convolution is the same as the circular convolution as long as channel spread is shorter than guard interval. After truncation, the DFT can be applied, resulting in a sequence of length N because the circular convolution of the two sequences has period of N.

Intuitively, an N-point DFT of a sequence corresponds to a Fourier series of the periodic extension of the sequence with a period of N. So, in the case of no cyclic extension we have

$$\sum_{i=-\infty}^{\infty} \sum_{m=0}^{N-1} d(m)h(n + iN - m), \tag{2.33}$$

which is equivalent to repeating a block of length $N + M - 1$ with period N. This results in aliasing or inter-symbol interference between adjacent OFDM symbols. In other words, the samples close to the boundaries of each symbol experience considerable distortion, and with longer delay spread, more samples will be affected. Using cyclic extension, the convolution changes to a circular operation. Circular convolution of two signals of length N is a sequence of length N so the inter-block interference issue is resolved.

Proper windowing of OFDM blocks, as shown later, is important to mitigating the effect of frequency offset and to control transmitted signal spectrum. However, windowing should be implemented after cyclic extension of the frame, so that the windowed frame is not cyclically extended. A solution to this problem is to extend each frame to $2N$ points at the receiver and implement a $2N$ FFT. Practically, it requires a $2N$ IFFT block at the transmitter, and $2N$ FFT at the receiver. However, by using partial FFT techniques, we can reduce the computation by calculating only the required frequency bins.

If windowing was not required, we could have simply used zero

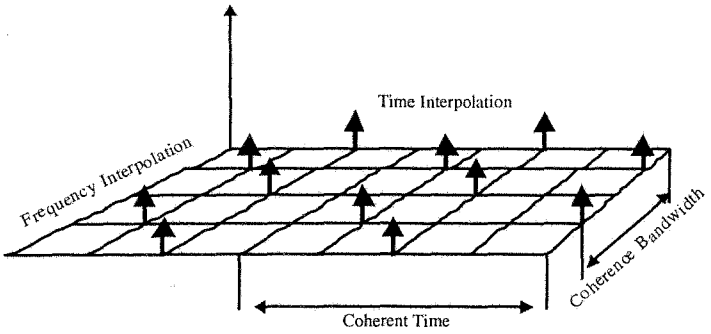


Figure 2.15: Pilot positioning in time and frequency

padded pre/postfix. Before the DFT at the receiver, we could copy the beginning and the end of frame as prefix and postfix. This creates the same effect of cyclic extension with the advantage of reducing transmit power and causing less ISI.

The relative length of cyclic extension depends on the ratio of the channel delay spread to the OFDM symbol duration.

2.7 Channel Estimation

Channel estimation inverts the effect of non-selective fading on each sub-carrier. The channel estimation in OFDM is usually coherent detection which is based on the use of pilot subcarriers in given positions of the frequency-time grid. Chapter 6 covers the channel estimation in more detail. Time varying channels are required to implement repeated pilot symbols. Nature of the channel such as coherence time and coherence bandwidth determines the pilot spacing in time and frequency. As it can be seen from Figure 2.15, the number of pilots necessary to use can be reduced by spacing the pilots less than coherence bandwidth in frequency and less than in coherence time in time

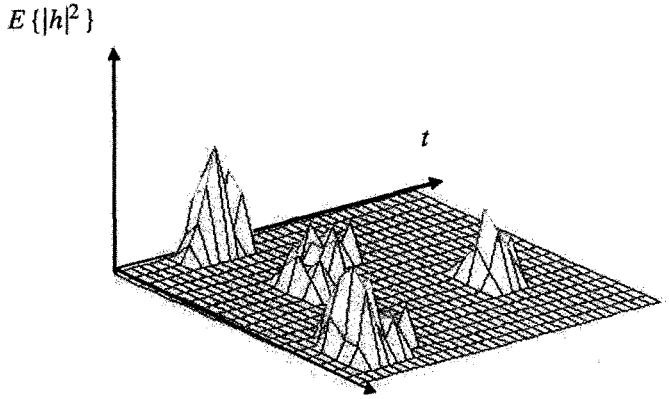


Figure 2.16: Typical impulse response of a wireless channel

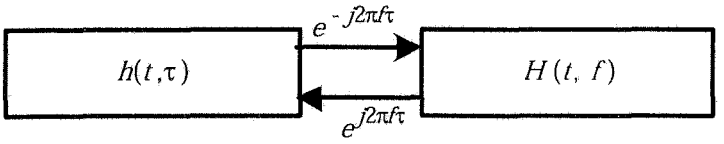


Figure 2.17: Relationship between system functions

domain.

Let

$$h_t = (h_1, h_2, \cdots, h_M)^T \tag{2.34}$$

represent the overall impulse response of the channel, including transmitter and receiver filters, with length M [15]. The N-element training sequence is represented by

$$X = [X_{ij}], 1 \leq i \leq M+N-1, \qquad 1 \leq j \leq M$$
$$X_{ij} = \begin{cases} x_{i-j+1}, & 1 \leq i-j+1 \leq N \\ 0, & elsewhere \end{cases} \tag{2.35}$$

The noise samples are

$$n = (n_1, n_2, \cdots, n_{M+N-1})^T, \tag{2.36}$$

with

$$R_n = E\{nn^*\} = LL^* \tag{2.37}$$

in which L is a lower triangular matrix, which can be calculated by the Cholesky method. The received signal is then

$$r = Xh_t + n \tag{2.38}$$

After correlation or matched filtering, the estimate of the impulse response is:

$$\hat{h}_t = X^* r = X^* X h_t + X^* n. \tag{2.39}$$

If the additive noise is white, the matched filter is the best estimator in terms of maximizing signal-to-noise ratio. However, the receiver filter colors the noise. With a whitening filter, the estimate is given in [111]. An unbiased estimator, which removes the effect of sidelobes, is

$$\hat{h}_t = (X^* R_n^{-1} X)^{-1} X^* R_n^{-1} r = h + (X^* R_n^{-1} X)^{-1} X^* R_n^{-1} n. \tag{2.40}$$

In the dual system architecture, circular convolution is replaced by multiplication. The repetition of the training sequence results in circular convolution. Therefore, Equation 2.39 is replaced by:

$$\hat{H}^\omega = X^\omega X^{\omega^*} H^\omega + X^{\omega^*} N^\omega, \tag{2.41}$$

where superscript ω shows Fourier Transform and all products are scalar. The same procedure can be applied to Equation 2.40. This procedure is also called frequency equalization.

2.8 Modelling of OFDM for Time-Varying Random Channel

2.8.1 Randomly Time-Varying Channels

Linear randomly time-varying channels can be characterized by their impulse response , which represents system response at time t due to an impulse at time $t - \tau$. So, for input $x(t)$, the output would be $y(t) = \int h(t,\tau)x(t-\tau)d\tau$.

Other kernels can be defined to relate input and output in time and frequency domains [13]. Since the impulse response is time-varying, it is expected that the system transfer function is time-varying. The transfer function is defined as

$$H(t,f) = \int e^{-j2\pi ft}h(t,\tau)d\tau, \tag{2.42}$$

$$y(t) = \int X(f)H(t,f)e^{j2\pi ft}df. \tag{2.43}$$

If the input is a single tone $e^{j2\pi mt}$ the output is:

$$\int e^{j2\pi m(t-\tau)}h(t,\tau)d\tau = e^{j2\pi mt}H(t,m), \tag{2.44}$$

which shows that the time-varying transfer function has an interpretation similar to time-invariant systems. Namely, the output of the system for a sinusoidal function is the input multiplied by the transfer function. In general, $h(t,\tau)$ is a random process and its exact statistical description requires a multi-dimensional probability distribution. A practical characterization of the channel is given by first and

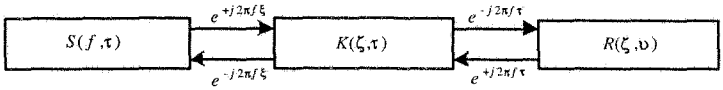


Figure 2.18: Relationship between correlation functions

second order moments of system functions. For example, many randomly time-varying physical channels are modelled as wide sense stationary non-correlated scatterers. Therefore, they are white and non-stationary with respect to delay variable τ and stationary with respect to t.

$$E\{h(t,T)h^*(t',\tau')\} = K(t-u,\tau)\delta(\tau-\tau'), \tag{2.45}$$

$K(t,\tau)$ is called the tap gain correlation. The scattering function of the channel is defined as

$$S(f,\tau) = \int K(\psi,\tau)e^{-j2\pi f\psi}d\psi. \tag{2.46}$$

For any fixed τ, the scattering function may be regarded as the complex-valued spectrum of the system. The relationship between different system functions is shown in the Figure 2.17 and Figure 2.18.

Coherence bandwidth of the received signal (or channel) is defined as the frequency distance beyond which frequency responses of the signal are uncorrelated. The reciprocal of delay spread is an estimate for coherence bandwidth. Coherence time is the time distance beyond which the samples of received signal (or channel impulse response) are uncorrelated. Since the channel and signal are Gaussian, independence of samples is equivalent to their being non-correlated. The reciprocal of Doppler spread is an estimate of coherence time [14].

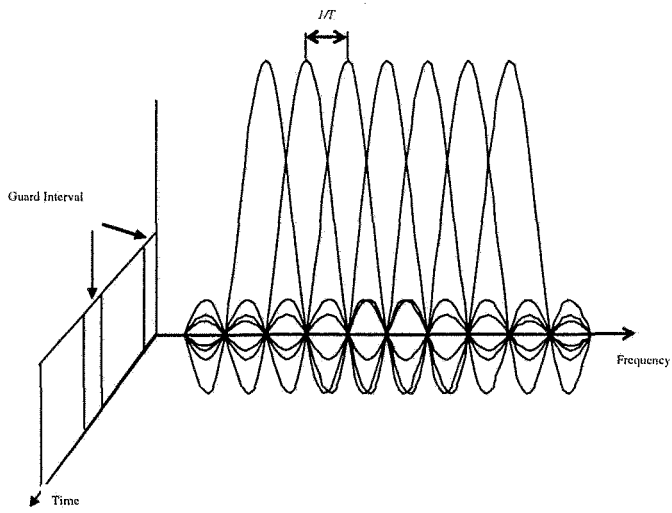


Figure 2.19: OFDM time and frequency span

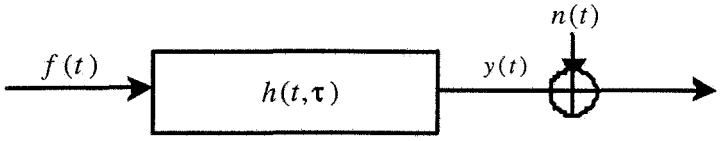


Figure 2.20: Time-varying channel

2.8.2 OFDM in Randomly Time-Varying Channels

In general, OFDM is a parallel transmission technique in which N complex symbols modulate N orthogonal waveforms $\psi_i(t)$ which maintain their orthogonality at the output of the channel. This requires that the correlation function of the channel output process, in response to different waveforms, has orthogonal eigenfunctions.

Let $h(t,\tau)$ be the Rayleigh distributed, time-varying complex impulse response of the channel, while $R(t,\tau)$ denotes the complex correlation of the output in response to a basic pulse shape $f(t)$:

$$R(t,\tau) = E\{y(t)y(t+\tau)^*\}. \tag{2.47}$$

Using Karhunen-Loeve expansion, the output of the channel $y(t)$ can be represented as

$$y(t) = \sum_i D_i y_i \varphi_i(t), \tag{2.48}$$

where the $\varphi_i(t)$s are eigenfunctions of the correlation function $R(t,\tau)$ and

$$y_i = \int y(t)\varphi_i(t)dt, \tag{2.49}$$

Using Karhunen-Loeve expansion, the output of the channel $y(t)$ can be represented as

$$y(t) = \sum_i y_i \varphi_i(t), \tag{2.50}$$

where the $\varphi_i(t)$s are eigenfunctions of the correlation function $R(t,\tau)$ and

$$y_i = \int y(t)\varphi_i(t)dt \tag{2.51}$$

The corresponding eigenvalues of the correlation function are λ_i where

$$\int R(t,\tau)\varphi_i(\tau)d\tau = \lambda_i \varphi_i(t). \tag{2.52}$$

For a Gaussian random process $y(t)$, the coefficients in Equation 2.50 are Gaussian random variables and their powers are the corresponding

eigenvalues. If the input consists of orthogonal modulating functions which are harmonics of the form

$$\psi_k(t) = u(t)e^{j2\pi f_k t}, \tag{2.53}$$

where $u(t)$ is the basic waveform such as time-limited gating function. The eigenfunctions of the received signal correlation function $R_k(t,\tau)$ are,

$$\varphi_i(t)e^{j2\pi f_i t}, \tag{2.54}$$

where again $\varphi_i(t)$ are the eigenfunctions' output correlation functions in response to basic waveform $u(t)$. Therefore, the orthogonality condition is

$$\int \varphi_i(t)\varphi_k^*(t)e^{j2\pi(f_i-f_k)}dt = 0 \tag{2.55}$$

which can be achieved by increasing $\delta f = f_{i+1} - f_i$. Analytically this means that l_2 sub-spaces spanned by different eigenfunctions are mutually orthogonal. In other words, the frequency response of the channel for different frequency bins is independent. For a pure tone signal of frequency f_k, the output of random channel is:

$$\begin{aligned} y_k(t) &= \int h(t,\tau)x(t-\tau)d\tau = \int h(t,\tau)e^{j2\pi f_k(t-\tau)}d\tau \\ &= e^{j2\pi f_k t}\int h(t,\tau)e^{-j2\pi f_k \tau}d\tau \end{aligned} \tag{2.56}$$

The transfer function of a time-varying channel is defined as

$$H(f,t) = \int e^{-j2\pi f\tau}h(t,\tau)d\tau, \tag{2.57}$$

so the output of channel can be represented as

$$y_k(t) = H(f_k,t)e^{j2\pi f_k t}, \tag{2.58}$$

and the orthogonality requirement of the output processes holds. If the filter gains $H(f_k,t)$ are orthogonal, and the tap filters are non-correlated and Gaussian, then they will be independent too. As an example, consider orthogonal functions of the form

$$\varphi_{ik}(t) = u(t-iT)e^{2j\pi f_k(t-jT)}. \tag{2.59}$$

This set of orthogonal functions is used to modulate data symbols and transmit them in a single OFDM symbol timing interval such that the orthogonality in the frequency domain is preserved. By choosing the above set of orthogonal functions, the eigenfunctions used in Karhunen-Loeve expansion will be related as shown in Equation 2.54. The optimum receiver maximizes a *posteriori* probability (conditional probability of the received signal given a particular sequence is transmitted). Usually, channel characteristics do not change during one or a few symbols. Therefore, the channel output can be represented as

$$y(t) = \sum_k H(f_k, t) e^{j2\pi f_k t}. \tag{2.60}$$

The received signal $r(t) = y(t) + n(t)$ with $n(t)$ representing additive white Gaussian noise can be expanded as

$$r(t) = \sum_i r_i \varphi_i(t), \tag{2.61}$$

where

$$\begin{aligned} r_i &= y_i + n_i \\ n_i &= \int n(t)\varphi_i(t)dt \\ y_i &= \int y(t)\varphi_i(t)dt \end{aligned} \tag{2.62}$$

Notice that the noise projection should be interpreted as a stochastic integral, because a Riemann integral is not defined for white noise. r_is are independent zero mean Gaussian random variables and their powers are

$$E|r_i|^2 = E|H_i|^2 + \bar{n}_i^2, \tag{2.63}$$

assuming that data symbols are of equal and unit power. So, the problem at hand reduces to the detection of a Gaussian random process corrupted by Gaussian noise.

Chapter 3

Performance over Time-Invariant Channels

3.1 Time-Invariant Non-Flat Channel with Colored Noise

Many channels, such as wireline, have transmission properties and noise statistics that do not vary rapidly with time. Therefore, over moderately long time intervals, they may be treated as being invariant. In such a case, the channel impulse response $h(t, \tau)$ is a function only of $t - \tau$, and may be written as $h(u)$ where $u = t - \tau$. Since it is a function of only one variable, we may take its 1-dimensional Fourier Transform to arrive at the channel transfer function

$$H(f) = \int_{-\infty}^{\infty} h(u) e^{-j2\pi f u} du. \tag{3.1}$$

Note that we did not constrain $h(u)$ to be causal. A common practice is to shift the time axis so that $t = 0$ coincides with some principal sampling point of $h(u)$.

Similarly the auto-correlation of the noise $R_N(t, \tau)$ is also a func-

tion only of $t - \tau$, and can be treated as $R_n(u)$. Its Fourier Transform $N(f)$ is the power spectral density. Since, in general, both $|H(f)|$ and $N(f)$ vary over the frequency band of interest, the signal-to-noise ratio also varies.

We are now ready to analyze the performance of a multi-carrier system over such a channel, and to optimize the transmitted signal so as to maximize performance.

3.2 Error Probability

Since a fully equalized OFDM system can be treated as N independent QAM signals with independent noise, error probability analysis is the same as for QAM [3]. The QAM signal with independent noise in turn consists of two orthogonal components, so we may deal with each of these components separately.

Each of the I and Q components can be considered to be an independent PAM signal. As noted previously, the signal points in a symmetric L-level PAM signal after demodulation can be written as

$$\{-A(L-1), \cdots, -A, A, \cdots, A(L-1)\}. \tag{3.2}$$

The average power of either the I or Q component of the jth subchannel is then

$$P = 2A^2 \sum_k p(k)(2k-1)^2, \tag{3.3}$$

where we have used the symmetry of the signal, and $p(k)$ is the probability of level k.

If the levels are equiprobable, as in an L^2-point QAM constellation, then

$$P = \frac{L^2 - 1}{3} A^2 \tag{3.4}$$

If the noise can be treated as white with power spectral density N_{0j} over the sub-channel of interest, then its variance will be

$$N_j = N_{0j}/2T \tag{3.5}$$

in the I or Q component.

An error will result if the noise amplitude D, after demodulation, is such that $D > A$ or $D < -A$ for an inner level, $D < -A$ for the most positive level, and $D > A$ for the most negative level.

So, the error probability is

$$P_{ej} = 2\frac{L-1}{L}Q\left(\frac{A}{\sqrt{N_j}}\right) \tag{3.6}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-x/2}. \tag{3.7}$$

It should be noted that for low error probabilities ($x > 3$), $Q(x)$ may be approximated by

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}}e^{-x/2}. \tag{3.8}$$

Equation 3.8 is the probability of error per I or Q sub-symbol. We are usually concerned with the bit error probability. If the assignment of bits to levels is such that adjacent levels differ by only one bit, as in a Gray code, then the bit error probability becomes

$$p_{ebj} = \frac{p_{ej}}{\log_2 L}. \tag{3.9}$$

Most radio systems use QPSK (a 4-point constellation) on each sub-carrier, so that L = 2 and

$$p_{ebj} = Q\left(\frac{A}{\sqrt{N_j}}\right). \tag{3.10}$$

For the more general equiprobable case, we substitute

$$p_{ebj} = 2\frac{L-1}{L\log_2 L}Q\left(\sqrt{\frac{3P_j}{N_j(L^2-1)}}\right) \tag{3.11}$$

We can often neglect the coefficient, which is always ≤ 1. Particularly for low error probability and a constellation size not too large, this provides a reasonably tight upper bound.

The quantity P_j/N_j is the signal-to-noise power ratio, where that noise power is measured in a bandwidth equal to the sub-carrier QAM symbol rate. It is also equal to E/N_{0j}, where E is the energy per I or Q sub-symbol. A common representation of the signal-to-noise ratio is E_b/N_0 where E_b is the energy per bit, and

$$E_b = \frac{E}{\log_2 L}. \tag{3.12}$$

The above analysis assumes that the noise power spectral density is constant at $N(f) = N_{0j}/2$ for each sub-channel. For multi-carrier systems with a large number of sub-channels, this should be an excellent approximation. When this is not the case, we should use

$$N_j = \int N(f)R_j^2(f)df \tag{3.13}$$

where $R_j(f)$ is the receiver response for the jth sub-channel. For straight DMT without windowing,

$$R_j(f) = \frac{sin\pi(|f| - f_j)T}{\pi(|f| - f_i)T} \tag{3.14}$$

where f_j is the center frequency of the sub-channel.

It is common in analyzing multi-carrier systems to deal with QAM sub-carriers rather than the I and Q components. Equation 3.11 can be written as

$$P_{ebj} = \frac{K}{b_j}Q\left(\sqrt{\frac{3P_j}{N_j(M_j-1)}}\right), \tag{3.15}$$

where M_j is the number of points in the QAM constellation, P_j is its power, N_j is the noise power in the that sub-channel after receiver processing, and K is the average number of its nearest neighbors, and is usually (but not necessarily) an integer.

Equation 3.15 is strictly applicable to square equiprobable constellations. However, very little error will result if we instead use the same analysis for more general equiprobable constellations. For example, a large QAM constellation with a circular boundary will have only 0.2 dB less average power than a square constellation with the same number of points and spacing between points. If further "shaping gain" is provided by using a multi-dimensional hyperspheric constellation boundary over several QAM sub-channels, then up to 1.53 dB average power reduction can be achieved. However, such shaping gain is rarely used in OFDM systems

$$p_{eb} = \frac{1}{\sum_j b_j} \sum_j b_j p_{ebj}. \tag{3.16}$$

For a complete OFDM system, the overall average bit error probability is

$$P_{ej} = 2\frac{L-1}{L} Q\left(\frac{A}{\sqrt{N_j}}\right). \tag{3.17}$$

Clearly if the p_{ebj} are all equal, that will also be the overall error probability. Otherwise, the higher error probabilities will dominate performance. A sensible design goal is, therefore, to try to achieve the same error probabilities on all sub-channels if possible.

3.3 Bit Allocation

We now address the problem of optimizing the performance of an OFDM system over a stationary, non-flat, linear channel by choosing

the transmitted signal. We may either seek to maximize the overall bit rate with a required error probability, or we may minimize the error probability for a given bit rate. We will first deal with the former optimization, that is, the maximization of

$$R_b = \frac{1}{T} \sum_j b_j, \tag{3.18}$$

where b_j is the number of bits carried by the constellation of the jth sub-carrier. We will deal primarily with the case of high signal-to-noise ratio (low error probability).

The total transmitted power is assumed to be constrained. We assume that the multi-carrier signal consists of a large number of non-interfering sub-channels, separated in frequency by $\delta = 1/T$. δf is assumed to be small enough so that the channel gain and the noise power spectral density are constant over each of the sub-channels. So, the channel gain and noise power on the jth sub-channel can be given by H_j and N_j respectively. It is assumed that the noise is Gaussian. This assumption may seem questionable when the noise is primarily due to crosstalk, as on most cabled wire-pair media. However, the assumption becomes valid when there are many such interferers, and/or when they are passed through a narrow filter, as is the case in OFDM.

The variables in the optimization are the individually transmitted sub-channel powers P_j and the individual constellation sizes M_j. Fortunately, these optimizations are separable. We may first optimize the power distribution, under the constraint

$$\sum_j P_j = P. \tag{3.19}$$

It would appear that the constraint of an overall error probability is best met by requiring the same error probability for each of the sub-channels. This is indeed true at low error probabilities, which is the condition under study here, however it has been shown in general

recently [19] that this is sub-optimum. A further simplification that is usually made is to constrain the symbol error probability of each sub-channel to the target value, rather than the bit error probability. This can be significantly pessimistic, but the difference in the required signal-to-noise ratio is not too great at very low error probabilities, or when trellis coding is used. When fine precision is desired, the difference between symbol and bit error probability can be accounted for.

Setting the symbol error probability to the target value p requires that

$$KQ\left(\sqrt{\frac{3P_j|H_j|^2}{N_j(M_j-1)}}\right) = p. \tag{3.20}$$

This requires that we set the quantity

$$\Gamma^2 = \frac{3P_j|H_j|^2}{N_j(M_j-1)} \tag{3.21}$$

to the constant value such that

$$Q(\Gamma) = \frac{p}{4}. \tag{3.22}$$

The problem of optimizing the transmit powers of the sub-carriers is similar to a classical problem in information theory: given a linear channel with transfer function $H(f)$, Gaussian noise of power spectral density $N(f)$, and a transmit power constraint P, find the transmit power distribution that maximizes the capacity of the channel. The answer is the well-known "water pouring" solution [20]:

$$P(f) = \begin{cases} \lambda - \frac{N(f)}{|H(f)|^2}, & f \in F \\ 0, & f \in F' \end{cases} \tag{3.23}$$

where

$$F = \{f : \frac{N(f)}{|H(f)|^2} \leq \lambda\}, \tag{3.24}$$

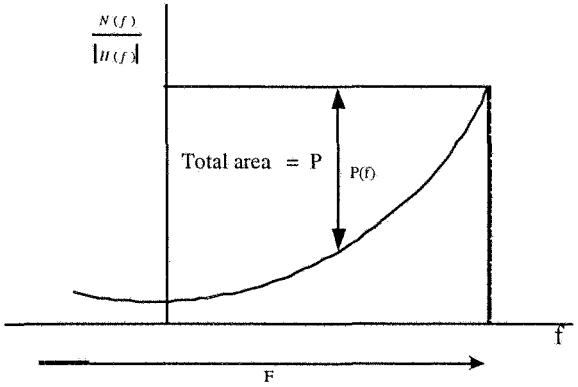


Figure 3.1: Optimum transmit power distribution by the water pouring theorem

and λ is the value for which

$$\int_F P(f)df = P. \tag{3.25}$$

This is illustrated in Figure 3.1. The curve of $N(f)/|H(f)|^2$ may be thought of as a bowl into which water whose area is equal to the allotted power is poured. The level of water is λ, and the distance between that level and the curve is the optimum transmit power spectrum. The range of F has been shown as contiguous for simplicity, but, in general, this need not be true.

It was shown in [21] that the optimum multi-carrier spectrum, under the assumptions given above, has a very similar form:

$$p_j = \begin{cases} \lambda - \frac{N_j\Gamma^2}{3P|H_j|^2}, & j \in F \\ 0, & j \in F' \end{cases} \tag{3.26}$$

where

$$F = \{j : \frac{N_j\Gamma^2}{3P|H_j|^2} \leq \lambda\}, \tag{3.27}$$

and λ is the value that causes

$$\sum_j P_j = P. \tag{3.28}$$

Note that over a region for which $N_j/|H_j|^2$ is small, P_j is approximately constant $\approx \lambda$. In a practical system with a low required error probability, we delete the region where $N_j/|H_j|^2$ is not low enough to support a sub-carrier with at least 4 constellation points at the required error probability. The result is an approximately flat spectrum over a region somewhat narrower than F. It was shown in [21] and verified in [22] that using a flat transmit spectrum results in very little degradation over the optimum when the required error probability is low.

A near-optimum design of a multi-carrier system therefore is to set all sub-carriers which cannot support a minimum size constellation to zero, and then, at least initially, to divide the allotted transmit power equally among the remaining sub-carriers. The next problem is to choose the constellation sizes of the non-zero sub-carriers.

Firstly, we will treat the constellation sizes as being continuous variables. Restricting M_j to an integer, which of course is essential, has almost no effect on the result. On the other hand, restricting $b_j = \log_2 M_j$ to an integer, which is convenient but not essential, does have a significant effect. Solving Equation 3.21,

$$M_j = 1 + \frac{3P_j|H_j|^2}{N_j\Gamma^2} \tag{3.29}$$

Rounding down to the next lower integer has little effect, and we have now found an approximation to the problem of maximizing the overall bit rate under an error probability constraint. The resultant bit rate is

$$R_b = \frac{1}{T}\sum_j \log_2\left[1 + \frac{3P_j|H_j|^2}{N_j\Gamma^2}\right] \tag{3.30}$$

Since the sub-channels are closely spaced, the above may be closely approximated by the integral

$$R_b = \int_F \log_2\left[1 + \frac{3P(f)|H_j|^2}{N(f)\Gamma^2}\right] df. \tag{3.31}$$

The above is the same as the well-known formula for the capacity of the channel, in which the signal-to-noise ratio,

$$\rho(f) = \frac{P(f)|H_j|^2}{N(f)} \tag{3.32}$$

has been decreased by a factor of $\frac{3}{\Gamma^2}$. This factor has been referred to as the "gap" [23]. In other words, the actual rate achievable, at a given error probability, on a given channel by an uncoded multi-carrier system is equal to the capacity of the same channel in which the signal-to-noise ratio has been reduced by the gap.

The above integral (Equation 3.31), except for a small change in the range of integration, is the same as the rate achievable over the same channel by a single carrier system using an ideal decision feedback equalizer (DFE). This leads to the very interesting result that at low error probability, the performance of an optimized multi-carrier system is approximately the same as that of a single carrier system with an ideal DFE [25]. This is particularly true when compared with an MMSE DFE as opposed to a zero-forcing DFE [26]. At higher error probabilities, the multi-carrier system has been shown to perform somewhat better [19].

We can rewrite Equation 3.30 as:

$$R_b = \frac{1}{T}\sum_j \log_2[1 + \frac{3\rho_j}{\Gamma^2}] = \frac{1}{T}\log_2\{\prod_j[1 + \frac{3\rho_j}{\Gamma^2}]\}. \tag{3.33}$$

Now define an average signal-to noise ratio [23] such that

$$R_b = \frac{1}{T}\log_2[1 + \frac{3\bar{\rho}}{\Gamma^2}], \tag{3.34}$$

Then, with n the number of sub-channels,

$$\bar{\rho} = \frac{\Gamma^2}{3}\{(\prod_{j=1}^{n}[1 + \frac{3\rho_j}{\Gamma^2}])^{\frac{1}{n}} - 1\} \tag{3.35}$$

and at high signal-to-noise ratio,

$$\bar{\rho} = [\prod_{j=1}^{n}\rho_j]^{\frac{1}{n}}. \tag{3.36}$$

Thus, at high signal-to-noise ratio, when a multi-carrier system is fully optimized, the average signal-to-noise ratio is the geometric mean of the signal-to-noise ratios of the sub-channels. This is an important quantity, and serves as a good quality measure of the channel. The integral form of the above appears in analysis of single carrier systems with DFE.

So far we have not required that the number of bits carried by each sub-carrier be an integer. Such non-integral values can be achieved by treating more than one sub-carrier as a constellation in a higher-dimensional space, and assigning an integral number of bits to that higher-dimensional constellation. However, to avoid that complexity, it is usually required that b_j be an integer. Simply performing the above optimization and rounding each b_j down results in the average loss of $\frac{1}{2}$ bit per sub-carrier.

The above loss can be mostly recovered by small adjustments in the sub-carrier powers. Those sub-carriers that required a large round-down can have their powers increased slightly in order to achieve one more bit. Those with a smaller round-down can have their power reduced while keeping the same number of bits. During this process of

power re-allocation, it is important to ensure that the constraint of total power is kept. After the process is completed, the power deviations will vary over a 3 dB range.

A further small increase in bit rate can be achieved by accounting for the bit error probability of each sub-carrier rather than the symbol error probability, which is higher. This involves calculating the bit error probability for each sub-carrier after the system has been optimized for symbol error probability. Those sub-carriers which can meet the bit error probability target with an addition bit can then have their constellations increased accordingly.

3.4 Bit and Power Allocation Algorithms for Fixed Bit Rate

A more common design requirement is to minimize the bit error probability for a fixed bit rate, where that error probability must be below some target value. Again, we will assume a transmitter power constraint, and integral number of bits per sub-channel. Several algorithms have been developed to solve this design problem, with varying degrees of precision, complexity, and computation time [26]. One approach is similar to the previously described procedure for maximizing a variable bit rate. As before, we first determine Γ such that $Q(\Gamma) = p/4$, where p is the desired error probability, and is conservatively chosen as the maximum symbol error probability for each sub-carrier. We now examine if the un-quantized total bit rate

$$R_b = \frac{1}{T} \sum_j \log_2[1 + \frac{3P_j|H_j|^2}{N_j\Gamma^2}] \tag{3.37}$$

is greater than the required rate. If not, the desired performance cannot be achieved. Otherwise, if $R_b - B \geq \frac{A}{T}J$, where B is the required bit rate, A is a positive integer, and J is the number of used sub-carriers,

then A bits are subtracted from each sub-carrier. This step adds 3 dB of noise margin.

After the above procedure, the number of bits on each sub-carrier is rounded down to the next lowest integer. The resultant bit rate $R^{'}$ is compared with B. If $R^{'} > B$, then a set of sub-carriers with the smallest round-off are each reduced by an additional bit. If $R^{'} < B$, then a set of sub-carriers with the largest round-off are each increased by one bit. Finally, the powers of the sub-carriers are adjusted as before to achieve the same error probability for each.

Chapter 4

Clipping in Multi-Carrier Systems

4.1 Introduction

It is widely recognized that a serious problem in OFDM is the possibility of extreme amplitude excursions of the signal. The signal is the sum of N independent (but not necessarily identically distributed) complex random variables, each of which may be thought of as a Quadrature Amplitude Modulated (QAM) signal of a different carrier frequency. In the most extreme case, the different carriers may all line up in phase at some instant in time, and therefore produce an amplitude peak equal to the sum of the amplitudes of the individual carriers. This occurs with extremely low probability for large N.

The problem of high peak amplitude excursions is most severe at the transmitter output. In order to transmit these peaks without clipping, not only must the D/A converter have enough bits to accommodate the peaks, but more importantly the power amplifier must remain linear over an amplitude range that includes the peak amplitudes. This

leads to both high cost and high power consumption.

Several researchers have proposed schemes for reducing peak amplitude by introducing redundancy in the set of transmitted symbols so as to eliminate those combinations which produce large peaks [28, 29]. Clearly the required redundancy increases with the desired reduction in peak amplitude. This redundancy is in addition to any other coding used to improve performance, and leads to a reduction in the carried bit rate.

In this chapter, we will examine the effects of clipping an unconstrained OFDM signal. Several previous papers have analyzed the clipping as an additive Gaussian noise. This approach is reasonable if the clipping level is sufficiently low to produce several clipping events during an OFDM symbol interval. In most realistic cases, however, particularly when the desired error probability is low, the clipping level is set high enough such that clipping is a rare event, occurring substantially less than once per OFDM symbol duration. Clipping under these conditions is a form of impulsive noise rather than a continual background noise, leading to a very different type of error mechanism. Here we evaluate the rate of clipping with a high clipping threshold, and the energy and spectrum of those events. This permits the determination of the error probability and interference into the adjacent channels.

The transmitter output of a multi-carrier system is a linear combination of complex independent random variables. It is generally assumed that the distribution of that output signal is Gaussian using the central limit theorem. This assumption is valid for large N over the range of interest.

The output of the transmitter Inverse Fourier Transform, $\{d_m\}$ is a linear combination of complex independent identically distributed symbols $\{D_n\}$ of the corresponding constellations:

$$d_m = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} D_n e^{j2\pi n \frac{m}{N}}. \tag{4.1}$$

Assuming Gaussian distribution for transmitted samples, signal "peak" should be defined carefully. Peak-to-average has commonly been used, yet a clear definition of peak is not explicitly presented. Absolute peak is not a proper parameter in our analysis because the probability of co-phasing N complex independent random variables is very small, even for moderate values of N. A proper definition should take into account statistical characteristics of the signal and the fact that Gaussian distribution tails are extended infinitely. A reasonable alternative to absolute maximum is to define peak of signal power (variance) such that probability of crossing that level is negligible. For example, a threshold of 5.2 times the average RMS (5.2σ), corresponds to a crossing probability of 10^{-7}. Such a definition of peak-to-average ratio of a Gaussian signal, unlike the absolute maximum, is independent of the number of sub-carriers.

4.2 Power Amplifier Non-Linearity

Several analytical models for power amplifier non-linearity are offered in the literature [36]. Amplitude and phase non-linearity of power amplifiers known as AM/AM and AM/PM non-linearity results in an inter-modulation distortion and requires operation of amplifiers well below a compression point as shown in Figure 4.1.

For a narrowband Gaussian input signal[1] $x(t)$ of the form:

$$x(t) = I(t)\cos\omega t - Q(t)\sin\omega t = R(t)cos(\omega t + \varphi(t)), \tag{4.2}$$

the output of the non-linear device using AM/AM and AM/PM non-linearity model is:

[1]It means the ratio of signal bandwidth to carrier frequency is small. The majority of wireless modulation systems meet this requirement.

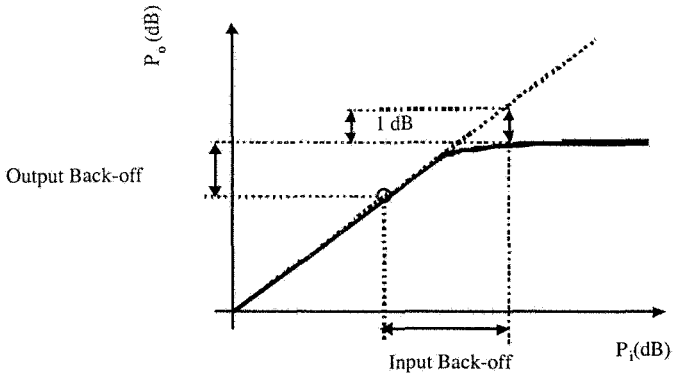


Figure 4.1: Power amplifier 1 dB compression point

$$y(t) = A(R(t))\cos(\omega t + \varphi(t) + \theta(t)). \tag{4.3}$$

Let $f(\cdot)$ denote the instantaneous mapping of the non-linear function. Then $A(\cdot)$ and $f(\cdot)$ are related through the Chebychev transformation [42]:

$$A(R) = \frac{2}{\pi}\int_0^{\pi} f(R\cos\varphi)\cos\varphi d\varphi \tag{4.4}$$

A proper clipping model for the signal and its envelope is critical in this analysis. We assume a soft amplitude limiter model for the power amplifier. The relationship between signal clipping and envelope clipping transfer function according to Equation 4.4 is shown in Figure 4.2.

In the first case, the Gaussian signal $x(t)$ is clipped and higher frequencies generated by the non-linearity are filtered out at the output of the amplifier. In the latter, the amplitude $R(t)$, which is a Rayleigh process, is clipped and then the output harmonics are filtered out.

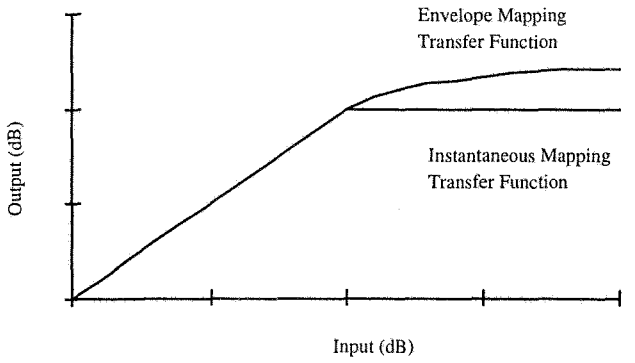


Figure 4.2: Relationship between instantaneous and envelope clipping

4.3 Error Probability Analysis

We[2] will be concerned with stationary Gaussian processes, possessing finite second-order moments, and continuous with probability one. The following *three* asymptotic properties of the large excursions of such processes will form the basis of our analysis in the sequel. These properties are originally due to Rice [37] and have also been elaborated on or discussed by several subsequent authors [34, 35, 38, 40, 42]. They are depicted conceptually in Figure 4.3.

- The sequence of upward level crossings[3] of a stationary and ergodic process, having continuous sample function with probability one, asymptotically approaches a Poisson process for large levels . For a Gaussian signal $x(t)$, the rate of this Poisson

[2]This section is taken from [44], which is a corrected version and extension of the material in the first edition of this book.

[3]A random process $x(t)$ is said to have an upcrossing of the level l at time t_0 if $\exists$ $(\epsilon > 0)$: $x(t) \leq l$ for $t \in (t_0 - \epsilon, t_0)$, and $x(t) > l$, and $x(t) > l$ for $t \in (t_0, t_0 + \epsilon)$

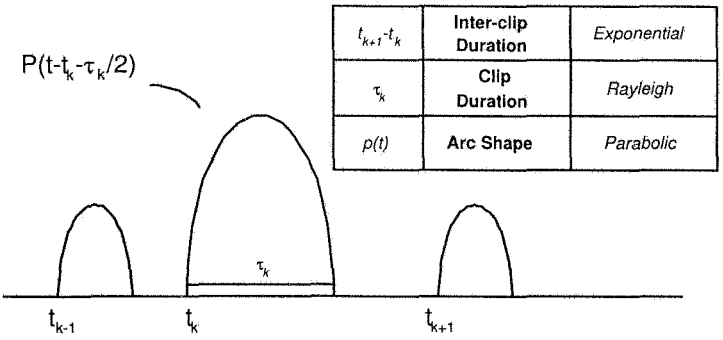


Figure 4.3: Excursions of a Gaussian random process above $|l|$

process is given by [40]

$$\lambda_l = \frac{1}{2\pi}\sqrt{\frac{m_2}{m_0}}\exp\left(\frac{-l^2}{2m_0}\right) \tag{4.5}$$

where

$$m_i \doteq \frac{1}{2\pi}\int \omega^i S_x \omega d\omega,\ i = 0, 2, \cdots \tag{4.6}$$

and $S_x(\omega)$ is the power spectral density of $x(t)$. Without loss of generality, we can normalize the power m_0 in the signal $x(t)$ to unity. The relevant quantity m_2 then represents the power in the first derivative of $x(t)$.

- The length of intervals τ during which the signal stays above the high level $|l|$ is (asymptotically) Rayleigh distributed [38], with density function

$$\rho_\tau(\tau) = \frac{\pi}{2}\frac{\tau}{\tau_m^2}e^{-(\pi/4)(\tau/\tau_m)^2},\ \tau \geq 0 \tag{4.7}$$

Here, τ_m denotes the expectation of τ. Since $\lambda_l \tau_m = Pr(x(t) \geq l)$, the expected value of the duration of a clip may be approximated as [38]

$$\tau_m = \frac{Q(l)}{\lambda_l} \approx \frac{\sqrt{2\pi}}{l\sqrt{m_2}} \tag{4.8}$$

where

$$\begin{aligned} Q(z) &= \tfrac{1}{\sqrt{2\pi}} \int_z^\infty e^{-u^2/2} du \\ &= \tfrac{1}{\sqrt{2\pi}z} e^{-z^2/2} (1 - \tfrac{1}{z^2} + \tfrac{3}{z^4} - \cdots) \end{aligned} \tag{4.9}$$

The approximation in Equation 4.8 is valid for $l > 3\sigma$.

- The shape of pulse excursions above level $|l|$ are parabolic arcs [35, 37] of the form

$$p_\tau(t) = \left(-\frac{1}{2} l m_2 t^2 + \frac{1}{8} l m_2 \tau^2 \right), |t| \leq \frac{T}{2} \tag{4.10}$$

where $rect(.)$ denotes a rectangular window function, and τ, the (random) duration of the clip, forms the support of the parabolic arc (Figure 4.3). The parabolic shape of the signal's crossings above high level l follows naturally from a Taylor's series approximation about the peak value (CW. [42], and references therein). We will also require the *instantaneous* spectrum of the clipped signal component for subsequent analysis. This can be obtained by taking the Fourier transform of the parabolic arc Equation 4.10, to get

$$g_\tau(\omega) = m_2 \frac{l\tau}{\omega^2} (sinc\frac{\omega\tau}{2} - \cos\frac{\omega\tau}{2}) \tag{4.11}$$

which is plotted in Figure 4.4 ($sinc\varphi \doteq \frac{\sin\varphi}{\varphi}$). Note that the instantaneous spectrum is not flat over all frequencies.

In this section, we will compute a bound on the bit-error floor caused by clipping in a multi-carrier system (in the absence of noise and other channel impairments). The basic system model used for this analysis is presented next.

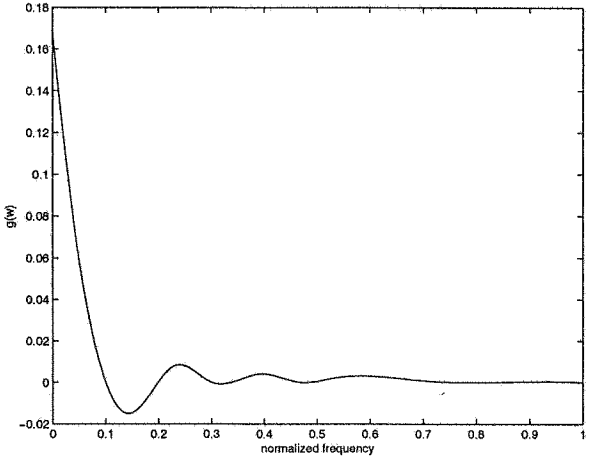


Figure 4.4: Normalized instantaneous distortion spectrum

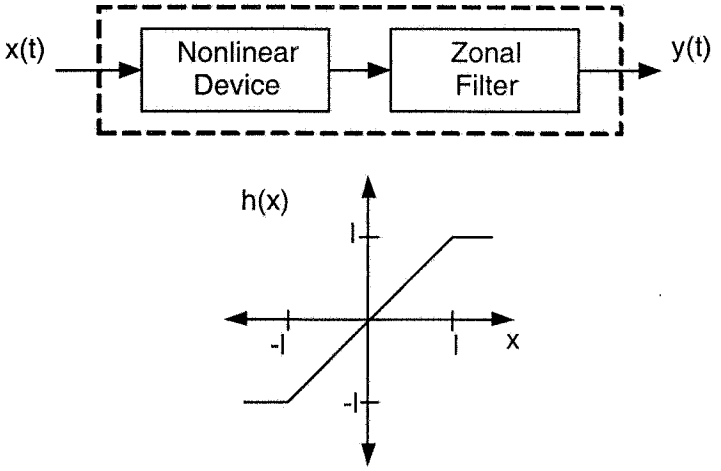


Figure 4.5: Memoryless nonlinear mapping

4.3.1 System Model

Consider a continuous time, unconstrained, baseband multi-carrier signal $x(t)$, input to a soft (clipping) nonlinearity $h(.)$, to produce the (clipped) output waveform (Figure 4.5). The non-linear transfer function may be described, e.g.,

$$y = h(x) = \begin{cases} -l, & x \leq l \\ x, & |x| < l \\ l, & x \geq l \end{cases} \tag{4.12}$$

where explicit dependence on time is omitted from the notation above. For a memoryless nonlinearity such as Equation 4.12, the output $y(t)$ may be decomposed by Bussgang's theorem [41] into two uncorrelated signal components

$$y(t) = \gamma x(t) + c_L(t) \tag{4.13}$$

where the value of γ that un-correlates the "error" and input signal $x(t)$ is given by

$$E[x(t+\tau)y(t)] = \gamma E[x(t+\tau)x(t)] \tag{4.14}$$

from which it is straightforward to show that $\gamma \approx 1$ for $l \geq 1$. We can, thus, use a linear model with uncorrelated components to analyze the effect of clipping on the transmitted signal.

Under the assumption of many subcarriers, the power spectrum of the $x(t)$ will tend to rectangular over a total frequency $f_0 = N/T$ (T = OFDM symbol duration). Hence, the rate of the Poisson process corresponding to the clipping of $x(t)$ can be defined from Equation 4.5 as

$$\lambda_l = \frac{1}{2\pi}\sqrt{m_2}e^{\frac{-l^2}{2}} = \frac{f_0}{\sqrt{3}}e^{\frac{-l^2}{2}} \tag{4.15}$$

since $m_2 = (2\pi f_0)^2/3$, according to Equation 4.6. Also, the expected value for the duration of a clip will be given [from Equation 4.8] by

$$\tau_m^{-1} \approx \sqrt{\frac{2\pi}{3}} f_0 l \tag{4.16}$$

4.3.2 BER Due to Clipping

Having defined the relevant parameters, we now turn to evaluating the conditional rate of symbol error due to clipping, which is required to compute the overall symbol error probability. Define as the event of clipping a signal above level $|l|$ in any given symbol interval (and A^c as its complement). Then, by the law of total probability

$$\begin{aligned} Pr(symbol\ error) &= Pr(symbol\ error|A)Pr(A) \\ &+Pr(symbol\ error|A^c)Pr(A^c) \end{aligned} \tag{4.17}$$

In order to compute this probability, we will first evaluate the effect of a clip on each of the individual subchannels. The effect of a clip of duration τ, occurring at time t_0, on the kth sub-channel is given by the Fourier transform of the clipped portion of the pulse,

$$F_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f_n e^{-j2\pi(kn/N)} \tag{4.18}$$

where f_n are samples of the clipped pulse $p(t)$ in Equation 4.10

$$f_n = p(t - t_0 - \tau/2)|_{t=(nT/N)} \tag{4.19}$$

and the factor $1/\sqrt{N}$ preserves total power. We can replace the Discrete Fourier Transform with the conventional one by substituting

$$f_n = \frac{1}{\delta} \int_{n\delta-\delta/2}^{n\delta+\delta/2} n(t)dt. \tag{4.20}$$

where $\delta = T/N$. This substitution is valid since $\delta \ll T$. Then

$$F_k = \frac{N}{T}\frac{1}{\sqrt{N}}\int_{t_0}^{t_0+\tau} p(t - t_0 - \tau/2)e^{-j2\pi(kt/T)}dt. \tag{4.21}$$

Substituting $u = t - t_0$ in Equation 4.21 yields

$$\begin{aligned} F_k &= \tfrac{\sqrt{N}}{T}e^{(-j2\pi kt_0/T)}\textstyle\int_0^\tau p(u - \tau/2)e^{-j2\pi(ku/T)}du \\ &= \tfrac{\sqrt{N}}{T}e^{-(j2\pi k(t_0+\frac{\tau}{2})/T)}g(\tfrac{2\pi k}{T}) \end{aligned} \tag{4.22}$$

where $g(\cdot)$ is pulse spectrum given by Equation 4.11. Thus

$$F_k = \frac{\sqrt{N}m_2Tl\tau}{4\pi^2k^2}e^{-(j2\pi k(t_0+\tau/2)/T)}\left(sinc\left(\frac{\pi k\tau}{T}\right) - \cos\left(\frac{\pi k\tau}{T}\right)\right). \tag{4.23}$$

For $\alpha \ll l(\tau \ll T)$, we can use the expansion

$$sinc(\alpha) - cos(\alpha) \approx \frac{\alpha^2}{3} \tag{4.24}$$

in Equation 4.23 in order to approximate the response of the kth sub-carrier to a clip of duration τ as

$$F_k = \frac{1}{12T}\sqrt{N}m_2l\tau^3\exp(j\theta). \tag{4.25}$$

In Equation 4.25, is uniformly distributed over $[0, 2\pi]$, and the Rayleigh probability distribution of τ is given by $\rho_\tau(\tau)$in Equation 4.7. Note that the expansion in Equation 4.24 will not be valid in general for the higher frequency sub-carriers in the multi-carrier signal. In fact, the nature of the distortion spectrum in Equation 4.22 would indicate that the probabilities of error due to a clip will vary across sub-carriers, the overall probability being dominated by that of the lower subcarriers[4]. We can, however, use the response in Equation 4.25 to

[4]This would be true if the constellation sizes on the different sub-carriers are equal, or even more so when the lower sub-carriers carry larger constellations, as is typically the case in ADSL (due to bit loading)

compute an upper bound on the error probability due to clipping. Rewriting Equation 4.25 as

$$F_k = \eta \exp(j\theta) \tag{4.26}$$

we are interested in obtaining the probability distribution of η

$$Pr(\eta > R) = Pr\left[\tau > \left(\frac{12RT}{\sqrt{N} m_2 l}\right)^{1/3}\right]. \tag{4.27}$$

Using the distribution of τ in Equation 4.7, and substituting for m_2 from Equation 4.6, we get, after some simplification

$$Pr(\eta > R) = exp - \left[\frac{3\pi^2 R^2 l^4 N}{8}\right]^{1/3}. \tag{4.28}$$

If we warp the complex plane by the mapping $ae^{j\theta} \rightarrow a^{1/3}e^{j\theta}$, then Equation 4.28 would represent a Rayleigh distribution in the warped complex plane. The real or imaginary component of F_k, therefore, has a normal distribution with

$$Pr(\eta \cos\theta > x) = Q\left(\frac{x^{1/3}}{\sigma}\right) \tag{4.29}$$

where

$$\sigma \doteq \left(\frac{2}{\sqrt{3N}\pi l^2}\right)^{1/3}.$$

To compute the error probability, we will assume that each subcarrier carries a square constellation of L^2 points. Each (real/imaginary) component has L levels, equally spaced and separated by $2d$, with total power equal to $1/2N$ (since we have normalized the total power to unity). Thus

$$d = \sqrt{\frac{3}{2N(L^2-1)}}. \tag{4.30}$$

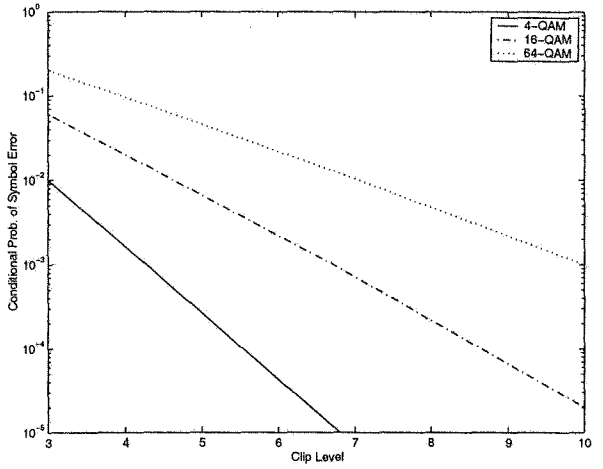


Figure 4.6: Conditional probability of symbol error

Combining Equation 4.28 and Equation 4.30, we get the probability of error in any particular subcarrier, given that a clip occurs, as

$$\begin{aligned} Pr(error|A) &= \frac{4(L-1)}{L} Q\left(\frac{d^{1/3}}{\sigma}\right) \\ &= \frac{4(L-1)}{L} Q\left(\left[\frac{3\pi l^2}{\sqrt{8(L^2-1)}}\right]^{1/3}\right) \end{aligned} \tag{4.31}$$

Note that this quantity is independent of N, the total number of sub-carriers. This is because although the effect of a clip increases with the number of sub-carriers as given by Equation 4.25, the signal power in each component is inversely proportional to N. Also, not only does the probability of occurrence of a clip decrease with the clipping level [from Equation 4.15], but so does the probability of error caused by each occurring clip. The conditional error probability

in Equation 4.31 is plotted in Figure 4.6 for different constellation sizes L^2.

To compute the overall probability of error, we will need to multiply the conditional probability in Equation 4.31 with the probability of occurrence of a clip in one symbol duration. This can be obtained using Equation 4.5, the Poisson rate of clipping events, as

$$Pr(clip) = 1 - e^{-\lambda_l T} \approx \lambda_l T \tag{4.32}$$

For two-sided clipping, we will have

$$Pr(A) = 2\lambda_l T \tag{4.33}$$

so that the overall probability of symbol error is upper bounded by

$$Pr(error) = \frac{8N(L-1)}{\sqrt{3}L} e^{-(l^2/2)} Q\left(\left[\frac{3\pi l^2}{\sqrt{8(L^2-1)}}\right]^{1/3}\right) \tag{4.34}$$

In particular, for $L = 2$ (QPSK), which is typical in several wireless OFDM systems, the error probability is extremely small, being given by

$$Pr(error) = \frac{4N}{\sqrt{3}} e^{-(l^2/2)} Q\left(\left[\sqrt{\frac{3}{8}}\pi l^2\right]^{1/3}\right) \tag{4.35}$$

On the other hand, for large constellation sizes, the error probability may be approximated as

$$Pr(error) = \frac{8N}{\sqrt{3}} e^{-(l^2/2)} Q\left(\left[\frac{3\pi l^2}{\sqrt{8}L}\right]^{1/3}\right). \tag{4.36}$$

The Poisson probability in Equation 4.32 would hold for a continuous time signal, i.e., an infinitesimal sampling of the multi-carrier waveform. Since practical transceivers operate at Nyquist frequency (or over-sampled versions of this frequency), we can tighten the error

bound in Equation 4.34 by replacing this clip probability with a corresponding discrete time value. Specifically, the probability of a single sample exceeding the level $|l|$ is $2Q(l)$, so assuming the samples to be i.i.d.[5], the probability of clipping m samples within an OFDM symbol will be given by the binomial distribution

$$\begin{aligned} p_m &= \binom{N}{m} (1-2Q(l))^{N-m}(2Q(l))^m \\ &\approx 0 \; for l \gg 1 \; and \; m > 1 \end{aligned} \tag{4.37}$$

and $p_m \approx 2NQ(l)$ for $m = 1$. The error probability in Equation 4.38 then becomes

$$Pr(error) = \frac{8N(L-1)}{L} Q(l) Q\left(\left[\frac{3\pi l^2}{\sqrt{8(L^2-1)}} \right]^{1/3} \right) \tag{4.38}$$

This probability $P(error)/N$ is plotted in Figure 4.7 for various typical constellation sizes L^2.

It is interesting to compare these results with those obtained using the "additive Gaussian noise" model, which would hold if the clip level were set low enough to permit several clipping events per OFDM symbol. In reference [32], the power in the clipped portion $c_L(t)$ of the input waveform $x(t)$ has been computed as

$$\begin{aligned} \sigma_c^2 &= 2\int_1^\infty (x-l)^2 e^{-x^2/2} dx \\ &= -\sqrt{\tfrac{2}{\pi}} l \exp(-\tfrac{l^2}{2}) + 2(1+l^2)Q(l) \end{aligned} \tag{4.39}$$

where we have normalized the result to unity signal power. This result is pessimistic since it does not account for the fact that some of the

[5]It is known that the IDFT of statistically independent, Gaussian spectral points produces statistically independent time domain samples. In the presence of filtering, such an independent identically distributed (i.i.d.) assumption would not be strictly valid, as filtering would introduce correlation among the time domain samples. The binomial distribution should, therefore, be seen as a simplifying approximation.

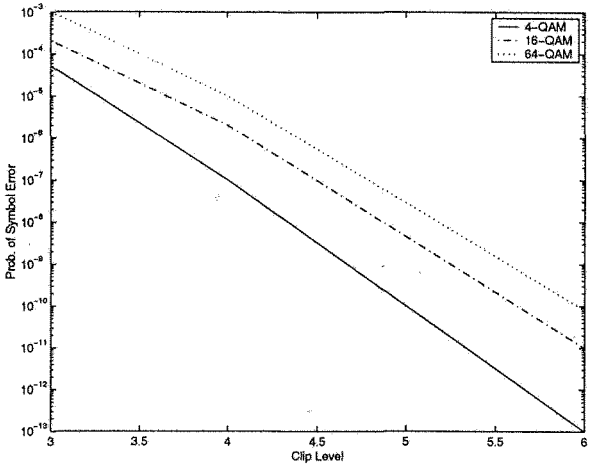


Figure 4.7: Symbol error probabilities due to clipping

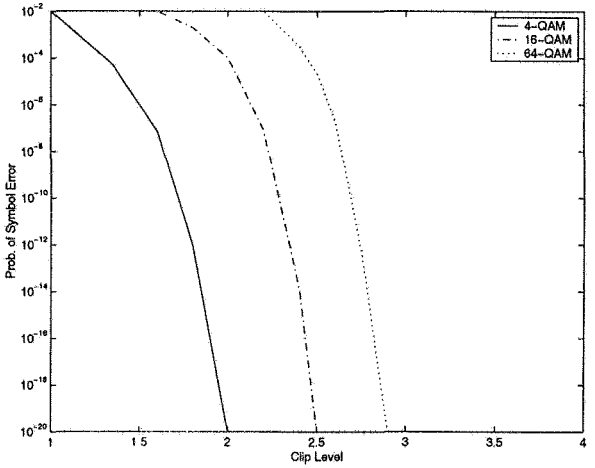


Figure 4.8: Symbol error probabilities using the additive Gaussian noise approach

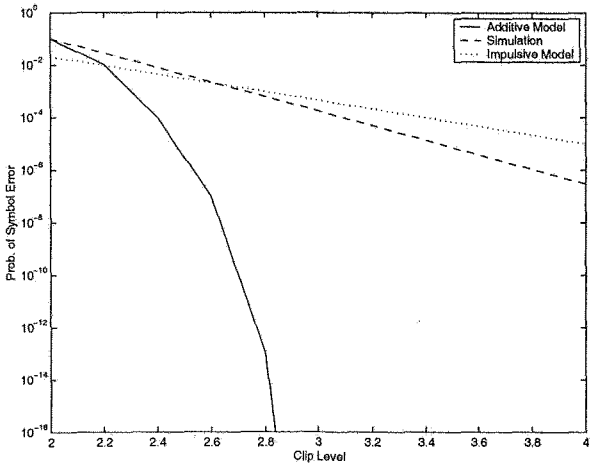


Figure 4.9: Simulated and analytical symbol error probability floor due to clipping

noise power will fall out-of-band for the OFDM band of interest, and will, therefore, not cause any in-band distortion. Nevertheless, for the (Gaussian) noise power level in 4.39, the probability of error for any one subcarrier would be given by

$$Pr(error) = 4\frac{L-1}{L}Q\left(\frac{\sqrt{3}}{\sigma_c\sqrt{(L^2-1)}}\right). \qquad (4.40)$$

This probability is plotted in Figure 4.8. It is easy to see from this curve that the error probabilities as computed by the additive noise model are substantially lower than those obtained using our analysis, even under the pessimistic approximation described above. This is because instead of being spread uniformly over time, as is assumed in [30, 31, 32, 33], noise due to clipping is in fact concentrated in time as impulses, leading to a greater probability of error over the various sub-carriers.

The validity of the two models discussed in this section was con-

firmed by simulating the error probability of a 64-tone OFDM signal modulated using a 64-QAM constellation, soft clipped at various levels of output amplitude. The results, shown in Figure 4.7, suggest that while the additive noise model is reasonable for low clipping thresholds, it underestimates the error probability by several orders of magnitude for the higher clip levels. This implies that we must account for the impulsive nature of the clipping process in estimating error probabilities for high clip thresholds.

We can also use Equation 4.38 to compute an approximate bound on the probability of error *per bit*. Assuming Gray coded levels in each subcarrier component, and noting that errors beyond adjacent levels are negligible, we get the probability of bit error

$$Pr_b(error) \approx \frac{4N(L-1)}{L\log_2 L} Q(l) Q\left(\left[\frac{3\pi l^2}{\sqrt{8(L^2-1)}}\right]^{1/3}\right). \quad (4.41)$$

Note that this value will also hold in the presence of fading, as long as the fading is slow compared with the OFDM symbol rate, and the clip occurs at the *transmitter*. In practice, however, we are mostly concerned with BER performance of the OFDM *receiver*. The analysis for clipping at the receiver (in the presence of channel impairments) is presented in Section 4.4

4.4 Performance in AWGN and Fading

Consider first the case when additive white Gaussian noise of (random) amplitude n is also present on the channel. In this case, the per-bit error probability in a lower subcarrier component would be [from Equation 4.29]

$$Pr_b(error|clip, noise) = \frac{2(L-1)}{L\log_2 L}\int Pr(r\cos\theta > d-n)p_N(n)dn$$
$$= \frac{2(L-1)}{L\log_2 L\sqrt{2\pi}\sigma_n}\int Q[\frac{(d-n)^{1/3}}{\sigma}]e^{(-n^2/2\sigma_n^2)}dn \quad (4.42)$$

where σ_n^2 is the power (variance) of the AWGN, and the effect of the AWGN is to reduce the effective distance between various constellation points (from d to $d-n$). From Equation 4.17, the overall error probability is upper bounded by

$$Pr_b(error|clip, noise) = \frac{2(L-1)}{L\log_2 L}\left(\begin{array}{l}\int \frac{\sqrt{2}NQ(l)}{\sqrt{\pi}\sigma_n}Q[\frac{(d-n)^{1/3}}{\sigma}] \\ .exp(-\frac{n^2}{2\sigma_n^2})dn + Q(\frac{d}{\sigma_n})\end{array}\right) \quad (4.43)$$

To examine the effect of fading on these results, consider a noiseless, flat-fading channel, modelled by a slowly varying Rayleigh random variable with long-term distribution function

$$p_Z(z) = \frac{\pi}{2}ze^{-(\pi z^2/4)} \quad z > 0 \quad (4.44)$$

where the mean has been normalized to unity, without loss of generality. We assume the channel to be constant over one OFDM symbol duration, and that this value is known in advance to the transmitter to appropriately set its clipping level. The error probability for a symbol during which the channel amplitude variable is z is then

$$Pr(error|z) \approx 8\frac{N(L-1)}{L}Q(lz)Q\left(\left[\frac{3\pi l^2 z^2}{\sqrt{8(L^2-1)}}\right]^{1/3}\right) \quad (4.45)$$

Thus, the overall error probability in this case is upper-bounded by

$$Pr(error) = 4\pi \frac{N(L-1)}{L} \int_0^\infty z e^{(-\pi z^2/4)} Q(lz) Q([\frac{3\pi l^2 z^2}{\sqrt{8(L^2-1)}}]^{1/3}) dz \quad (4.46)$$

which can be expressed (for purposes of numerical computation) in definite integral form as

$$Pr(error) = \frac{4N(L-1)}{\pi L} \int_0^\infty \int_0^{\pi/2} \int_0^{\pi/2} z .exp(\frac{-z^2}{4}[\frac{\pi}{2} + \frac{2l^2}{\sin^2\varphi} + \frac{\xi^{1/3}}{\sin^2\psi}]) d\varphi d\psi dz \quad (4.47)$$

where

$$\xi \doteq \frac{9\pi^2 l^4}{z^2(L^2-1)}. \quad (4.48)$$

Note that the alternate representation of the Q function [43] has been used to derive Equation 4.47 from Equation 4.46. This error probability [in Equation 4.47] is computed numerically and plotted in Figure 4.10. It is clear from this figure that the performance with fading is far worse than that obtained in the absence of fading, discussed in the previous section. Therefore, it appears that while the OFDM signal is relatively insensitive to clipping at the transmitter, the performance is somewhat degraded in the presence of channel fades, together with clipping at the receiver. Of course, Equation 4.46 only represents an upper bound on the probability of error due to clipping, that is tight for lower sub-carriers of the multi-carrier signal.

In order to validate the analytical results presented in this section, as well as to generalize these results for the case of a multi-path fading channel, extensive, $bit - true$ simulations were performed using the OFDM baseband equivalent system model shown in Figure 4.11. A 64-point FFT signal with 52 "active" tones was simulated, and a guard interval of length 16 was added to each OFDM symbol as cyclic prefix to compensate for ISI. Two channels were considered: 1) a Rayleigh

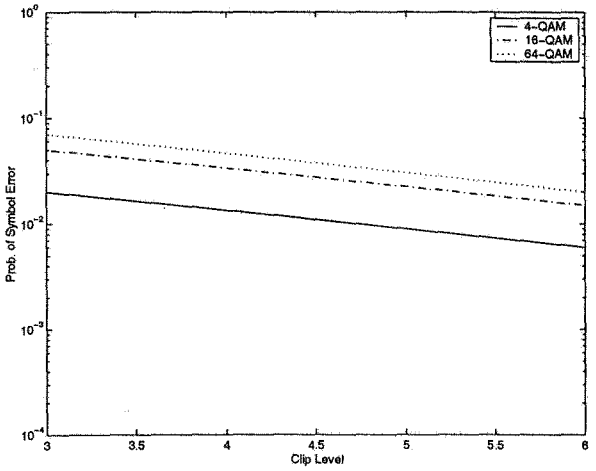


Figure 4.10: Symbol error probability floor in a Rayleigh fading channel

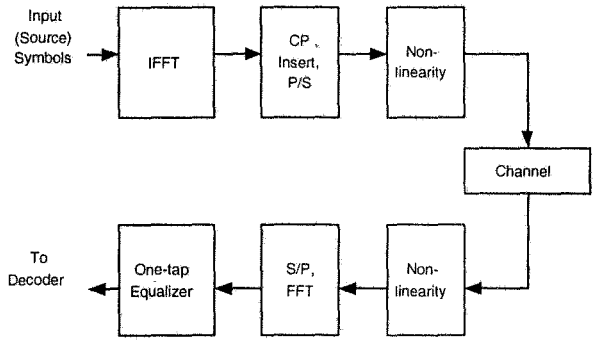


Figure 4.11: Baseband equivalent system model for clipping at OFDM receiver

(no diversity) channel modelled using a single "tap" of average power 0 dB, corresponding to a severe, flat fading environment and 2) a three-path WSSUS multi-path channel with an exponentially decaying power profile and RMS delay spread of 100 ns, corresponding to a typical indoor environment. Perfect synchronization and a one-tap, post-FFT equalizer with perfect channel estimates were assumed in order to observe the effect of the clips alone. In each case the bit-error probability was recorded as a function of the received signal SNR, averaged over a large number (10^4) of fading instances. The resulting error probability curves are presented in Figures 4.12-4.15. Several observations can be made from these curves. First, as expected, the bit-error performance due to clipping deteriorates with increasing constellation size L^2, and improves as the clip level is increased. Second, the performance in Rayleigh (flat) fading is observed to be consistently worse than that seen in AWGN, for a given clip level and constellation size. Also, a comparison of Figures 4.12 and 4.13 with the analytical curves in Figures 4.8 reveals that the residual symbol error floor due to clip-ping (even at the high SNRs) is within an order of magnitude of the theoretical bound computed in Equation 4.46.

Other results, not shown here, also indicated that the BER performance was not equal on all sub-carriers; the lower sub-carriers being subject to about 1 to 1.5 dB greater degradation than those at the extremes. This is in accordance with our observation about the nature of the instantaneous clipping spectrum. The error floor for the "worst" affected sub-carrier was calculated and used to compare with the analytical bound for error probability described in Equation 4.46. As evident from Figures 4.10 and 4.16, the theoretical and simulated results are in good agreement for the higher clip levels, where the conditions discussed in Section 2 for the rare event model are satisfied. This shows that the derived analytic bound is tight for lower sub-carriers of the multi-carrier signal.

Apart from the above discussion about error performance, there are several other practical (implementation) issues associated with clipping at the receiver as well. For example, in a packet commu-

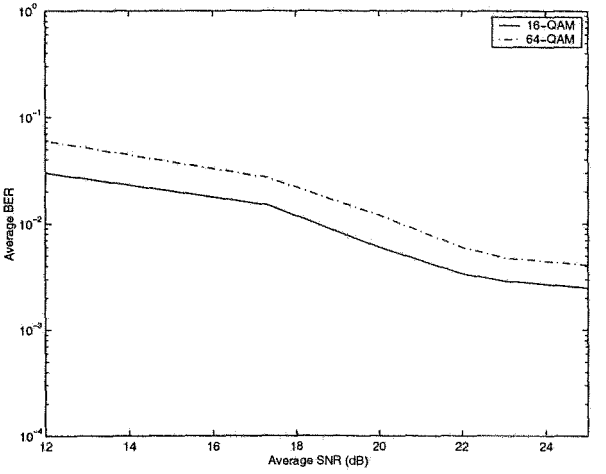


Figure 4.12: BER in a Rayleigh fading channel, clip level = 3σ

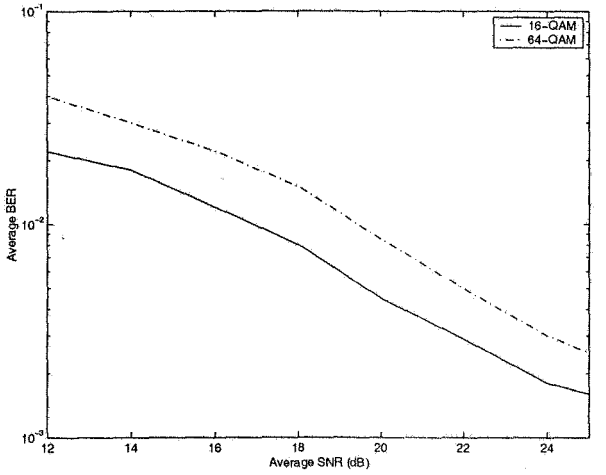


Figure 4.13: BER in a Rayleigh fading channel, clip level = 6σ

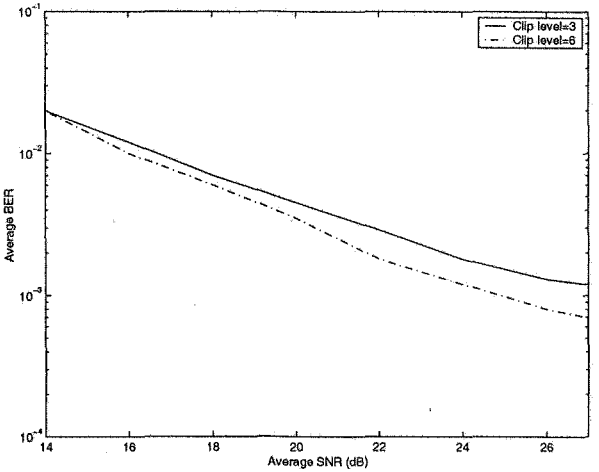


Figure 4.14: BER for uncoded 16-QAM in a multipath fading channel

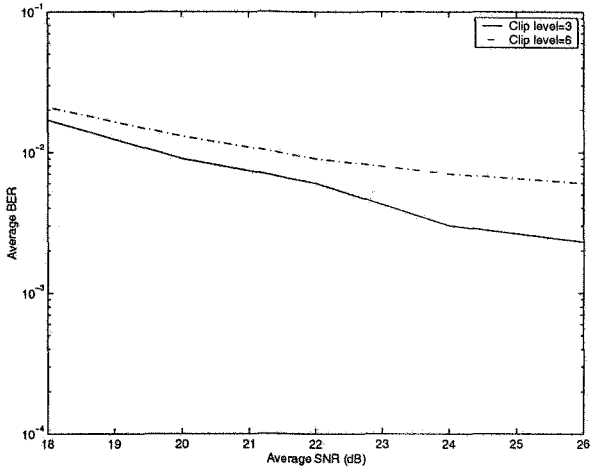


Figure 4.15: BER for uncoded 64-QAM in a multipath fading channel

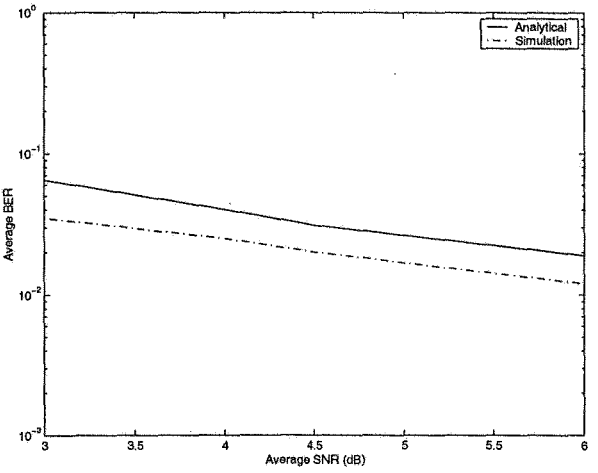


Figure 4.16: Comparison of simulated and analytical symbol error probability floor for 64-tone OFDM signal in Rayleigh fading

nication system, the AGC and clip level are set during the training phase or "preamble" preceding a data packet which is usually designed to have a low peak value; consequently, the signal variations (e.g., caused by fading) in the packet itself are not known *a priori* to the receiver. The choice of an appropriate clip level is, therefore, complicated and involves a tradeoff between excessive clipping noise on the one hand, and a dominant A/D quantization noise floor on the other. Also, even though most analysis, including the one in this paper, deals with memoryless nonlinearities, some devices have memory that arises when various circuit elements are driven into saturation. This memory would only worsen the error performance in the presence of clipping.

4.5 Bandwidth Regrowth

Bandwidth regrowth or out-of-band leakage due to clipping is a critical issue for wireless communications system designers. The amount of out-of-band spill determines filtering requirements and adjacent channel interference. Of course analog filtering is also required at the receiver to reduce aliasing. The spectrum of an instantaneously clipped signal has out-of-band energy at harmonics of the original fundamental band which reduces the energy of in-band signal. Envelope clipping does not result in any harmonics but the energy spill in the adjacent channel is higher.

Using proper filtering of the power amplifier output mitigates both effects. However, proper choice of filtering and its feasibility is important for overall system performance and cost. Interestingly, filtering of harmonics results in higher fluctuation, and therefore, higher peak-to-average ratio in the received signal.

In this analysis, we use two different approaches to evaluate the impact of clipping on OFDM signals. The first technique is based on calculation of the correlation function of the clipped signal at arbitrary level in order to find the average power spectral density of the leakage. Secondly, we outline how to take into account the impulsive nature of the leakage.

Comparing these two techniques one can verify the validity of asymptotic techniques and provide a basis for a system designer as to how much bandwidth spread can be tolerated.

One should notice that the "bursty" nature of clipping should be taken into account as in the previous section. Specifically, spectrum estimation requires an averaging, therefore conditional spectral estimation is a more realistic measure of out-of-band leakage. Otherwise, instantaneous power spill is averaged to a smaller value and may mislead system designers.

Following Rice approximation of the clipping function:

$$f(x) = \begin{cases} x, & if|x| < l \\ \frac{|x|}{x}, & elsewhere \end{cases} \tag{4.49}$$

by a contour integral of the form:

$$f(x) = \frac{1}{2\pi}\int_C \frac{e^{i(x-l)z} - e^{ixz}}{z^2}dz - \int_{C'} \frac{e^{i(x+l)z} - e^{ixz}}{z^2}dz, \tag{4.50}$$

where C is a path of integration along real axis from $-\infty$ to ∞ and indented downwards near origin and C' is the same path but indented upward near origin, the relationship between input correlation function $r(t)$ and output correlation function $R(t)$ is:

$$\frac{1}{\sqrt{N}}R(t) = r\{erf(l/\sqrt{2})\}^2 + \sum_{n=3,5,\cdots}^{\infty} \frac{r^n}{n!}[H_{n-2}(l)e^{-l^2/2}]^2 \tag{4.51}$$

where $H_n(x)$ are Hermite polynomials. Reference [29] uses this approach to arrive at the power spectrum

$$S(f) = erf^2(\frac{1}{\sqrt{2}})S_{xx}(f) + \sum_{n=2,4,6,\cdots}^{\infty} C_n S_{xx}^{(n+1)}(f) \tag{4.52}$$

where $S_{xx}(f)$ is the power spectral density before clipping, $S(f)$ is the spectrum after clipping, $S_{xx}^{(n+1)}(f)$ denotes $(n+1)$-fold convolution, and

$$C_n = \frac{4H_{n-1}^2(l/\sqrt{2})}{\pi 2^n(n+1)!}e^{-l^2}. \tag{4.53}$$

We are interested in evaluating the above in the region $f > f_0$, where the un-clipped spectrum is zero.

In Reference [35], another expression for the power spectrum is given, including high and low frequency asymptotes. In terms of the notation used in this text, the high frequency asymptotic expression is

$$S(f) = \frac{4\sqrt{6}f_0^3 e^{-l^2/2}}{9\pi^2 f^4} \tag{4.54}$$

The out-of-band spectrum discussed above must be multiplied by the transmitter filter function $|H_T(f)|^2$. The analysis applies when the clipping level l is low. It is also applicable in any case for determining the long term average power spectral density, for example to examine whether the signal meets regulatory requirements.

In another approach, as shown in the previous section clipping is considered as a rare event, hence approaches a Poisson process. We use the approach outlined in [35] where the random pulse train is

$$r(t) = \sum_{k=-\infty}^{\infty} P_k(t - t_k - \frac{\tau_k}{2}) \tag{4.55}$$

then the spectrum is the expected value of Fourier Transform with respect to Poisson arrival time t_k and Rayleigh distributed clip duration τ_k:

$$S(f) = \lim_{T\to\infty} \frac{E_{\tau_k,t_k}|\zeta\{r_T\}|^2}{2T} \tag{4.56}$$

Analysis of out-of-band power leakage using the spectrum of clipped signal amounts to integrating Equation 4.56 over the desired frequency band interval. As shown in Figure 4.3 the power amplifier is followed by a filter to control out-of-band leakage. Using the argument presented in the previous section, we can decompose the clipped signal into two components, the original signal and a non-correlated clipped portion as shown in Equation 4.13. Let $H(f)$ denote transfer function of the *linear* filter representing cascade of transmit filter, channel, and receive filter. Then, the spectrum of received signal is

$$S_R(f) = S(f)|H(f)|^2. \tag{4.57}$$

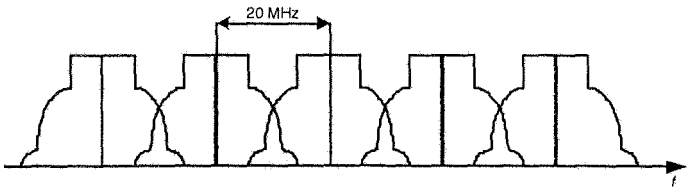


Figure 4.17: Wireless LAN adjacent channels

Using the same argument as the previous section, conditional bit error rate is a more appropriate measure of clipping distortion. To analyze the effect of clipping on the performance of a receiver using an adjacent channel, assume another similar OFDM system occupies the adjacent channel with a guard interval of Δf_g. For example in wireless LAN OFDM the system channel bandwidth is 100 MHz shown in Figure 4.7.

The adjacent channel interference due to clipping is

$$ACI = \int \int_{\Delta\omega} |g_\tau(\omega) H_R(\omega - \Delta\omega)|^2 d\omega f(\tau) d\tau. \tag{4.58}$$

Using the above expression we can calculate conditional error probability $P(e|clip)$ by following the same procedure as previous section.

Following the same reasoning used in the last section to evaluate error probability, treating the above average leakage power to determine the effect on an adjacent channel may be very optimistic at high clipping level. As before, Equation 4.23 describes the effect of a clip on a sub-channel k, only we are now interested in an adjacent system so that $k > N$. That F_k must be multiplied by H_k, the value of the composite filtering function at that sub-carrier frequency.

F_k again is a probabilistic quantity dependent on τ, which in turn is a function of l. Then numerical techniques must be used to calculate

the probability that this quantity exceeds the value which will produce an error for the particular constellation used.

Chapter 5

Synchronization

Multi-carrier modems require a reliable synchronization scheme like any other digital communication modems. Parallel transmission of N symbols results in a longer symbol duration, consequently there is less sensitivity to the timing offset. In other words, unlike single carrier systems in which a timing jitter can create inter-symbol interference (ISI), it does not violate orthogonality of transmitted waveforms in a multi-carrier system. However, frequency offset is detrimental to OFDM systems and has an important role in system design. Phase noise is another critical impairment in wireless OFDM. In this chapter we will review the effect of timing and frequency jitter in OFDM, and discuss some of the synchronization techniques.

5.1 Timing and Frequency Offset in OFDM

Binary bits are grouped into symbols and then they are converted to analog waveforms which are sent across the physical channel. Transmitted signals are associated with timing, frequency, and phase ref-

erence parameters. Proper detection at the receiver requires a knowledge of these parameters. The receiver processes the received waveforms and tries to identify a stream of signal samples. Apparently, first task of the receiver is to decide the symbol boundaries. If the receiver could not clearly identifies the symbols then ISI occurs. Commonly, a sequence of known symbols-preamble are used and receiver seeks for the preamble.

Transmission in the air take place by mixing up the baseband signal with the desired radio frequency spectrum in the transmitting side and the task of the receiver is to bring down the received RF signal from carrier to the baseband. Once the presence of symbol is detected, the next task is to estimate the frequency offset. Frequency offset occurs due to the unmatched local oscillators at the two ends of the communication link, the received signal has an extra factor of $e^{j2\pi\triangle ft}$. Therefore, subcarriers could be shifted from its original position and receiver experiences non-orthogonal signal with the result that inter-carrier interference (ICI) occurs after the FFT since the FFT output will contain interfering energy from all other subcarriers.

Assuming that interference between OFDM blocks is negligible, a transmitted OFDM signal can be represented as

$$x(t) = \sum_{k=-k_1}^{N+k_2+1} \sum_{n=0}^{N-1} D_k e^{j2\pi \frac{nk}{N}} w(t - \frac{k}{f_s}),\ for \frac{-k_1}{f_s} < t < \frac{(N+k_2)}{f_s}, \tag{5.1}$$

where k_1 and k_2 are pre-and postfix lengths, and $w(t)$ is the time domain window function. The received signal $r(t)$ is filtered and sampled at the rate of multiples of $1/T$.

The sampled signal at the output of the receiver FFT with ideal channel can be represented as

$$y_n = \left[\sum_{k=-\infty}^{\infty} X_c(f + \frac{Nk}{T}) \right] \otimes TW(fT) \Big|_{f = \frac{n}{T}} \tag{5.2}$$

where $X_c(f)$ is the Fourier Transform of the periodically repeated analog equivalent of the signal generated by transmitter's IFFT. Therefore, $X_c(f)$ is a line spectrum at $\pm k/T$ and $W(f)$ is the Fourier Transform of window function $w(t)$.

Assuming that the sampling time has a relative phase offset of τ and the offset does not change during one OFDM symbol, the sampled received signal for a non-dispersive channel can be simplified to:

$$y_k = \sum_{n=0}^{N-1} D_n e^{j\varphi} e^{j2\pi \frac{n}{N} f_s t} \Bigg| \; t = \frac{k+\tau}{f_s} \tag{5.3}$$

where ϕ represents envelope delay distortion. After the Fourier Transform at the receiver,

$$\begin{aligned} \tilde{D}_m &= \tfrac{1}{N} \textstyle\sum_{k=0}^{N-1} \sum_{n=0}^{N-1} D_n e^{j2\pi \frac{n}{N} k} e^{j(\varphi + 2\pi \frac{n}{N} \tau)} e^{-j2\pi \frac{m}{N} k} \\ &= \textstyle\sum_{n=0}^{N-1} D_n e^{j(\varphi + 2\pi \frac{n}{N} \tau)} \sum_{k=0}^{N-1} e^{-j2\pi \frac{k}{N}(n-m)} \\ &= \begin{cases} 0, & n \neq m \\ D_m e^{j(\varphi + 2\pi \frac{n}{N} \tau)}, & n = m \end{cases} \end{aligned} \tag{5.4}$$

As shown, the timing phase, or envelope delay distortion, does not violate the orthogonality of the sub-carriers and the effect of the timing phase offset is a phase rotation which linearly changes with sub-carriers' orders. On the other hand, envelope delay results in same amount of rotation for all sub-carriers. In a more general term, it can be shown [45] that with a pulse shaping filter of roll-off α and in the presence of a dispersive channel with impulse response $h(t)$, the detected data is attenuated and phase rotated such that

$$\tilde{D}_m = \gamma_m(\tau) D_m, \tag{5.5}$$

where

$$\gamma_m(\tau) = \begin{cases} H(\frac{m}{NT})e^{j2\pi\frac{m}{NT}\tau}, & 0 \le \frac{m}{N} \le \frac{1-\alpha}{2} \\ H(\frac{m}{NT})e^{j2\pi\frac{m}{NT}\tau} + H(\frac{m-N}{NT})e^{j2\pi\frac{m-N}{NT}\tau}, & \frac{1-\alpha}{2} \le \frac{m}{N} \le \frac{1+\alpha}{2} \\ H(\frac{m-N}{NT})e^{j2\pi\frac{m-N}{NT}\tau}, & \frac{1+\alpha}{2} \le \frac{m}{N} \le 1 \end{cases} \tag{5.6}$$

and $H(m/NT)$ is the Fourier Transform of $h(t)$ at frequency m/NT. Later in this chapter, we discuss estimation of τ and φ.

Frequency offset is a critical factor in OFDM system design. It results in inter-carrier interference and violates the orthogonality of sub-carriers. Equation 5.2, is a mathematical expression of frequency domain samples using a line spectrum representation of the analog signal. Clearly, the effect of frequency offset amounts to inter-channel interference which is similar to inter-symbol interference of a single carrier signal due to a timing jitter. In a non-dispersive channel with rectangular pulse shaping, the interference caused by frequency offset could be too constraining. The sampled signal is

$$\left.\begin{aligned} y_k &= \textstyle\sum_{n=0}^{N-1} D_k e^{j2\pi t(\frac{n}{N}f_s+\delta f)} \\ &= \textstyle\sum_{n=0}^{N-1} D_k e^{j2\pi(\frac{n}{N}+\frac{\delta f}{f_s})k} \end{aligned}\right| \; t = \frac{k}{f_s} \tag{5.7}$$

After DFT we have

$$\tilde{D}_m = D_m\left(\frac{e^{j2\pi\Delta f}-1}{e^{j2\pi\frac{\Delta f}{N}}-1}\right) + \sum_{n=0}^{N-1} D_n \sum_{k=0}^{N-1} e^{j2\pi\frac{k}{N}(n-m+\Delta f)} + N_m, \tag{5.8}$$

where $\Delta f = \frac{n\delta f}{f_s}$ is the relative frequency deviation. As Equation 5.8 shows, In addition to attenuation of desired signal, there is an interference between different symbols of several sub-carriers. This effect creates a significant ICI and is one of the important impairments in OFDM systems, as shown in Figure 5.1. The second term represents the interference from other sub-carriers which is dual to ISI in time domain due to timing offset.

In order to avoid ICI caused by frequency offset, the window function $w_n = w(t)|_{t=nT}$ must be such that zero crossings of its Fourier Transform are at multiples of symbol frequency.

$$W_m = W(\omega)\,|_{\omega=2\pi m f_s} = \delta_m \tag{5.9}$$

A generalized *sinc* function of the form

$$\frac{\sin \omega n}{\omega n} \times g(n) \tag{5.10}$$

for any differentiable function, $g(n)$ satisfies the condition of Equation 5.9. Another desired property of a sinc function is its low rate of change in the vicinity of sampling points in frequency. One common choice is a raised cosine function in time

$$w_{rc}(t) = \begin{cases} T, & 0 \le |t| \le \frac{1-\beta}{2T} \\ \frac{\sin \pi \frac{t}{T}}{\pi \frac{t}{T}} \times \frac{\cos \beta \pi \frac{t}{T}}{1-4\beta^2 \frac{t^2}{T^2}}, & \frac{1-\beta}{2T} \le |t| \le \frac{1+\beta}{2T} \\ 0, & elsewhere \end{cases} \tag{5.11}$$

where β is the roll-off factor for time domain pulse shaping. Notice that a higher roll-off factor requires a longer cyclic extension and guard interval, which consumes higher bandwidth.

Sensitivity of OFDM systems to frequency offset is sometimes too constraining. Therefore, in practice, the zero ICI condition is partly relaxed to provide more robustness against frequency offset. In other words, we tolerate a small amount of ISI at the cost of much less ICI. A Gaussian window is one example in which a single parameter adjustment can control the ISI and ICI trade-off.

The effect of frequency offset on ICI is depicted in Figure 5.1 disregarding ISI degradation. Proper windowing has a significant impact on the sensitivity of the receiver to frequency offset. A reasonable design technique is to optimize between ISI and ICI degradation by choosing an acceptable frequency offset and a proper window.

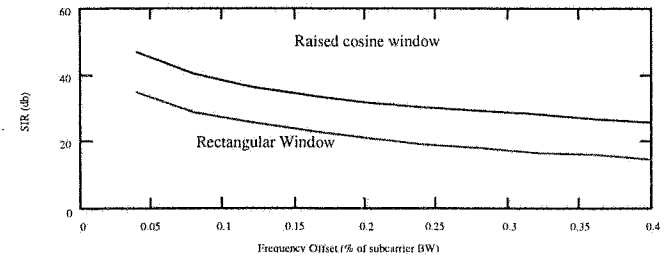


Figure 5.1: ACI caused by frequency offset

5.2 Synchronization & System Architecture

A synchronization sequence in an OFDM system is shown in Figure 5.2. Initially, a coarse frame (or packet) synchronization is required to provide timing information regarding the OFDM symbol. Then, a frequency correction prior to FFT is implemented to reduce the effect of inter-channel interference. Consequently, further stages of fine timing and frequency offset are implemented. These may be combined in a joint estimation block. In general, frame timing and frequency correction are more complicated and require higher hardware/software resources. Timing and frequency correction techniques for OFDM can be classified into two categories: data-aided and non-linear techniques. Data aided techniques use a known bit pattern or pilot signal to estimate the timing or frequency offset. Non-linear techniques utilize the cyclostationarity characteristics of the signal to extract the desired harmonic component by using a non-linear operation. Use of data-aided techniques is applicable to many OFDM digital communication systems such as HDTV and Wireless LAN. Pilot signal fields are provided in existing standards.

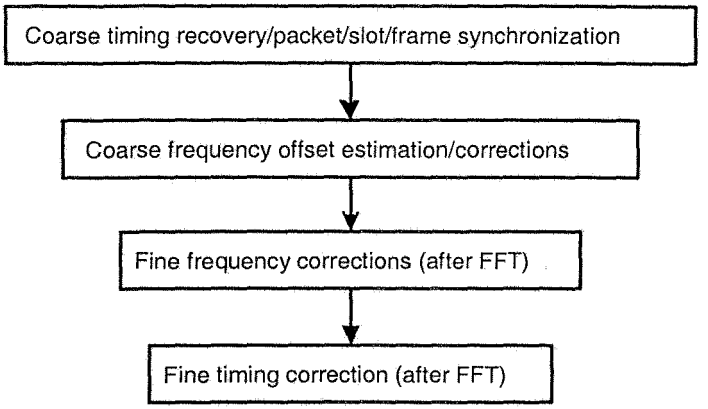


Figure 5.2: Synchronization sequence in OFDM

5.3 Timing and Frame Synchronization

Initially a frame (packet or slot) synchronization circuit is required to detect the frame starting point. This is usually achieved by correlating the incoming signal with a known preamble. The same circuit is sometimes used to adjust for initial gain control. Therefore, the threshold of the detection circuit should be adjusted accordingly. A general block diagram is shown in Figure 5.3. Since timing offset does not violate orthogonality of the symbols it can be compensated after the receiver FFT.

As shown in Equation 5.8, the effect of the timing offset is a phase rotation which linearly increases with sub-carrier order. In order to estimate the timing offset, we should solve a regression problem with proper weighting to represent the effect of fading in the channel.

Frequency offset estimation in packet data communications requires fast acquisition. Hence, phase-lock techniques are not applicable for initial coarse correction. For this purpose, a pilot or training

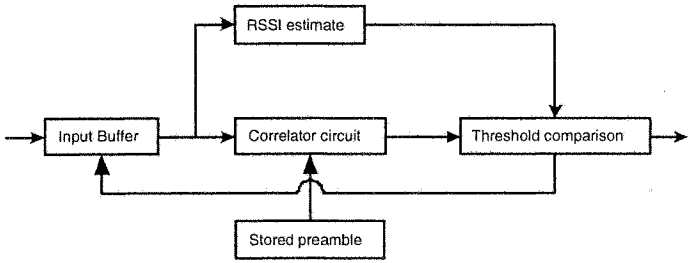


Figure 5.3: Frame synchronization

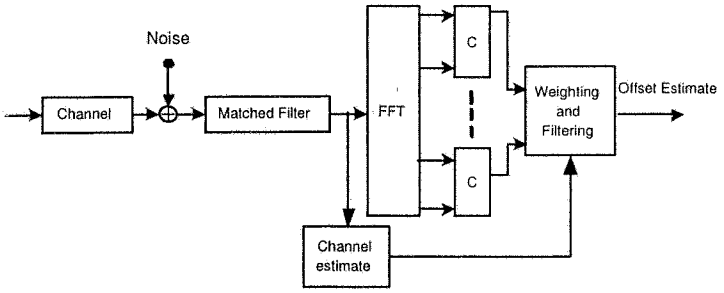


Figure 5.4: Timing offset estimate

sequence should be transmitted. This can be achieved by sending a pilot in one or two bins and estimate the frequency offset by measuring the energy spread in other bins, or by using a training sequence or pilot in every bin and measuring the phase difference by repeating the same pattern. Usually, the second approach is more reliable, as additive noises in different bins are independent and the estimate convergence will be faster. In the following, we discuss both techniques and compare the result analytically and through simulation.

5.4 Frequency Offset Estimation

Frequency offset should be corrected before the receiver FFT. The FFT can be used as a frequency offset detector. Before correction for frequency offset, C/I will not be at the acceptable level for reliable data transmission. Therefore, It is preferable that entire bins should be used for frequency estimation purposes initially. We assume a known data pattern is used for synchronization. After FFT, at the receiver, we have:

$$D_m = \sum_{k=0}^{N-1} \sum_{n=0}^{N-1} D_n e^{j2\pi \frac{k}{N}(n-m+\Delta f)} + N_m. \tag{5.12}$$

If the same block is repeated, the result would be

$$D_m e^{j2\pi \Delta f} + N_m^{'} \tag{5.13}$$

Therefore, we can obtain a proper estimate of Δf by averaging the product of every bin from one symbol and the same bin from the previous symbol, and then averaging the result.

$$\frac{1}{N} \sum_{m=0}^{N-1} \left(\frac{1}{l} \sum_{i=0}^{l-1} y_{mi} \bar{y}^{'}_{m(i+1)} \right) \tag{5.14}$$

This process is shown in Figure 5.5:

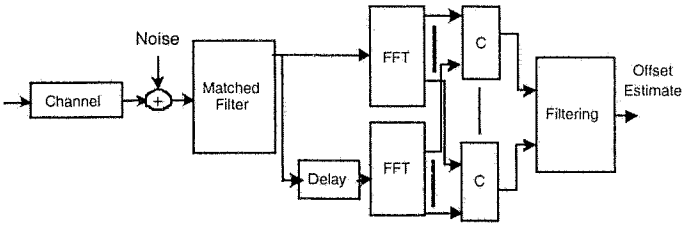


Figure 5.5: Frequency offset estimation

The number of repetitions depends on the signal-to-noise ratio and the required accuracy of frequency estimation. Double averaging accelerates the convergence of the estimate significantly. This scheme works in low signal-to-noise ratios as well. However, this technique is limited to frequency offset of less than $\Delta f < 1/2T$. The initial frequency offset could be higher than that limit, and requires an initial coarse acquisition. A typical coarse frequency detection in single and multi-carrier systems is the Maximum Likelihood estimation technique

$$f_{est} = max_f \sum D_{n+l+1} D^*_{n+l} \tag{5.15}$$

A digital implementation of the technique is shown in the Figure 5.6 where $C(f_i)$ denotes a matched filter tuned to transmit filter impulse response shifted by frequency offset of f_i.

5.5 Phase Noise

Intermediate stages of modulation, such as between RF and IF, can introduce degradation due not only to frequency error as described previously, but also to phase noise in the local oscillators involved in such modulation. Phase noise is the zero mean random process of the deviation of the oscillator's phase from a purely periodic function

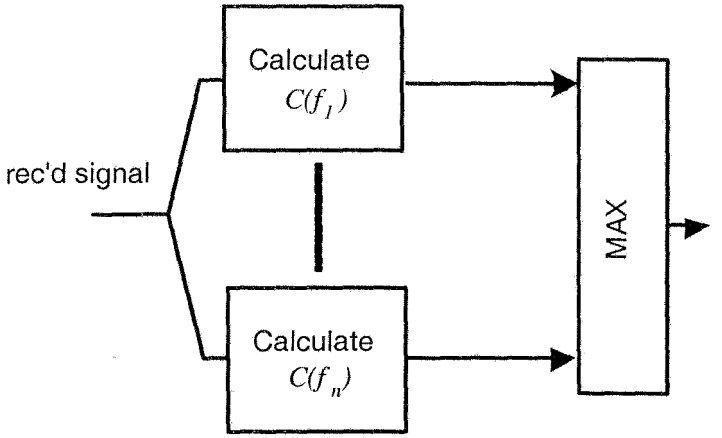


Figure 5.6: A digital implementation of coarse frequency offset estimation using maximum likelihood criteria

at the average frequency. The oscillator output may be considered to be an ideal sine wave phase modulated by the random process. The power spectral density is generally normalized to the power of the sine wave. Higher quality (and more costly) oscillators will typically have lower phase noise than cheaper ones. Therefore, one may expect greater phase noise at the mobile set in a wireless system, as opposed to the base station, and similarly in the customer terminal as opposed to the central office in a subscriber line transmission system.

An approximation of a typical phase noise power spectral density (PSD) is shown in Figure 5.7:

The portion with its slope steeper than f^{-2} exists only at a very low frequency. Actually, the PSD may be mathematically inaccurate in that the phase process in this region is typically non-stationary. The f^{-2} slope is equivalent to white frequency noise. The f^{-1} slope in PSD is difficult to model, but agrees with measurements and is commonly referred to as "flicker noise". Finally, the white phase noise at

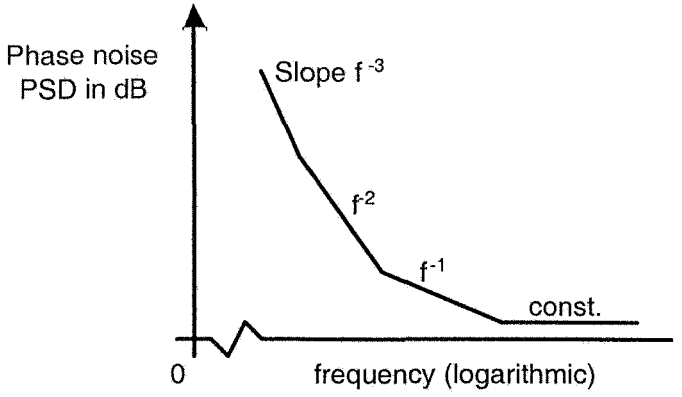


Figure 5.7: A typical oscillator phase noise power spectrum

a higher deviation frequency is usually quite low.

The above describes a free-running oscillator. In practice, any oscillator used for intermediate modulation should be phase-locked to eliminate frequency error, for the reasons previously discussed. When a phase-locked loop is employed, the phase noise spectrum is modified. Figure 5.8 shows a classic phase-locked loop and its mathematical model.

The variables ϕ are phase deviations from nominal. ϕ_0,ϕ_i, and ϕ_n are the input, output, and internal oscillator phase deviations respectively. The low-pass filter $F(s)$ must be chosen such that the feedback loop is stable. The voltage-controlled oscillator (VCO) has an output frequency deviation from its nominal value that is proportional to the control voltage. Since phase is the integral of frequency, the transfer function of the VCO is K/s as shown. The equation of the loop is

$$\phi_0 = \frac{KF(s)}{s + KF(s)}\phi_i + \frac{s}{s + KF(s)}\phi_n \tag{5.16}$$

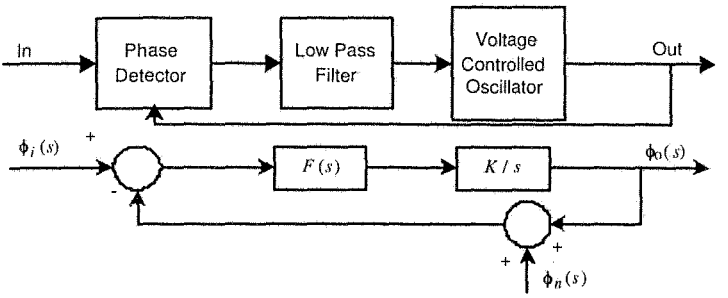


Figure 5.8: Classical phase-locked loop

Because the open loop transfer function has a pole at zero frequency, the frequency of the VCO will lock exactly to the average input frequency, provided the initial frequency difference is within the lock-in range of the loop. Substituting the frequency variable $j\omega$ for the Laplace variable s, the power transfer function is

$$|\phi_0|^2 = \frac{K^2|F(\omega)|^2}{|j\omega + KF(\omega)|^2}|\phi_i|^2 + \frac{\omega^2}{|j\omega + KF(\omega)|^2}|\phi_n|^2 \qquad (5.17)$$

The phase-locked loop thus acts as a low-pass filter of its input phase noise, and a high-pass filter of the internal VCO noise. Note that this is true even without a low-pass filter $[F(s) = 1]$. More modern phase-locked loops are implemented digitally [45], except at very high frequencies. An example of a digital phase-locked loop is shown in Figure 5.9.

In this implementation, the input and output are in complex form. The complex multiplier eliminates the need to remove double frequency components, an easy task in analog implementation, but not in digital. Only the imaginary output of the multiplier is used. Instead of a VCO, the same function is performed by an accumulator and phase look-up table. The loop is implemented as a sampled data system. The sampling and computation rate $1/\tau$ need not be particularly accurate or related to the frequency to be locked. As long as τ is much smaller

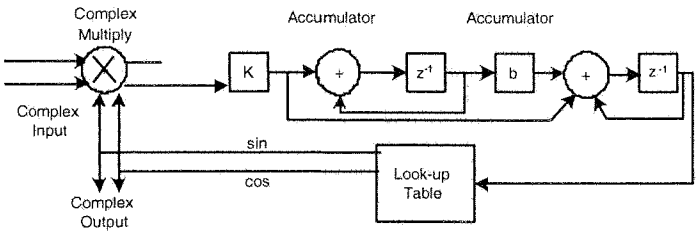


Figure 5.9: A second order digital phase-locked loop

than any of the loop time constants, the operation is equivalent to an analog loop. The particular loop shown is a second order one, where the additional accumulator in the digital filter produces a second pole at zero frequency.

When a phase-locked loop is employed to lock the local oscillator used for intermediate modulation, the resultant phase noise spectrum will be a high-pass filtered version of Figure 5.7, such as the simplified spectrum shown in Figure 5.10.

If the constant high frequency portion can be neglected, then the above (two-sided) PSD is a Lorenzian one,

$$|\phi_0(f)|^2 = \frac{1}{\pi} \frac{K\beta}{f^2 + \beta^2} \tag{5.18}$$

The 3 dB bandwidth is β, and the total power is K. The corresponding phase noise is a Weiner process, in which all disjoint phase increments are independent. This noise model is frequently used in analysis of phase noise impairment.

We are now ready to examine the effects of phase noise on an OFDM system [47, 48, 49, 50]. Following the same analysis that was used to examine frequency offset, and assuming there are no other

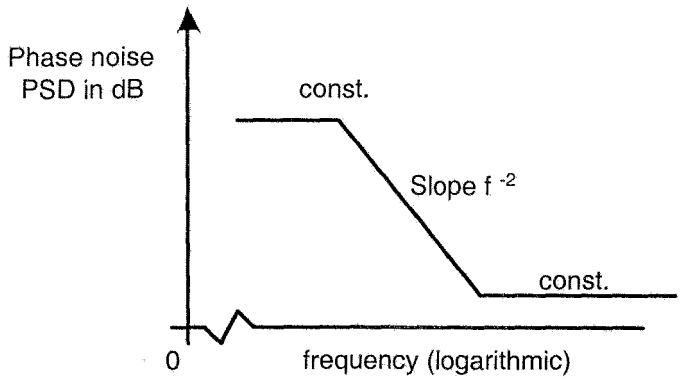


Figure 5.10: Typical spectrum of a phase-locked oscillator

impairments, the output of the receiver demodulator is

$$\tilde{D}_m = \frac{1}{N}\sum_{n=0}^{N-1} D_n \sum_{k=0}^{N-1} e^{j\phi(k)} e^{-j2\pi\frac{k}{N}(n-m)},\ m = 0, \cdots, (N-1) \tag{5.19}$$

where $\phi(k)$ is the sample of the phase noise at the receiver input at the kth sample time. If the phase noise is small, as is almost always the case,

$$e^{j\phi(k)} \approx 1 + j\varphi(k), \tag{5.20}$$

so,

$$\tilde{D}_m \approx D_m - \frac{jD_m}{N}\sum_{k=0}^{N-1}\varphi(k) - \frac{j}{N}\sum_{n=0,n\neq m}^{N-1} D_n \sum_{k=0}^{N-1}\varphi(k)e^{-j2\pi\frac{k}{N}(n-m)} \tag{5.21}$$

The first term in Equation 5.21 is the desired output, the second is a random phase rotation of the sub-carrier, and the third term is inter-carrier interference due to a loss of orthogonality among the sub-carriers.

Let us first examine the second term. When the bandwidth of the phase noise is less than the OFDM symbol rate $1/T$, then

$$\sum_{k=0}^{N-1} \varphi(k) = \frac{N}{T} \int_T \varphi(t) dt \tag{5.22}$$

or in the frequency domain,

$$\sum_{k=0}^{N-1} \varphi(k) = N \int_{-\infty}^{\infty} \phi(f) sinc(fT) df \tag{5.23}$$

This phase rotation is the same for all sub-carriers. It can readily be eliminated, for example by measuring the phase variation of a pilot sub-carrier and subtracting that rotation from all sub-carriers. Another technique to solve this, and other problems, is the use of differential phase modulation among the sub-carriers. In this technique, one sub-carrier carries a fixed phase and the information on other sub-carriers is carried by phase differences. This is applicable only to constellations of constant amplitude. Furthermore, there is a noise performance penalty when compared with coherent demodulation.

If the above correction is performed, we are left with the inter-carrier interference. Again, assuming the phase noise bandlimited to less than $1/T$, that interference is

$$\begin{aligned} I &= \tfrac{j}{T} \textstyle\sum_{n=0,n\neq m}^{N-1} D_n \int_T \varphi(t) e^{j2\pi(n-m)\frac{t}{T}} dt \\ &= j \textstyle\sum_{n=0,n\neq m}^{N-1} D_n \int_{-\infty}^{\infty} \phi(f) sinc(n-m+fT) df \end{aligned} \tag{5.24}$$

If the data symbols on the different sub-carriers are independent, then the interference may be treated as a Gaussian noise of power

$$I^2 = \sum_{n=0,n\neq m}^{N-1} \sigma_n^2 \int_{-\infty}^{\infty} |\phi(f)|^2 sinc^2(n-m+fT) df \tag{5.25}$$

where σ_n^2 is the average power of the nth sub-carrier. For any phase noise PSD that is decreasing with frequency, near sub-carriers will

contribute greater to this interference than sub-carriers further removed from the sub-carrier of interest. For this reason, sub-carriers near the center of the frequency band will be subject to more interference than sub-carriers at the band edge by up to a factor of two. More importantly, the interference increases as the spacing between sub-carriers decreases.

The noise-to-signal ratio for a sub-carrier near the middle of the band, assuming equal power on all sub-carriers, is

$$\frac{I^2}{\sigma_m^2} = 2 \sum_{n=1}^{N/2-1} \int_{\infty}^{\infty} |\phi(f)|^2 sinc^2(p + fT) df \qquad (5.26)$$

For the Lorenzian noise spectrum given previously, this quantity is plotted as a function of βT in Figure 5.11, normalized by the ratio of phase noise power to power in the sub-carrier. Here, all sub-carrier powers were assumed to be equal. Note that βT is the ratio of the 3 dB bandwidth of the phase noise to the OFDM symbol rate. To find the un-normalized noise-to-signal ratio, the plotted quantity should be multiplied by the phase noise to sub-carrier power ratio.

It is seen that as long as the OFDM symbol rate is greater than approximately twice the phase noise bandwidth, the performance degrades as that symbol rate, or, equivalently, the sub-carrier spacing, decreases. Therefore, under these conditions, the use of a large number of sub-carriers to carry a given bit rate is detrimental. However, when the OFDM symbol rate increases beyond approximately twice the phase noise bandwidth, there is little further degradation., assuming that the line rate is still much greater than the noise bandwidth.

Another form of phase noise that can exist even in OFDM systems without intermediate modulations, such as in wireline applications, results from jitter in the sampling time at the receiver input. In Reference [51], it is shown that when the timing jitter is constant over an OFDM symbol interval, no inter-carrier interference results. In this case, each sub-carrier constellation is subjected to rotation pro-

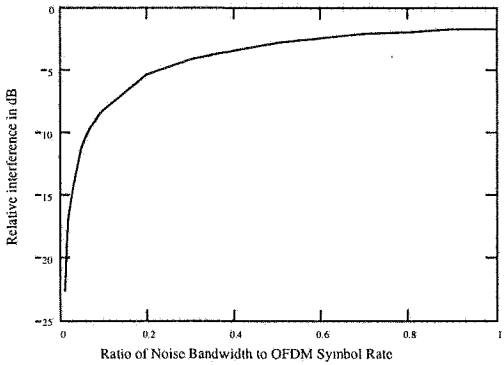


Figure 5.11: Inter-carrier interference noise

portional to the timing offset and to the frequency of that sub-carrier.

Chapter 6

Channel Estimation and Equalization

6.1 Introduction

Time-variant and frequency-selective fading channels present a severe challenge to the designer of a wireless communication system. An OFDM receiver implements dual role to tackle this problem:channel estimation or equalization. Several choices are possible for the implementation of a receiver depending on the modelling of the channel and complexity invested in each task. The nature of the OFDM enables powerful estimation and equalization techniques.

6.2 Channel Estimation

In OFDM link, modulated bits are disturbed during the transmission through the channel since the channel introduces amplitude and phase shifts due to frequency selective and time-varying nature of the radio

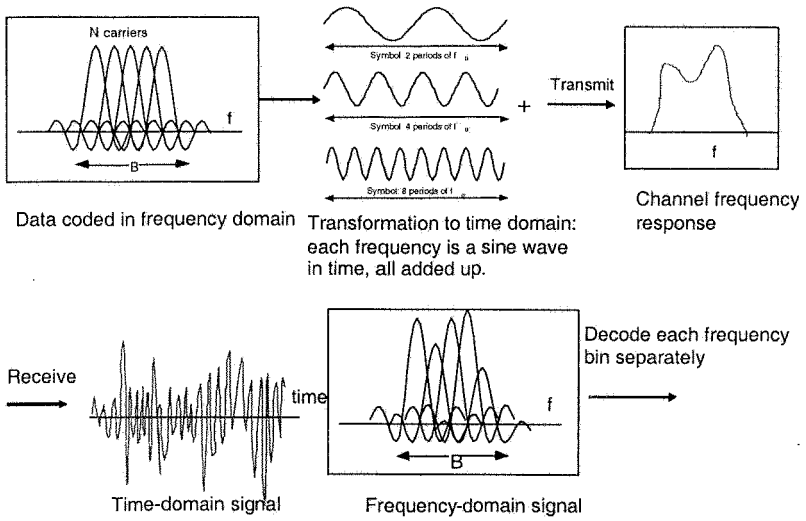


Figure 6.1: Observation of the distortion by channel [87]

channel. In order for the receiver to acquire the original bits, it needs to take into account these unknown changes. Figure 6.1 illustrates a typical OFDM transmission and the effects of the channel. The receiver applies either coherent detection or non-coherent detection to recover the original bits. Coherent detection uses reference values that are transmitted along with data bits. The receiver can tune into those values and estimate the channel only at the locations where the reference values are located. The entire channel can be estimated by using several interpolation techniques. Non-coherent detection on the other hand, does not use any reference values but uses differential modulation where the information is transmitted in difference of the two successive symbols. The receiver uses two adjacent symbols in time or two adjacent subcarriers in frequency to compare one with another to acquire the transmitted symbol [52].

In this section, we will review channel estimation techniques. We

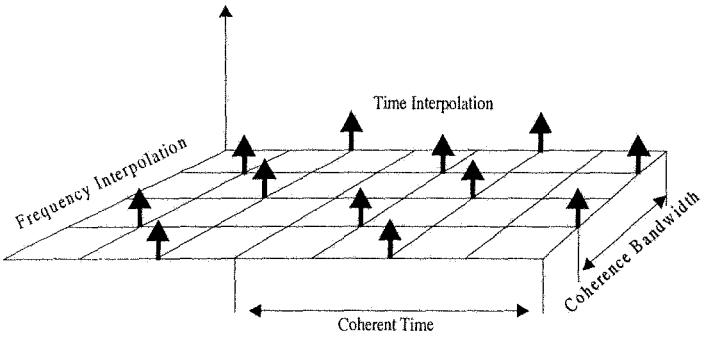


Figure 6.2: Pilot positioning in time and frequency

start with coherent and non-coherent detection wherein we describe the block type and comb type channel estimation and continue with the interpolation techniques. Finally, channel estimation in MIMO-OFDM system is discussed.

6.2.1 Coherent Detection

The coherent detection can be performed by either inserting pilot tones into all of the subcarriers of OFDM symbols with a period in time or inserting pilot tones into each OFDM symbol with a period in frequency. In the case of time-varying channels, the pilot signal should be repeated frequently. The spacing between pilot signals in time and frequency depends on coherence time and bandwidth of the channel. One can reduce the pilot signal overhead by using a pilot signal with a maximum distance of less than the coherence time and coherence bandwidths. Then, by using time and frequency interpolation, the impulse response and frequency response of the channel can be calculated. Figure 6.2 illustrates the pilot arrangement in time-frequency plane.

If B_d is Doppler spread and σ_τ is delay spread, a suitable choice

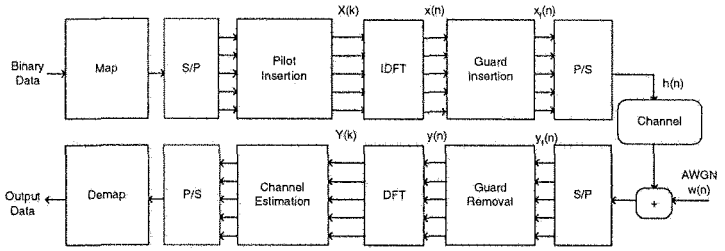


Figure 6.3: Baseband OFDM system

for pilot spacing in time N_p^t, and in frequency N_p^f, is as follows:

$$N_p^t \approx \frac{1}{B_d T_{symb}} \quad N_p^f \approx \frac{1}{\triangle f \sigma_\tau} \tag{6.1}$$

where $\triangle f$ is subcarrier bandwidth, T_{symb} is symbol time.

Various types exist in the pilot arrangement depending on the channel condition. Those types can be classified into two at higher level. The first one is the block type pilot arrangement for slow fading channels. All subcarriers are used once and the estimated channel is used for the coming symbols since the channel changes very slowly for slow fading channels, and the channel is highly correlated for the consecutive symbols. The second approach is the comb type pilot arrangement, which interleaves the pilots over time and frequency. Comb-type estimation uses several interpolation techniques to estimate the entire channel. Flat fading channels can be estimated by comb-type pilot arrangement and pilot frequency increases if the channel is frequency selective fading.

The OFDM system based on pilot channel estimation is given in Figure 6.3. The binary information is first grouped and mapped according to the modulation in "signal mapper". After inserting pilots either to all sub-carriers with a specific period or uniformly between the information data sequence, IDFT block is used to transform the data sequence of length N $\{X(k)\}$ into time domain signal $\{x(n)\}$

with the following equation:

$$\begin{aligned} x(n) &= IDFT\{X(k)\} \qquad n = 0, 1, 2, \ldots, N-1 \\ &= \sum_{k=0}^{N-1} X(k) e^{j\frac{2\pi kn}{N}} \end{aligned} \tag{6.2}$$

where N is the DFT length. Following IDFT block, guard time, which is chosen to be larger than the expected delay spread, is inserted to prevent inter-symbol interference. This guard time includes the cyclically extended part of OFDM symbol in order to eliminate inter-carrier interference (ICI). The resultant OFDM symbol is given as follows:

$$x_f(n) = \begin{cases} x(N+n), & n = -N_g, -N_g+1, \ldots, -1 \\ x(n), & n = 0, 1, \ldots, N-1 \end{cases} \tag{6.3}$$

where N_g is the length of the guard interval. The transmitted signal $x_f(n)$ will pass through the frequency selective time varying fading channel with additive noise. The received signal is given by:

$$y_f(n) = x_f(n) \bigotimes h(n) + w(n) \tag{6.4}$$

where $w(n)$ is Additive White Gaussian Noise (AWGN) and $h(n)$ is the channel impulse response.The channel response h can be represented by [56]:

$$h(n) = \sum_{i=0}^{r-1} h_i e^{j\frac{2\pi}{N} f_{Di} T n} \delta(\sigma_\tau - \tau_i)\, 0 \leq n \leq N-1 \tag{6.5}$$

where r is the total number of propagation paths, h_i is the complex impulse response of the ith path, f_{Di} is the ith path Doppler frequency shift, σ_τ is delay spread index, T is the sample period and τ_i is the ith path delay normalized by the sampling time. At the receiver, after passing to discrete domain through A/D and low pass filter, guard time is removed:

$$\begin{aligned} &y_f(n) && -N_g \leq n \leq N-1 \\ &y(n) = y_f(n+N_g) && n = 0, 1, \ldots N-1 \end{aligned} \tag{6.6}$$

Then $y(n)$ is sent to DFT block for the following operation:

$$\begin{aligned} Y(k) &= DFT\{y(n)\} \qquad k = 0, 1, 2 \ldots N-1 \\ &= \tfrac{1}{N} \textstyle\sum_{n=0}^{N-1} y(n) e^{-j\frac{2\pi kn}{N}} \end{aligned} \tag{6.7}$$

Assuming there is no ISI, [59] shows the relation of the resulting $Y(k)$ to $H(k) = DFT\{h(n)\}$, $I(k)$ (ICI because of Doppler frequency), and $W(k) = DFT\{w(n)\}$, with the following equation [55]:

$$Y(k) = X(k)H(k) + I(k) + W(k) \; k = 0, 1 \ldots N-1 \tag{6.8}$$

where

$$H(k) = \sum_{i=0}^{r-1} h_i e^{j\pi f_{D_i} T} \frac{sin(\pi f_{D_i} T)}{\pi f_{D_i} T} e^{-j\frac{2\pi\tau_i}{N}k},$$

$$I(k) = \sum_{i=0}^{r-1} \sum_{\substack{K=0, \\ K \neq k}}^{N-1} \frac{h_i X(K)}{N} \frac{1 - e^{j2\pi(f_{D_i}T - k + K)}}{1 - e^{j\frac{2\pi}{N}(f_{D_i}T - k + K)}} e^{-j\frac{2\pi\tau_i}{N}K},$$

Following DFT block, the pilot signals are extracted and the estimated channel $H_e(k)$ for the data sub-channels is obtained in channel estimation block. Then the transmitted data is estimated by:

$$X_e = \frac{Y(k)}{H_e(k)} \qquad k = 0, 1, \ldots N-1 \tag{6.9}$$

Then the binary information data is obtained back in "signal demapper" block.

6.2.2 Block-Type Pilot Arrangement

Block type pilot channel estimation has been developed under the assumption of slow fading channels. The estimation of the channel for this block-type pilot arrangement can be based on Least Square (LS) or Minimum Mean-Square (MMSE). The MMSE estimate has been shown to give 10-15 dB gain in signal-to-noise ratio (SNR) for the

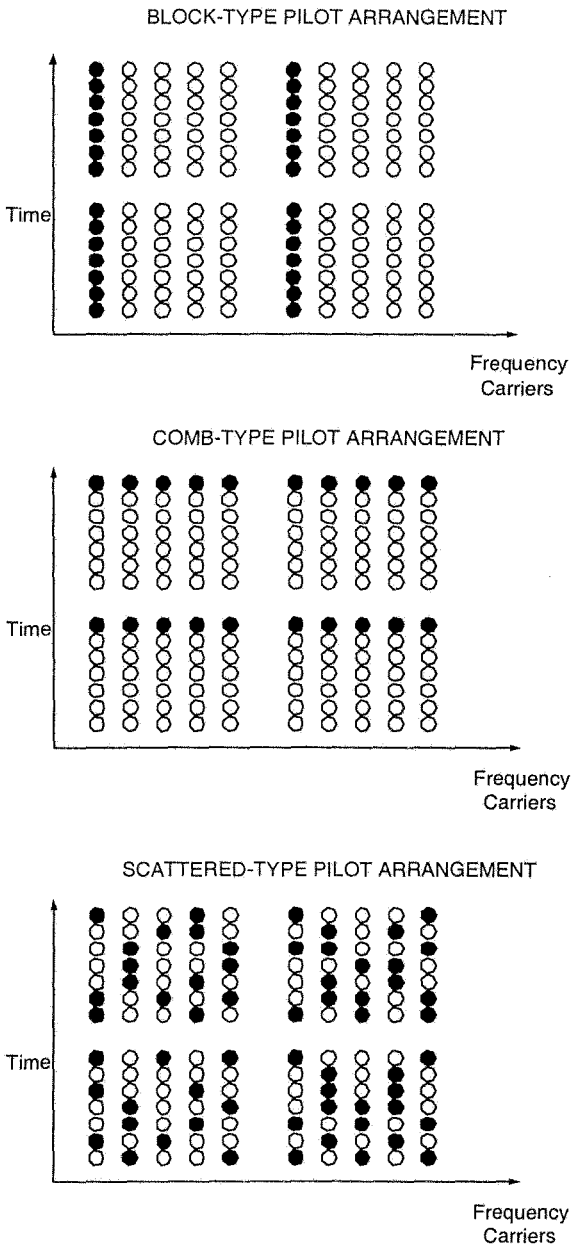


Figure 6.4: Pilot arrangement

same mean square error of channel estimation over LS estimate [53]. In [54], a low-rank approximation is applied to linear MMSE by using the frequency correlation of the channel to eliminate the major drawback of MMSE, which is complexity.

In block-type pilot based channel estimation, OFDM channel estimation symbols are transmitted periodically, in which all sub-carriers are used as pilots as seen in Figure 6.4. If the channel is constant during the block, there will be no channel estimation error since the pilots are sent at all carriers.

If inter symbol interference is eliminated by the guard interval, the Equation (6.8) is written in matrix notation as

$$Y = XFh + W \tag{6.10}$$

where

$$\begin{aligned} X &= diag\{X(0), X(1), \ldots, X(N-1)\} \\ Y &= [Y(0)Y(1)\ldots Y(N-1)]^T \\ W &= [W(0)W(1)\ldots W(N-1)]^T \\ H &= [H(0)H(1)\ldots H(N-1)]^T = DFT_N\{h\} \end{aligned}$$

$$F = \begin{bmatrix} W_N^{00} & \cdots & W_N^{0(N-1)} \\ \vdots & \ddots & \vdots \\ W_N^{(N-1)0} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix} \tag{6.11}$$

$$W_N^{nk} = \frac{1}{N} e^{-j2\pi \frac{n}{N} k}$$

If the time domain channel vector h is Gaussian and uncorrelated with the channel noise W, the frequency domain MMSE estimate of h is given by [54]:

$$H_{MMSE} = FR_{hY}R_{YY}^{-1}Y \tag{6.12}$$

where

$$R_{hY} = E\{hY\} = R_{hh}F^H X^H$$

$$R_{YY} = E\{YY\} = XFR_{hh}F^H X^H + \sigma^2 I_N \tag{6.13}$$

are the cross covariance matrix between h and Y and the auto covariance matrix of Y. R_{hh} is the auto-covariance matrix of h and σ^2 represents the noise variance $E\{|W(k)|^2\}$.

The LS estimate on the other hand minimizes the difference between the received value and estimated value. Thus, LS minimizes $(Y - XFh)^H(Y - XFh)$.

The estimated channel in LS is represented by

$$H_{LS} = X^{-1}Y, \tag{6.14}$$

Decision-feedback channel estimation

When the channel is slow fading, the channel estimation inside the block can be updated using the decision feedback equalizer at each sub-carrier. Decision feedback equalizer for the kth sub-carrier can be described as follows:

- The channel response at the kth sub-carrier estimated from the previous symbol $\{H_e(k)\}$ is used to find the estimated transmitted signal $\{X_e(k)\}$.

$$X_e(k) = \frac{Y(k)}{H_e(k)} \qquad k = 0, 1, \ldots N-1 \tag{6.15}$$

- $\{X_e(k)\}$ is mapped to the binary data through "signal demapper" and then obtained back as $\{X_{pure}(k)\}$ through "signal mapper".

- The estimated channel $H_e(k)$ is updated by:

$$H_e(k) = \frac{Y(k)}{X_{pure}(k)} \qquad k = 0, 1, \ldots N-1 \tag{6.16}$$

Since the decision feedback equalizer has to assume that the decisions are correct, the fast fading channel may cause the complete loss of estimated channel parameters. Therefore, as the channel fading becomes faster, there happens to be a compromise between the estimation error due to the interpolation and the error due to loss of channel tracking. For fast fading channels the comb-type based channel estimation performs better.

6.2.3 Comb-Type Pilot Arrangement

The comb-type pilot channel estimation, has been introduced to satisfy the need for equalizing when the channel changes even in one OFDM block. The comb-type pilot channel estimation consists of algorithms that estimate the channel at pilot frequencies and that interpolate the channel.

The estimation of the channel at the pilot frequencies for comb-type based channel estimation can be based on LS, MMSE or Least Mean-Square (LMS). MMSE has been shown to perform much better than LS. In [55], the complexity of MMSE is reduced by deriving an optimal low-rank estimator with singular-value decomposition.

The interpolation of the channel for comb-type based channel estimation can depend on linear interpolation, second order interpolation, low-pass interpolation, spline cubic interpolation, and time domain interpolation. In [55], second-order interpolation has been shown to perform better than the linear interpolation. In [56], time-domain interpolation has been proven to give a lower bit-error rate (BER) compared to linear interpolation.

In comb-type pilot based channel estimation, the N_p pilot signals are uniformly inserted into $X(k)$ according to the following equation:

$$\begin{aligned} X(k) &= X(mL + l) \\ &= \begin{cases} x_p(m), & l = 0 \\ inf.data & l = 1, \ldots L - 1 \end{cases} \end{aligned} \tag{6.17}$$

where $L = (\text{number of carriers}/N_p)$ and $x_p(m)$ is the mth pilot carrier value. We define $\{Hp(k), \quad k = 0, 1, \ldots Np\}$ as the frequency response of the channel at pilot sub-carriers. The estimate of the channel at pilot sub-carriers based on LS estimation is given by:

$$H_e = \frac{Y_p}{X_p}, \qquad k = 0, 1, ...N_p - 1 \tag{6.18}$$

where $Y_p(k)$ and $X_p(k)$ are output and input at the kth pilot sub-carrier respectively.

Since LS estimate is susceptible to noise and ICI, MMSE is proposed while compromising complexity. Since MMSE includes the matrix inversion at each iteration, the simplified linear MMSE estimator is suggested in [57]. In this simplified version, the inverse only needs to be calculated once. In [55], the complexity is further reduced with a low-rank approximation by using singular value decomposition.

6.2.4 Interpolation Techniques

In comb-type pilot based channel estimation, an efficient interpolation technique is necessary in order to estimate channel in time and frequency. The pilot symbols can be interleaved within a symbol or between the symbols. The estimated pilot values should be connected in order to get an estimated value for the data values. The following techniques are written for interpolation within the symbol but it can also apply to interpolation between the symbols.

Linear Interpolation:

The linear interpolation method is shown to perform better than the piecewise-constant interpolation in [58]. The channel estimation at the data-carrier k, $mL < k < (m + 1)L$, using linear interpolation is

The pilot symbols do not need to be periodic, they can be arranged according to the channel condition dynamically with additional overhead. the number of pilots and their positions can be changed dynamically. This type estimation is called scattered estimation as shown in Figure 6.4.

6.2.5 Non-coherent Detection

Non-coherent detection uses differential modulation technique wherein the information to be transmitted is contained in the phase difference ψ_k of two successive modulation symbols x_k and x_{k-1} as in M-DPSK and DQPSK. and the received phase differences φ_k of the corresponding received symbols y_k and y_{k-1}[65].

$$\psi_k = \arg(\frac{x_k}{x_{k-1}}), \quad \varphi_k = \arg(\frac{y_k}{y_{k-1}}) \tag{6.24}$$

In non-coherent detection, there is no requirement for the channel estimation and consequently, there is no overhead introduced by the pilot symbols. Coherent detection, on the other hand, performs better than non-coherent detection since coherent detection is not affected by phase changes since training values are used to estimate the reference phases and amplitudes[66].

6.2.6 Performance

The performance results are from a system with the parameters indicated in Table 6.2.6. Perfect synchronization and no ISI is assumed. The HE channel model of the ATTC (Advanced Television Technology Center) and the Grande Alliance DTV laboratory's ensemble E model is used. The impulse response of the Rayleigh channel is given by:

Parameters	Specifications
FFT Size	1024
Number of Active Carriers(N)	128
Pilot Ratio	1/8
Guard Interval	256
Guard Type	Cyclic Extension
Sample Rate	44.1KHz
Bandwidth	17.5KHz
Signal Constellation	16QAM
Channel Model	Rayleigh Fading
Coding Rate	1/1

Table 6.1: Simulation parameters for channel estimation

$$\begin{aligned} h(n) &= \alpha(n) + 0.3162\alpha(n-2) \\ &\quad + 0.1995\alpha(n-17) + 0.1296\alpha(n-36) \\ &\quad + 0.1\alpha(n-75) + 0.1\alpha(n-137) \end{aligned} \tag{6.25}$$

In block type estimation, each block consists of a fixed number of symbols, which is 30 in the simulation. Pilots are sent in all the sub-carriers of the first symbol of each block and channel estimation is performed by using LS estimation. The channel is estimated at the beginning of the block and used for all the following symbols of the block.

In decision feedback equalizer, in addition to the block type estimation, the successive symbols after the pilots are used to calibrate the estimated values of the pilot symbols. In this way, channel can be tracked through the channel variations.

Values is estimated in either LS or Least Mean Square (LMS) method. The LMS estimator uses one tap LMS adaptive filter at each

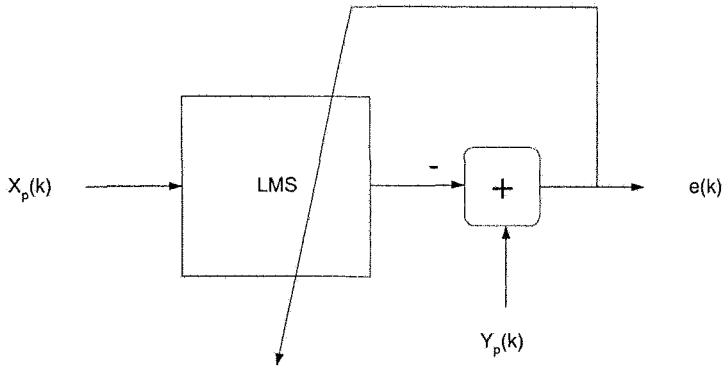


Figure 6.6: LMS scheme

pilot frequency. The first value is found directly through LS and the following values are calculated based on the previous estimation and the current channel output as shown in Figure 6.6.

All of the possible interpolation techniques (linear interpolation, second order interpolation, low-pass interpolation, spline cubic interpolation, and time domain interpolation) are applied to LS estimation result to investigate the interpolation effects and linear interpolation are applied to LMS estimation results to compare with the LS overall estimation results.

The legends "linear, second-order, low-pass, spline, time domain" denote interpolation schemes of comb-type channel estimation with the LS estimate at the pilot frequencies, "block type" shows the block type pilot arrangement with LS estimate at the pilot frequencies, "decision feedback" means the block type pilot arrangement with LS estimate at the pilot frequencies and with decision feedback, and "LMS" is for the linear interpolation scheme for comb-type channel estimation with LMS estimate at the pilot frequencies.

Figures 6.7 gives the BER performance for Rayleigh fading channel, with static channel response given in equation (6.25).

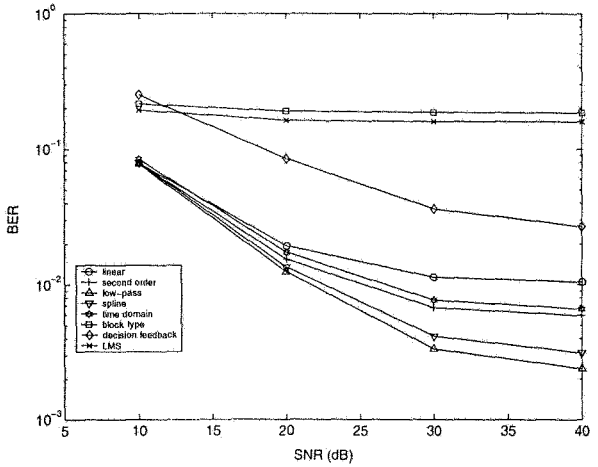


Figure 6.7: 16QAM modulation with Rayleigh fading (Doppler Freq. 70Hz)

These results show that the block-type and decision feedback estimation BER values are 10-15dB higher than that of the comb-type estimation type. This is because the channel transfer function changes so fast that there are even changes for adjacent OFDM symbols.

The comb-type channel estimation with low pass interpolation achieves the best performance among all the estimation techniques for 16QAM modulation. The performance among the comb-type channel estimation techniques usually range from the best to the worst as follows: low-pass, spline, time-domain, second-order and linear. The result was expected since the low-pass interpolation used in simulation does the interpolation such that the mean-square error rate between the interpolated points and their ideal values is minimized.

Figure 6.8 shows the performance of channel estimation methods for 16QAM modulation in Rayleigh fading channel for different Doppler frequencies. The general behavior of the plots is that BER increases as the Doppler spread increases. The reason is because of

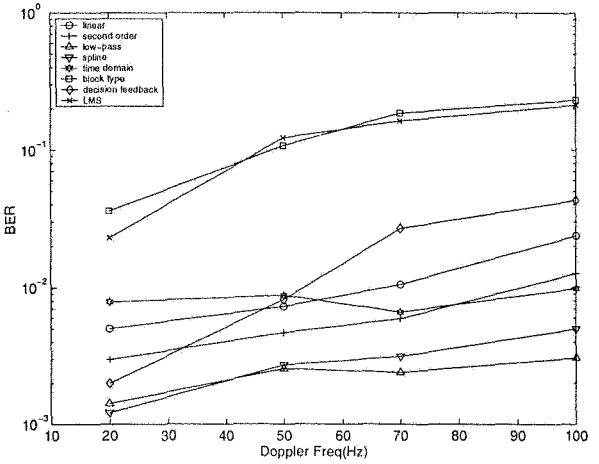


Figure 6.8: 16QAM modulation with Rayleigh fading (SNR 40dB)

the existence of severe ICI caused by Doppler shifts. Another observation from this plot is that the decision feedback block type channel estimation performs better than the comb-type based channel estimation for low Doppler frequencies except low-pass and spline interpolation. One can observe that time-domain interpolation performance improves compared to other interpolation techniques as Doppler frequency increases [52].

6.2.7 Channel Estimation for MIMO-OFDM

Multiple transmit and receive antennas are a strong choice to improve capacity in OFDM systems. MIMO systems leverage the transmitter and receiver diversity. Channel estimation is an important issue for the MIMO-OFDM systems since computational complexity increases rapidly with the increase in the number of antennas. Any estimation error becomes crucial since the system becomes more sensitive.

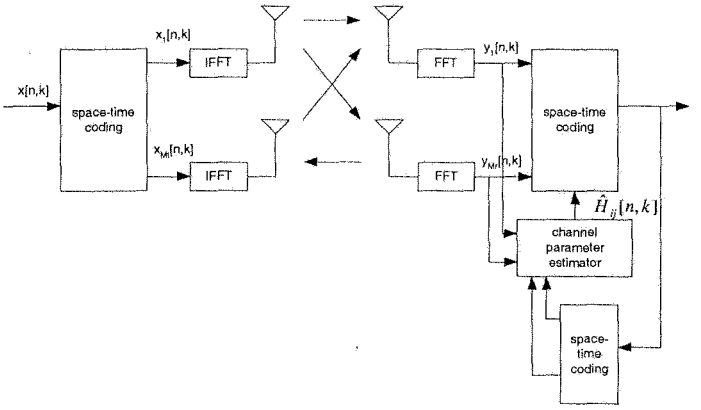


Figure 6.9: MIMO-OFDM architecture

In the estimation process of the channel in any transmitter receiver pair, the signals transmitted by other transmitters act as an interference. The challenge can be overcome by inserting pilot tones orthogonal in time, meaning whenever a pilot tone is inserted in a subcarrier, all other transmitters do not send anything in that subcarrier [73]. This method is suitable to the cases when the symbol time is larger than the coherence time. If the symbol time is smaller than the coherence time, then the correlation can be leveraged to improve the channel estimation. Assume $H_n(t)$ is the estimate of the channel at time t at the nth pilot tone. The $H_n(t+1)$ can be estimated given $H_n(t)$ and $H_n(t+1)$ by MMSE or LS estimation [62].

For a MIMO-OFDM system with M_t transmit and M_r receive antennas, there are $M_t \times M_r$ channels to be estimated. All transmitters send their signals simultaneously and the received signal is a superposition of the transmitted signals that are distorted by the channel. The received signal at the kth subcarrier of the nth block from the jth

receive antenna described in [75], [73] is expressed as:

$$y_j[n,k] = \sum_{i=1}^{M_t} H_{ij}[n,k]x_i[n,k] + w_j[n,k], \tag{6.26}$$

for $j = 1, \ldots, M_r$ and $k = 0, \ldots, K-1$, where $x_i[n,k]$ is the symbol transmitted from the ith transmit antenna at the kth subcarrier of the nth block, $H_{ij}[n,k]$ is the channel's frequency response at the kth subcarrier of the nth block corresponding to the ith transmit and the jth receive antenna, and, $w_j[n,k]$ is additive (complex) Gaussian noise that is assumed to be independent and identically distributed (i.i.d.) with zero-mean and variance σ.

The received symbol in each antenna is independent from one another. Following the notation in [75] and omitting the receiver antenna subscript j the estimated channel is as follows:

Define

$$\begin{aligned} y[n] &\triangleq (y[n,0], y[n,1], \cdots, y[n,K-1])^T, \\ \hat{D}_i &\triangleq diag\{\hat{x}_i[n,0], \cdots, \hat{x}_i[n,K-1]\}, \\ w[n] &\triangleq (w[n.0], w[n,l], \cdots, w[n,K-1])^T, \end{aligned}$$

where $\hat{x}_i[n,k]$'s are detected symbols. Let rank-χ approximation of the channel response be:

$$(H_i[n,0], \cdots, H_i[n,K-1])^T = Uh_i[n], \tag{6.27}$$

where

$$h_i[n] \triangleq (h_i[n,0], \cdots, h_i[n,\chi-1])^T,$$

and U is the rank-χ transform matrix which is

$$U = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_K & \cdots & W_K^{\chi-1} \\ \vdots & \cdots & \ddots & \vdots \\ 1 & W_K^{(K-1)} & \cdots & W_K^{(\chi-1)(K-1)} \end{pmatrix},$$

where $W_K = e^{-j\frac{2\pi}{K}}$ and equation (6.26) can be written as

$$x[n] = \sum_{i=1}^{M_t} \hat{D}_i[n] U h_i[n] + w[n]. \tag{6.28}$$

Let

$$\hat{h}[n] \triangleq (\hat{h}_1^T[n], \hat{h}_2^T[n], \cdots, \hat{h}_{M_t}^T[n]),$$

from [73], the LS estimate is given as

$$\hat{h}[n] = (T^H[n]T[n])^{-1} T^H[n] y[n] \tag{6.29}$$

where

$$T[n] = (\hat{D}_1[n]U, \hat{D}_2[n]U, \cdots, \hat{D}_{M_t}[n]U). \tag{6.30}$$

The matrix inversion introduces complexity to the system. There are proposals that use training sequences which optimize and simplify the channel estimation during the training period [74].

6.3 Equalization

One of the advantages of a multi-carrier system is its robustness against inter-symbol interference. The longer duration of OFDM symbols provides higher immunity against delay spread and ISI. As long as channel dispersion is not longer than the OFDM symbol guard interval, system performance does not degrade due to ISI and use of the time domain equalization is not usually mandated.

However, in case of higher data rates and channels with extensive time dispersion, an equalizer is unavoidable. Albeit, the structure of the equalizer is different from that of single carrier systems. The purpose of equalization is not the complete removal, but the restriction of inter-symbol interference to a tolerable extent.

In frequency domain equalization, the absence of inter-channel interference (ICI) is used to compensate for channel complex gain

at each sub-carrier frequency. Using time-frequency duality, complex gain compensation after FFT is equivalent to a convolution of a FIR filter in time domain (residual equalization).

In this section, we review main differences of equalization in single and multi-carrier and follow up with analytical and practical aspects of time domain equalization. Then, frequency domain equalization and echo cancellation are discussed. Using duality principle, we present a unified approach to time and frequency domain equalization.

6.3.1 Time Domain Equalization

A brief review of decision feedback equalizer structure and performance should help the reader to readily understand the analysis and design of equalizers for multi-carrier.

Decision Feedback Equalizers

In a decision feedback equalizer, the forward filter whitens the noise and produces a response with post-cursor ISI only, and the feedback filter then cancels that post-cursor ISI. The optimum forward filter for a zero-forcing DFE can be considered as cascade of a matched filter followed by a linear equalizer and a causal whitening filter whose transfer function can be found by spectral factorization of the channel power spectrum. The forward filter, therefore, first minimizes ISI, then noise power at the input of slicer, reintroducing causal ISI. Notice that the separation and ordering has only analytic significance [79]. In practice, all forward filters are usually lumped together in one adaptive filter.

In an MSE decision feedback equalizer, the optimum forward filter minimizes the mean-square residual ISI and noise power. In both MSE-DFE and ZF-DFE we intend to minimize the error term at the input of the slicer, which in the absence of error propagation includes

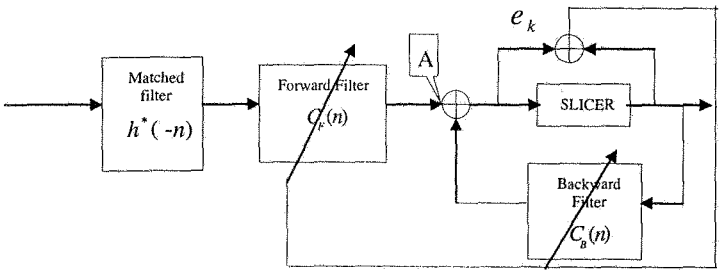


Figure 6.10: General decision feedback equalizer

residual ISI and noise. Let $h(n)$ represent the channel impulse response, hence $h^*(-n)$ is the impulse response of the matched filter. Forward filter $c_F(n)$ performs ISI suppression and noise whitening. The feedback filter $c_B(n)$ eliminates the interference of previous symbols, hence $1+c_B(n)$ has to be a monic and causal filter. Therefore, its output should have post-cursor inter-symbol interference only, namely

$$h(n) * h^*(-n) * c_F(n) = 1 + c_B(n) \tag{6.31}$$

On the other hand the variance of the noise, at the output of the matched filter is

$$\sigma^2 \int_{-\pi/2}^{\pi/2} S_h(e^{j\omega t}) d\omega \tag{6.32}$$

where S_h is power spectrum of the channel impulse response. In order to minimize noise power at the input of the slicer, $c_F(n)$ should whiten the noise. Rewriting Equation 6.31 in the frequency domain, the forward filter transfer function is

$$C_F(z) = \frac{1 + C_B(z)}{H(z)H^{*-1}(z^{-1})}. \tag{6.33}$$

The forward filter may apply further linear transformation to the input signal to meet some optimality criteria such as minimizing square

error or peak distortion. To whiten a stochastic process with spectrum of $\sigma^2 S_h(z)$ it should satisfy the following equation:

$$\sigma^2 S_h(z)\frac{1+C_B(z)}{H(z)H^{*^{-1}}(z^{-1})}\frac{1+C_B^*(z^{*^{-1}})}{H^*(z^{*^{-1}})H(z)} = \sigma_{DFE}^2 \tag{6.34}$$

Since

$$S_h(z) = H(z)H^{*^{-1}}(z^{-1}) \tag{6.35}$$

from Equation 6.34 we have:

$$S_h(z) = (1+C_B(z))(1+C_B^*(z^{*^{-1}})). \tag{6.36}$$

The same procedure [79] can be followed for the MSE equalizer to obtain

$$S_h'(z) = (1+C_B(z))(1+C_B^*(z^{*^{-1}})). \tag{6.37}$$

In the following, we discuss some important issues about the DFE with finite tap FIR feedback and forward filters as a preface for future discussions.

The input to the feedback filter is noise free, therefore the feedback filter eliminates the post-cursor ISI both in ZF-DFE and MSE-DFE. The error term used for training is memoryless and defined as:

$$e_k = x_k - \hat{x}_k. \tag{6.38}$$

where x_k is transmitted symbol, which is the same as the output of the slicer when there is no error and $\hat{x}_k$ is the output of the equalizer. Minimizing $|e_k|^2$ requires that the input signal be orthogonal to the error term. Since the input to the feedback filter is noise free, we design the forward filter to minimize the error power which leads to the following equations:

$$E(\varepsilon_k y_{k-i}^*) = E(x_k - \sum_j c_{F_j} y_{k-j} - \sum_{j'} c_{B_{j'}} x_{k-j'}) y_{k-i}^* = 0 \tag{6.39}$$

where $-\infty < i < \infty$ and:

$$y_k = \sum_{n=0} p_n x_{k-n} + \eta_k. \tag{6.40}$$

and

$$p_i = \sum_{n=0}^{l-i} h_n h_{n+i}^* \quad i = 0, \cdots, l$$

$$p_{-i} = p_i^*.$$

where l is channel spread in symbols.

Therefore,

$$\sum c_{F_i} E\{\sum_n p_n x_{k-n-j} \sum_{n'} p_{n'}^* x_{k-n'-i}^*\} + \sum_j c_{F_j} E\{\eta_{k-j}\eta_k^*\}$$

$$= \sum_{n=0}^{N} p_n E(x_k x_{k-n-i})$$

$$\Rightarrow \sum_j c_{F_j} \sum_n p_n p_{n+j-i}^* + N_o \sum_j c_{F_j} p_j = p_{-i}^*. \tag{6.41}$$

η_k is a colored Gaussian noise with the following auto-correlation function

$$E(\eta_i \eta_j^*) = \begin{cases} N_o p_{i-j}, & |i-j| \leq l \\ 0, & otherwise \end{cases} \tag{6.42}$$

The solution to Equation 6.39 is

$$C_F = \Gamma^{-1}\theta, \tag{6.43}$$

where

$$\Gamma_{ij} = E\{y_{k+i} y_{k+j}^*\}, \ 0 \leq i, j < N. \tag{6.44}$$

N is the number of forward taps and

$$\theta_i = E\{x_k y_{k+i}\}, \ 0 \leq i < N. \tag{6.45}$$

Coefficients of the backward filter can be found by

$$C_B = P^T C_F, \tag{6.46}$$

where

$$C_F = \begin{pmatrix} c_{F_0} \\ c_{F_{-1}} \\ \\ c_{F_{-N}} \end{pmatrix} \quad C_B = \begin{pmatrix} c_{B_1} \\ c_{B_2} \\ \\ c_{B_M} \end{pmatrix} \quad P = \begin{pmatrix} p_0 & p_1 & \cdots & p_M \\ p_1 & p_2 & & p_{M+1} \\ & & & \\ p_N & & & p_{N+M} \end{pmatrix} \tag{6.47}$$

The forward filter c_F is anti-causal which means that the main tap is assumed to be the closest one to the slicer. The forward taps should remove the entire post-cursor. Practically, the matched filter is usually absorbed in the forward equalizer[1] then replaces is Equations 6.42 and 6.46. Notice that the feedback filter is designed to minimize the post-cursor ISI at the input of the slicer. In practice, c_B and c_F are adjusted adaptively.

6.3.2 Equalization in DMT

Multi-carrier systems are robust against inter-symbol interference as long as orthogonality of adjacent symbols are preserved in the frequency domain. In other words, the symbol duration of the OFDM signal is extended to beyond T. However, if channel impulse response spread is more than the prefix extension, inter-symbol interference can degrade the performance.

There are several fundamental differences between time domain equalization requirements of OFDM channels and classic decision feedback equalizers:

a) An equalizer of an OFDM system does not need to cancel the ISI entirely, but to limit its length. Unlike a single carrier system in which equalization minimizes ISI, we only need to reduce it to a time span less than the length of guard interval.

b) The channel impulse response is modelled as an ARMA system. Usually, the channel impulse response is too long. Therefore, a model of the form

$$\frac{A(z)}{1+B(z)} \tag{6.48}$$

is appropriate.

[1] We can interpret this as replacing the matched filter with a whitening matched filter.

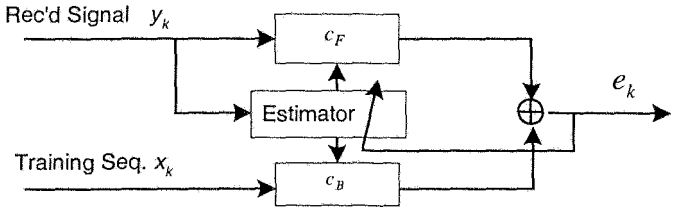


Figure 6.11: Equalizer configuration in training mode

c) The equalizer does not work as a DFE in data mode. Unless channel characteristics change rapidly, equalizer taps are adjusted during the training sequence and the same tap values are preserved during data mode.

d) The feedback filter is not in the loop and need not to be monic or causal. In single carrier systems, the slicer output is fed to the feedback filter to remove the post-cursor effect. However, truncating the impulse response does not require a monic feedback filter. A typical structure is shown in Figure 6.11

e) The error is not memoryless. The error term used in typical DFE is derived from the scalar input and output of the slicer. In a DMT equalizer the error is derived from the output of an impulse shortening filter and reference filter.

f) The location of the impulse response window has a big impact on the performance of equalizer.

The overall structure of the equalizer in training mode is shown in the following diagram:

In steady state mode, the equalizer will include only c_F . The error term is

$$e_k = c_F * Y_k - c_B * X_k \tag{6.49}$$

where X_k and Y_k are vectors of training and received samples respectively. In general, we can use adaptive techniques to find near optimum equalizer coefficients. A large number of taps prevents us from using RLS type of algorithms. On the other hand, a wide spread of eigenvalues of the input signal covariance matrix can slow down convergence of the equalizer. A different technique for the equalizer tap adjustment is based on the linear prediction theory as explained below.

During acquisition, the frame is repeated and can be used for equalizer training. In this approach, we use second order statistics of the channel to produce a one shot estimate of equalizer coefficients. In order to use the predictor format of linear regression techniques, without loss of generality, assume c_F is monic. So, the predictor format is:

$$\hat{y}(n|n-1) = \varphi^T(n)C \tag{6.50}$$

where

$$\begin{aligned} \varphi(n) \quad = [&-y(n-1) \\ &-y(n-2) \\ &\vdots \\ &-y(t-N) \\ &x(n-1) \\ &x(n-2) \\ &\vdots \\ &x(n-M)] \end{aligned} \tag{6.51}$$

and,

$$C = [\ C_F \quad C_B\ ]^T. \tag{6.52}$$

The least square solution of linear regression is:

$$\hat{C} = \Gamma^{-1}\theta \tag{6.53}$$

with:

$$\Gamma = \begin{pmatrix} A_{N\times N} & B_{N\times M} \\ B^*_{M\times N} & I_{M\times N} \end{pmatrix}, \tag{6.54}$$

where

$$A_{ij} = R_{yy}(i-j) = \frac{1}{N}\sum_{n=1}^{N} y(n-i)y(n-j), \tag{6.55}$$

and

$$B_{ij} = R_{xy}(i-j) = \frac{1}{N}\sum_{n=1}^{N} x(n-i)y(n-j), \tag{6.56}$$

and

$$\theta = [\ D_N \quad E_M\ ]^T \tag{6.57}$$

with

$$D_i = R_{yy}[i],\ E_i = R_{yx}[i], \tag{6.58}$$

where R_{yy} and R_{xy} are auto-correlation and cross correlation functions. We can use channel impulse response to calculate coefficients directly

$$A_{ij} = \sum_{n} h_n h^*_{n+j-i} \tag{6.59}$$

$$B_{ij} = h^*_{-i+j}. \tag{6.60}$$

The estimation of channel impulse response can help in choosing the optimum window location. As explained previously, c_B is not necessarily causal, therefore, delay parameter plays an important role in the performance of the equalizer. Other important parameters are the lengths of the two filters c_F and c_B.

6.3.3 Delay Parameter

Unlike classic decision feedback equalizers, an ARMA equalizer can have a feedback filter with non-causal transfer function i.e. the general error term is:

$$c_F * Y_k - c_B * X_{k-d} \tag{6.61}$$

the delay unit d has significant effect on overall performance. The number of taps M of filter c_B is approximately the same as the prefix length and number of taps for the forward filter $N \leq M$. The delay unit should be chosen properly to capture the most significant window of the impulse response. Brute force trial and error is one option. A more intelligent technique uses the estimate of the impulse response to pick a proper starting point where the energy of the impulse response is above a threshold. The threshold should be the ratio of the tap power to the overall window power. Since the length of the window function and the forward equalizer should be as short as possible, the proper choice of starting point is of significant importance. In other words, in an ARMA model, the number of zeroes are not assigned properly.

6.3.4 AR Approximation of ARMA Model

The Direct solution of Equation 6.53 requires a non-Toeplitz matrix inversion and is numerically sensitive and intensive. One technique to alleviate the numerical difficulty of least square solution is by using techniques for AR approximation of the channel response ARMA model, which results in a numerically attractive Toeplitz correlation matrix. A description of embedding techniques is presented in the Appendix. In the following, the procedure for this analysis is briefly described. By using embedding techniques, as explained in the Appendix, the ARMA model of Equation 6.48 is approximated by a two-channel AR model.

If the covariance matrix is available the problem can be reduced to the solution of a normal (Yule-Walker) equation. The multi-channel version of the Levinson algorithm is a computationally efficient solution for the two-channel AR model at hand. As a reference, the recursive technique is presented here:

Initialization:

$$\begin{aligned} C_1^f[1] &= -R^{-1}[0]R[1] \\ C_1^b[1] &= -R^{-1}[0]R[-1] \end{aligned} \tag{6.62}$$

where

$$C_i^f[j] = \begin{bmatrix} -a_i^j & b_i^j \\ 0 & 0 \end{bmatrix}$$

is the jth forward prediction coefficient of two-channel model at i^{th} iteration and $C_i^b[j]$ is the jth backward prediction coefficient at i^{th} iteration and

$$R[i] = \begin{bmatrix} R_{yy}[i] & R_{yx}[i] \\ R_{xy}[i] & R_{xx}[i] \end{bmatrix}.$$

If the input signal x is white then

$$R[i] = \begin{bmatrix} R_{yy}[i] & R_{yx}[i] \\ 0 & 0 \end{bmatrix}$$

with

$$R[0] = \begin{bmatrix} R_{yy}[i] & R_{yx}[i] \\ R_{xy}[i] & I \end{bmatrix}$$

Recursion:

$$\begin{aligned} C_i^f[i] &= K_i^f \\ C_k^f[i] &= C_{k-1}^f + K_k^f[i]C_{k-1}^b[k-i] \\ C_i^b[i] &= K_i^b \\ C_k^b[i] &= C_{k-1}^b + K_k^b[i]C_{k-1}^f[k-i] \\ C_k^b[i] &= C_{k-1}^b + K_k^b[i]C_{k-1}^f[k-i] \end{aligned} \tag{6.63}$$

where the reflection coefficients are,

$$\begin{aligned} K_k^f &= -\Delta_k^{f^T} \textstyle\sum_{k-1}^{b^{-T}} \\ K_k^b &= -\Delta_k^{f^T} \textstyle\sum_{k-1}^{b^{-T}} \\ \Delta_k^f &= \textstyle\sum_{i=0}^{k-1} R[k-i]C_{k-1}^{f^T}[i]. \end{aligned} \tag{6.64}$$

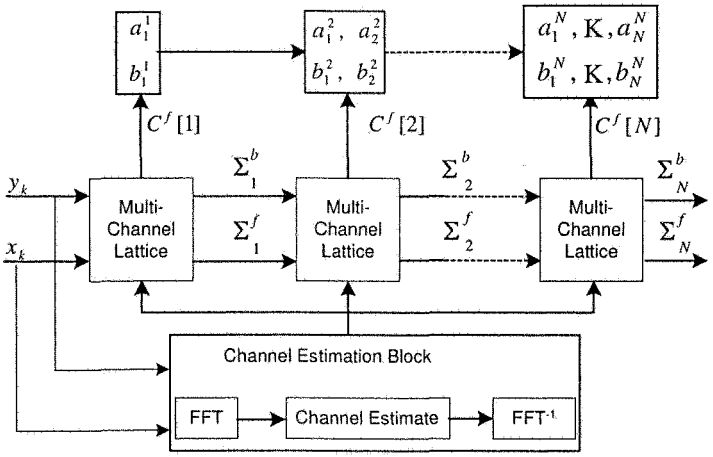


Figure 6.12: Equalization using order update

Figure 6.12 shows the structure of an equalizer using a lattice filter. This structure represents the case of equal number of poles and zeros. The extension to unequal number of poles and zeroes is straightforward and explained in [79].

The above approach is convenient for batch or frame based processing where the entire frame is buffered and processed for equalization. Instead of inverting the large matrix of Equation 6.43 for example, this structure can be used in the training mode of an ADSL receiver. This technique can be implemented iteratively in time and order when the covariance is not known in advance. The time/order updating algorithm increases the order while updating the coefficients of the equalizer. In some cases, we can preserve the order and continue updating the parameters using adaptive techniques. This technique is useful when the delay profile of the channel does not change significantly but some slow adaptation improves the performance.

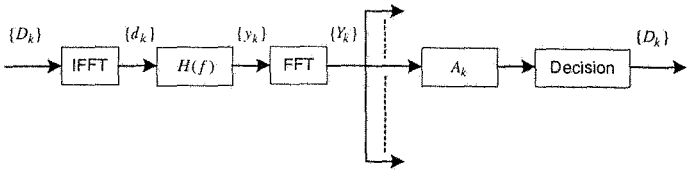


Figure 6.13: An OFDM system with frequency domain equalization

6.3.5 Frequency Domain Equalization

Once it is assured that orthogonality of the sub-carriers is maintained, possibly through the use of the cyclic prefix and time domain equalization as previously discussed, then the final frequency domain equalization of an OFDM signal is an extremely simple process. This is certainly one of the key advantages of OFDM.

After demodulation, the sub-carriers will be subjected to different losses and phase shifts, but there will be no interaction among them. Frequency domain equalization therefore consists solely of separate adjustments of sub-carrier gain and phase, or equivalently of adjusting the individual decision regions. In the case where the constellations consist of equal amplitude points, as in PSK, this equalization becomes even simpler in that only phase needs to be corrected for each sub-carrier, because amplitude has no effect on decisions.

A simplified picture of the place of frequency domain equalization is shown in Figure 6.13, where the equalizer consists of the set of complex multipliers, $\{A\}$, one for each sub-carrier.

Here the linear channel transfer function $H(f)$ includes the channel, the transmit and receive filters, and any time domain equalization if present. $H(f)$ is assumed bandlimited to less than N/T for a complex channel. Cyclic extension is not shown although it is almost certain to be present. The following analysis assumes that both amplitude and phase need to be corrected, and that equalization consists

of multiplying each demodulated component by a quantity such that a fixed set of decision regions may be used.

The signal presented to the demodulator is

$$y(t) = \sum_{n=0}^{N-1} d_n h(t - n\frac{T}{N}). \tag{6.65}$$

The kth output of the demodulator is then

$$Y_k(f) = D_k H_k, \; k = 0, 1, \cdots, N-1, \tag{6.66}$$

where the H_k are samples of $H(f)$

$$H_k = h(\frac{k}{T}). \tag{6.67}$$

Thus each output is equal to its associated input data symbol multiplied by a complex quantity which differs among the outputs, but are uncoupled. Equalization at its simplest then consists of setting the multipliers to $1/H_k$ for each non-zero channel.

The above approach is optimum in every sense under high signal-to-noise conditions. It also produces minimum probability of error at any noise level, and is an unbiased estimator of the input data D_k. However if the criterion to be optimized is minimum mean-square error particularly, then the optimum multipliers are modified to

$$A_k = \frac{1}{H_k} \frac{1}{1 + \frac{\sigma_k^2}{|D_k H_k|^2}} \tag{6.68}$$

where σ_k^2 is the noise power in the demodulated sub-channel. However, this value produces a biased estimator and does not minimize error probability.

As a practical implementation issue, for variable amplitude constellations, it is frequently desirable to have a fixed grid of decision

regions. The A_ks can then be scaled in amplitude such that the separation of constellation points is constant.

Since a shift τ in timing phase is equivalent to a phase shift

$$H_k = e^{j2\pi \frac{k}{T}\tau} \tag{6.69}$$

frequency domain equalization readily corrects for such timing shift.

In principle frequency domain equalization could be employed when orthogonality is lost due to interference among OFDM symbols. In this case, rather than a simple multiplier per sub-channel, a matrix multiplication would be required. This approach is bound to require more computational load than the combination of time and frequency domain equalizers.

During system initialization, the any time domain equalizer must be adjusted before frequency domain equalization is performed. Then any periodic test signal with full frequency content, such as a repeated segment of a PN sequence without cyclic prefix, may be used to adapt the frequency domain equalizer. The demodulator outputs should be averaged over many periods to minimize the effects of noise, and the reciprocal of these measured values used to set the multipliers. During steady state data operation adaptation may be performed using a single variable LMS algorithm as shown in Figure 6.14. This LMS algorithm will adapt to the values given by Equation 6.68.

An interesting interpretation of time domain and frequency domain equalizer can be obtained by studying their role in channel distortion compensation. As discussed earlier, a typical channel model for DMT and channels with long impulse response is an ARMA model of the form in Equation 6.48. Time domain equalization shortens the impulse response to a tolerable level for an OFDM system. Mathematically, it's equivalent to compensating the AR part of the channel impulse response $1/(1 + B(z))$. So after successful time domain equalization, the equivalent impulse response of the channel is reduced to a FIR filter of short duration $A(z)$. Since it does not violate orthogonal-

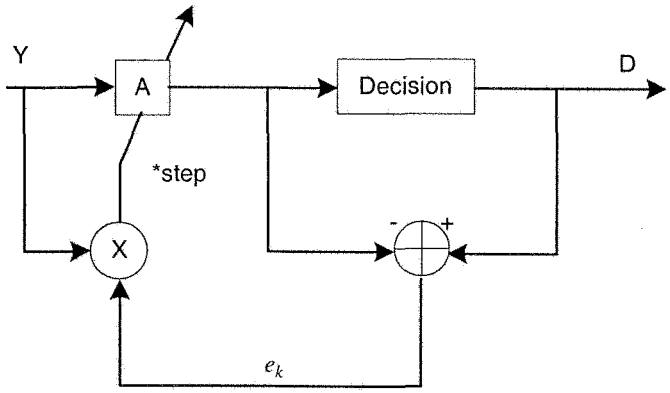


Figure 6.14: LMS adaptation of a frequency domain equalizer multiplier

ity of sub-carriers, we can remove its effect after the FFT by frequency domain equalization. The above process is shown in Figure 6.15.

6.3.6 Echo Cancellation

In wireline applications, it is usually highly desirable to communicate full duplex (simultaneous transmission in both directions) over a single wire-pair. A classic sub-system known as the hybrid provides partial separation, but is not sufficient for high performance digital applications.

Two possible techniques are frequency division and time division duplexing. In the former, different frequency bands are used in each direction, with analog filtering to provide separation. In the latter, alternating time slots are used to transmit alternately in each direction. For symmetric communication, that is equal rates for the two directions, both of these techniques require at least double the bandwidth

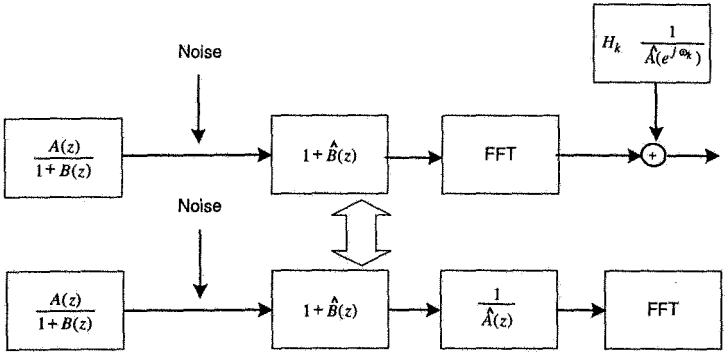


Figure 6.15: Time and frequency domain equalization

as unidirectional transmission, which is a very severe performance penalty.

Far preferable is the technique of echo cancellation, where at each end the leakage of transmitted signal into the co-located receiver is cancelled by creating and subtracting a replica of that leakage. The precision required is quite high, as illustrated in the example of a bad case wireline system shown in Figure 6.16

In this example, the loss of the line is 40 dB while the hybrid provides only 15 dB of loss. The leakage of the transmitted signal, or "echo", is therefore 25 dB higher than the desired received signal. One consequence is that the level which the front end of the receiver, in particular the A/D converter, must accommodate is dominated by the echo rather than the received signal. Of more interest here is the high degree of echo cancellation required. Assume, as in the figure, that a signal-to noise ratio of 30 dB is required, and that uncancelled echo can be treated as noise. If we use the reasonable design rule that the uncancelled echo should be at least 10 dB below the line noise, thereby increasing the noise by less than 0.4 dB, the level of the uncancelled echo must be below -80 dB. Because the hybrid provides only 15 dB of reduction, the echo canceller must provide an additional

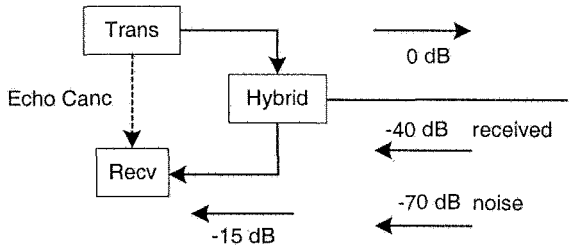


Figure 6.16: An example system illustrating echo cancellation requirement

65 dB of rejection. This requirement is far more severe than that of equalization, so that a much greater arithmetic precision and number of taps is required.

In principle, the same form of echo canceller used in single carrier systems could be used in OFDM. This is shown in Figure 6.17.

Residual echo is minimized when the tap coefficients of the echo canceller are equal to the corresponding samples of the echo path response $p_k = p(kT)$, where the sampled echo is

$$r_n = r(nT) = \sum_k d_k p_{n-k} \tag{6.70}$$

Adaptation is similar to that of equalization. Residual error results from echo samples beyond the time span of the canceller, and also from finite precision in signal sampling and in the coefficients. Unlike the case of single carrier equalization, the inputs to the canceller are samples of signals quantized by an A/D converter as opposed to exact data symbols. In one variation [80], adaptation is performed in the frequency domain, followed by an IFFT and time domain cancellation.

In any variation, time domain cancellation requires extensive computation, reducing a prime advantage of OFDM. Another possibility is

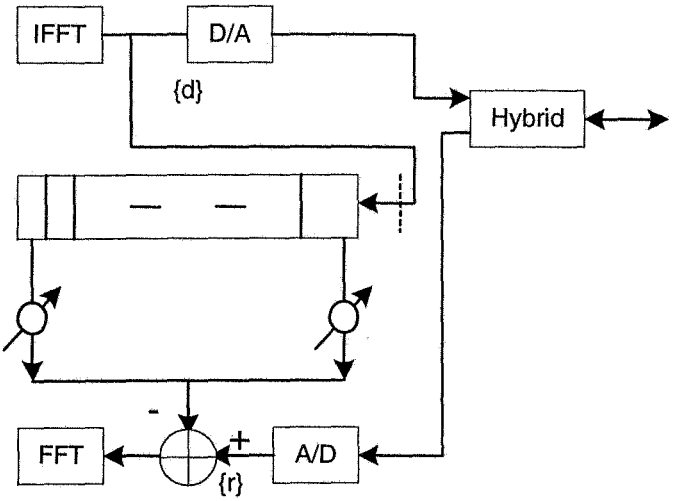


Figure 6.17: A basic echo canceller

frequency domain cancellation. If the echo of the transmitted OFDM signal maintains orthogonality, then echo cancellation may be performed very easily after demodulation in the receiver. An advantage of this approach is the use of exact data values as inputs to the canceller rather than quantized signal samples. Furthermore, orthogonality leads to simple adaptation. The canceller consists of one complex multiplication per sub-carrier, as in the frequency domain equalizer. In order for this technique to be valid, there must be no inter-symbol interference among OFDM signals. Otherwise a matrix multiplication is required rather than a single multiply per sub-carrier [81]. In another approach described in Reference [82], the time domain equalization is modified so as to shorten both the channel response and the echo response to a time span less than the duration of the cyclic prefix. The resulting equalizer and echo canceller can then both be performed quite simply in the frequency domain as shown in Figure 6.18.

Techniques have been proposed for echo cancellation in OFDM

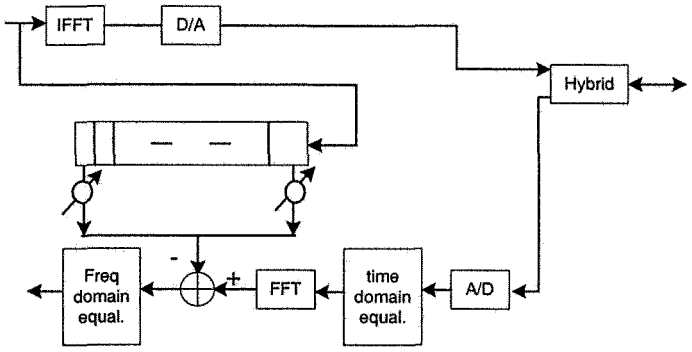


Figure 6.18: Frequency domain echo cancellation

that combine time and frequency domain components in ways to reduce total required computation [82], [84]. Shown in Figure 6.19 is the approach of Reference [82] for a symmetric system.

The time domain canceller is a short one if the echo response is not much longer than the duration of the cyclic prefix. Its functions are to undo the effect of cyclic extension, and to cancel the overlap from the previous symbol. The frequency domain canceller can now deal with orthogonal symbols as before.

In asymmetric applications such as ADSL, the penalty for using frequency division duplexing is not as great because the bandwidth used in one direction is relatively small. However for optimum performance, the use of overlapping bands and echo cancellation is employed. Shown in Figure 6.20 and Figure 6.21 are modifications of the previous structure [82] when the same OFDM symbol rate but different numbers of sub-carriers are used in each direction, with overlap of the frequencies of some of the sub-carriers.

Figure 6.11 illustrates a case in which the transmit rate is lower than the receive rate, as in the remote terminal of ADSL. Because the transmit spectrum is smaller than the receive spectrum, it is replicated

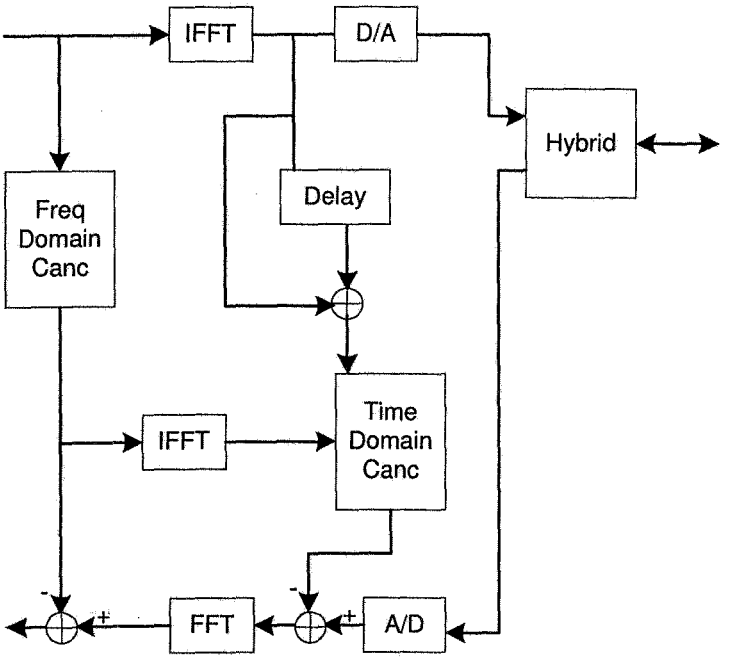


Figure 6.19: Combination echo cancellation for symmetric transmission

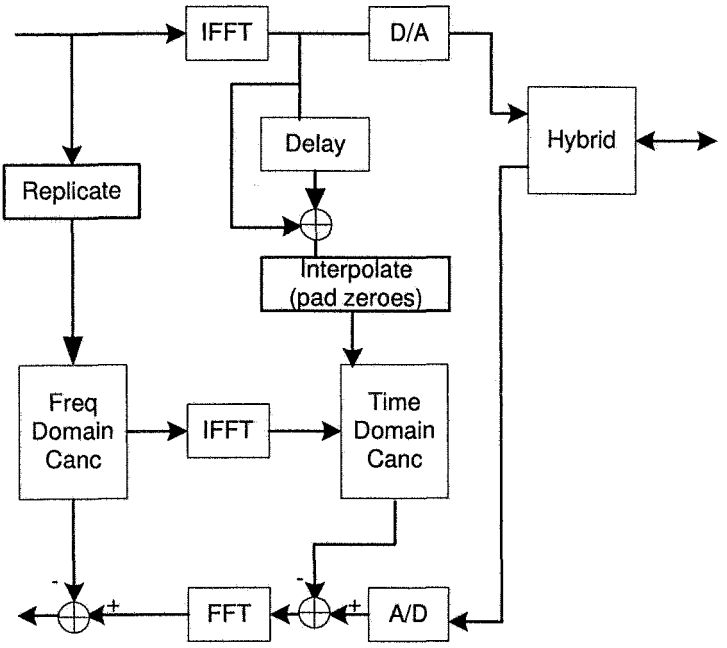


Figure 6.20: Echo cancellation where the transmit rate is lower than receive rate

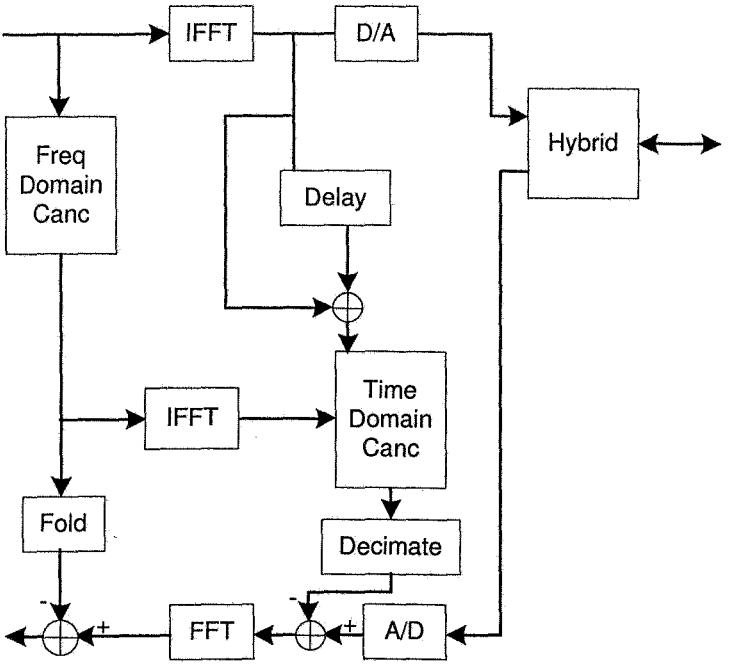


Figure 6.21: Echo cancellation where the receive rate is lower than transmit rate

so that the number of sub-carriers is increased to equal that of the received signal. For the time domain portion of the echo canceller, the transmit signal is interpolated to the higher rate by inserting the appropriate number of zeroes between the time samples.

The converse case, in which the transmit rate is higher than the receive rate, as in the central office terminal of an ADSL link, is shown Figure 6.21. The size of the input to the frequency domain canceller is reduced by folding subsets of the OFDM symbol on top of each other and adding them to produce a number of sub-carriers equal to the number received. The time domain sampling rate is reduced by computing the short echo at further spaced points.

Initial adaptation, for any echo cancellation structure, may be performed at the same time and with the same training signals as the equalization at the other end. Having the remote transmitter silent improves this initial adaptation.

6.3.7 Appendix - Joint Innovation Representation of ARMA Models

Some of the least squares estimation algorithms developed for AR models such as ladder form realization, can be extended to ARMA models by using embedding technique. Namely, an ARMA model can be embedded into a two channel AR model of joint input/output process [85]. The same approach can be used to model an equalizer for an ARMA channel by extending the results of the equalizer design for AR channels. Unlike AR models, the covariance matrix of an ARMA model is not Toeplitz and that results in some numerical computational difficulties. Therefore, by using embedding techniques we can transform an ARMA model to a multi-channel AR model with a block-Toeplitz covariance matrix.

First, we present a geometric interpretation of prediction techniques for ARMA processes which is the basis for understanding Levinson Algorithm.

In linear prediction theory, we predict x_n by $\hat{x}_n^m$, a linear combination of previous m samples $\{x_{n-1}, x_{n-2}, \cdots, x_{n-m}\}$ as a mth order linear predictor, such that an error criteria like

$$E\{|x_n - \hat{x}_n^m|^2\} \tag{6.71}$$

is minimized. The concept of an innovation process is a critical one in prediction and estimation. It represents the amount of new information obtained by observing a new sample x_n which was not available in previous observations:

$$v_n = x_n - x_{n|n-1} \tag{6.72}$$

Estimation of sample x_n using the previous m samples leads to an mth order forward linear predictor

$$\hat{x}^m_{n|n-1} = -\sum_{k=1}^{m} c[k]x_{n-k} \tag{6.73}$$

while using future m samples leads to a backward m^{th} order linear predictor

$$\hat{x}^m_{n|n-1} = -\sum_{k=1}^{m} c[k]x_{n+k} \tag{6.74}$$

Error terms for backward and forward predictors represent an innovation process.These error terms are defined as:

$$\begin{aligned} e^m_f[n] &= x_n - \textstyle\sum_{k=1}^{m-1} c[k]x_{n-k} \\ e^m_b[n] &= x_{n-m} - \textstyle\sum_{k=1}^{m-1} c[k]x_{n-m+k}. \end{aligned} \tag{6.75}$$

A recursive implementation of linear predictors utilizes Gram-Schmidt orthogonalization to transform samples into orthogonal random variables [86]

By interpretation of the expression in Equation 6.71 as a metric of the Hilbert space spanned by process x, a useful geometrical representation for linear prediction can be offered as shown in Figure 6.22. Next we briefly review the Levinson algorithm which is the origin of important concepts such as lattice filters and equalizers. Originally, the Levinson algorithm was proposed as an efficient numerical technique for inverting Toeplitz matrices.

Its order-recursive nature leads to important concepts in estimation and equalization. Minimizing the error criteria 6.71 for the predictor of order results in the normal equation

$$\Phi_m c_m = \phi_m \tag{6.76}$$

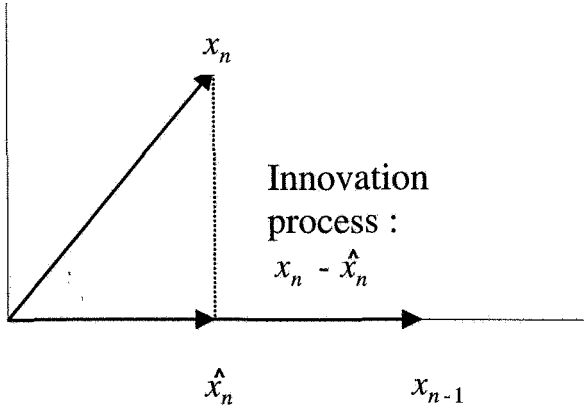


Figure 6.22: Geometric interpretation of linear predictors

where Φ_m is Toeplitz matrix of correlations $\Phi_{kl} = \phi(k-l)$ with

$$\phi(n) = E\{x_l x_{l+n}\} \tag{6.77}$$

The Levinson algorithm calculates the n coefficients of nth order linear predictor using $n-1$ coefficients of $(n-1)^{th}$ order linear predictor. As a reference we provide the scalar Levinson equations:

Initialization:

$$c_1[1] = \frac{-\phi[1]}{\phi[0]}, \; e_0 = \phi[0] \tag{6.78}$$

Recursion:

$$\begin{aligned} c_k[k] &= -\frac{\phi[k]+\sum_{l=1}^{k-1} c_{k-l}[l]\phi[k-l]}{\varepsilon_{k-1}} \\ c_k[i] &= c_{k-1}[i] + c_k[k]c^*_{k-1}[k-i] \\ e_k &= (1-|c_k[k]|^2)e_{k-1} \end{aligned} \tag{6.79}$$

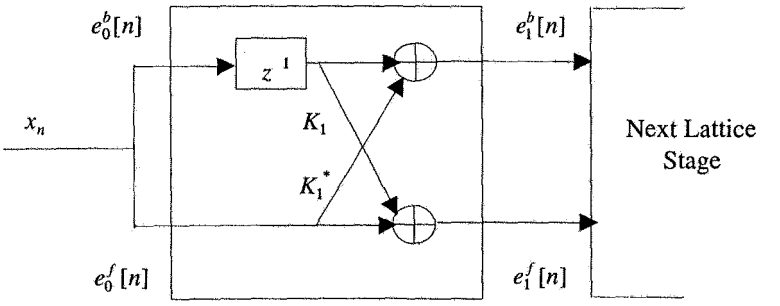


Figure 6.23: Lattice filter representation of predictors

where $c_k[i]$ represents i^{th} coefficient of the kth order predictor. In the above equations, we did not explicitly distinguish between backward and forward coefficients because optimal backward coefficients are reversed in time and complex conjugate of optimal forward predictor coefficients.

When the exact order of the predictor is unknown, we can monitor the error term e_k after each iteration. An extensive treatment of this issue is discussed in model order estimation theory.

The above order recursive technique is the basis of lattice representation of prediction filter as shown in Figure 6.23.

In general, an ARMA model can be represented as:

$$y_k + b_1 y_{k-1} + \cdots + b_N y_{k-N} = a_0 + a_1 x_k + \cdots + a_N x_{k-N} \quad (6.80)$$

where x_k is the input process and y_k represents the output process. The innovation process of y_k which represents additional information obtained from observation of sample y_k is:

$$e_{k,N} \equiv y_k - \hat{y}_{k|k-1} = \begin{bmatrix} B_N & -A_N \end{bmatrix} \begin{bmatrix} Y_N \\ X_N \end{bmatrix} \quad (6.81)$$

where

$$X_N^T = [x_k, \cdots, x_{k-N}]$$
$$Y_N^T = [y_k, \cdots, y_{k-N}]$$
$$A_N^T = [a_0, \cdots, a_N]$$
$$B_N^T = [1, b_1, \cdots, b_N]$$

The problem of linear prediction is: given a set of observed samples of $\{y_{n-1}, y_{n-2}, \cdots, y_{n-m}, x_{n-1}, \cdots, x_{n-m}\}$ an estimate of unobserved sample y_n is required, usually without covariance information. The prediction filter is chosen to minimize the power of prediction error which means designing a filter of the form

$$H_N(z) = \frac{a_0 z^N + a_1 z^{N-1} + \cdots + a_N}{z^N + b_1 z^{N-1} + \cdots + b_N} \tag{6.82}$$

such that

$$y(z) = H_N(z) e_{k,N}(z) \tag{6.83}$$

where $e_{k,N}$ is as close to a white noise process as possible.

One approach to linear prediction of ARMA models is to introduce an innovation process for input x_k, then embed it into a two-channel AR process. The general approach for embedding is discussed extensively in the literature. Here we discuss the embedding technique for the special case where x_k is white. By introducing the augmented process $w_k = [\ y_k \quad x_k\ ]$, the new prediction error is

$$e_{k,N} = C_N w_k, \tag{6.84}$$

where

$$e_{k,n} = \begin{bmatrix} e_N^x \\ e_N^y \end{bmatrix} = \begin{bmatrix} a_0 x_k \\ x_k \end{bmatrix} \tag{6.85}$$

and

$$C_N = \left[\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \begin{bmatrix} b_1 & -a_1 \\ 0 & 0 \end{bmatrix}, \cdots, \begin{bmatrix} b_N & -a_N \\ 0 & 0 \end{bmatrix} \right] \tag{6.86}$$

Geometric interpretation of the ARMA model using the embedded technique is straightforward and useful. In the ARMA model the

space of observations is spanned by augmented process $[x, y]$. Then the joint backward and forward innovations can be interpreted as orthogonal projections on the sub-spaces spanned by future and past observation respectively.

The Levinson recursion algorithm can be applied to the augmented (multi-channel) system as well. However, unlike the scalar case, backward and forward predictor coefficients are not time-reversed and complex conjugate versions of each other.

Chapter 7

Channel Coding

7.1 Need for Coding

In almost all applications of multi-carrier modulation, satisfactory performance cannot be achieved without the addition of some form of coding. In wireless systems subjected to fading, extremely high signal-to-noise ratios are required to achieve reasonable error probability. In addition, interference from other wireless channels is frequently severe. On wire-line systems, large constellation sizes are commonly employed to achieve high bit rates. Coding in this case is essential for achieving the highest possible rates in the presence of crosstalk, impulsive and other interference.

Proper coding design is extremely important for a digital communication link [89]. A designer should take several design factors into account. Those include required coding gain for intended link budget, channel characteristics, source coding requirements, modulation, etc. Coding in OFDM systems has an additional dimension. It can be implemented in time and frequency domain such that both dimensions are utilized to achieve better immunity against frequency and time

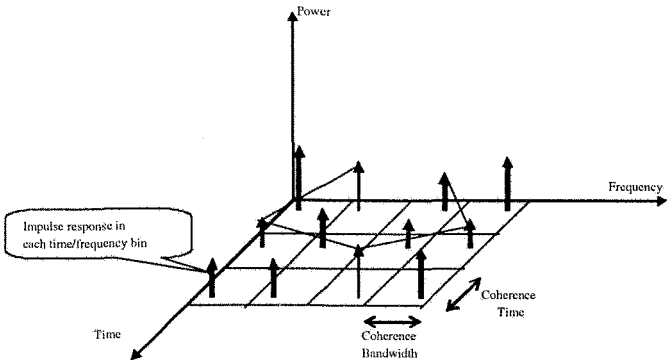


Figure 7.1: Two-dimensional coding for OFDM

selective fading. Obviously, interleaving plays an important role to achieve the above goal as shown in Figure 7.1. A combination of block coding and convolutional coding along with proper time/frequency interleaving constitutes a concatenated coding strategy. A new generation of parallel concatenated coding, namely turbo coding, also seems promising for some OFDM applications.

In this chapter, performance of block, convolutional, and trellis coding, and issues concerning concatenated and turbo coding are discussed. A comparison of their performance with theoretical bounds is also presented.

7.2 Block Coding in OFDM

In classical block coding, input data are blocked into groups of k bits, and each block is mapped into an output block of n bits, where $n > k$. In the canonical form, $n - k$ parity checks are computed among the input bits according to some algebraic procedure, and then appended to the original block. This is illustrated in Figure 7.2. This requires an

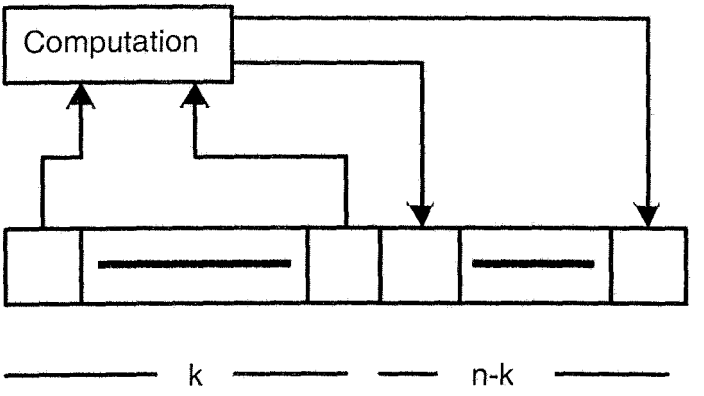


Figure 7.2: Construction of a canonical block code

increase in bandwidth by a factor of n/k. The reciprocal of this factor is the efficiency, or rate, of the code.

Only 2^k of the possible 2^n output blocks are legitimate code words. The code is chosen such that the minimum *Hamming distance*, which is the number of bits in which code words differ, is maximized. The code is described by the set of numbers $[n, k, d]$, where d is the minimum Hamming distance.

At the receiver, the n-bit block is recovered, possibly with errors, by demodulation and framing. The decoder finds the permissible code word that is closest in Hamming distance to this received block. The $n - k$ check bits may then be deleted and the result output as a replica of the original input. If $d = 2t + 1$, then any set of t or fewer errors in the block can be corrected. In OFDM, if n agrees with the number of bits in an OFDM symbol, then each symbol can be treated in the physical; separately and no memory beyond a symbol is required for decoding.

The performance improvement achieved by coding is illustrated in

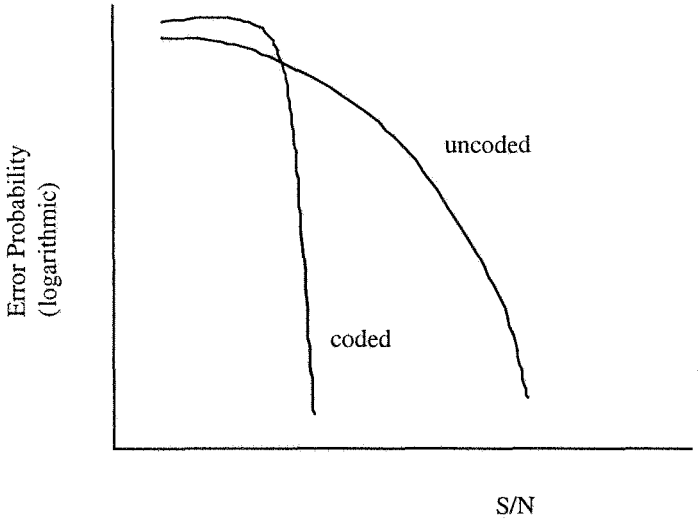


Figure 7.3: Performance with block coding over a Gaussian channel

Figure 7.3 for the Gaussian channel. Not only is the error probability substantially reduced, but the shape of the curve changes to one in which the error probability decreases very sharply with small increase in signal-to-noise ratio. For a relatively large block size, the sharp decrease occurs at a signal-to-noise ratio only a few dB higher than that which would yield an ideal capacity of the same bit rate. The sharpness of the curve implies that when such channel coding is used, little further benefit can be achieved by source coding that reduces the effects of errors.

The decoding described above is "hard decision", that is the decoder operates on the binary output of the demodulator. The actual size of error in the analog domain in the demodulator is lost. Better performance could be achieved if a measure of that analog error were used in the decoding operation. This is referred to as "soft decoding". Its complexity is clearly much greater than hard decoding.

An intermediate approach between hard and soft decoding is the introduction of erasures. In this technique, the demodulator not only detects the coded bits, but also measures the reliability of those decisions. When the reliability is below some threshold, the demodulator outputs an erasure symbol instead of a bit decision. The decoder now operates on 3-level inputs. Any combination of t errors and e erasures can be corrected if $d = 2t + e + 1$. Approximately half of the performance difference between hard and soft decision decoding can be recovered over a Gaussian channel if the erasure threshold is optimized.

Block codes can operate over symbol alphabets higher than binary. A particularly powerful and widely used family of codes is the Reed-Solomon code [90]. The alphabet size is any of the form $m = q^p$ where q is a prime number, almost always 2 in practice. An $[n, k, d]$ code maps k M-ary symbols into n M-ary symbols. d is a generalized Hamming distance, the minimum number of symbols that differ between code words. The block size n must be less than or equal to $m - 1$. A shortened block size leads to simpler implementation at the cost of performance.

As an example, consider a Reed-Solomon code that treats 8-bit sequences (bytes) as the underlying alphabet. Then $m = 2^8 = 256$, so that $n \leq 255$. A $\{255, 235, 20\}$ code maps 235-bytes (1880 bits) blocks into 255-bytes (2040 bits) blocks. It can correct up to 10 errors or 20 erasures, or an intermediate combination of both.

Block codes, in particular the Reed-Solomon class, are effective in combating burst errors. In single carrier systems, such bursts result from impulsive noise of duration greater than a symbol period, and from correlation due to fading when the time coherence of the channel is longer than a symbol period. The same considerations apply for a multi-carrier system, but in addition their frequency dual must be considered. An impulse will almost always have wide frequency content so as to affect several sub-carriers. A frequency selective fade whose coherence bandwidth is wider than the sub-carrier spacing will

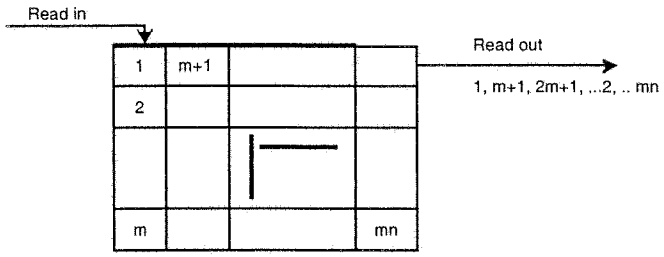


Figure 7.4: Implementation of periodic interleaving

lead to correlated errors among the sub-carriers.

The performance of codes in the presence of bursts can be improved by the process of interleaving. Rather than operating on symbols that are adjacent in time and/or frequency, the code is made to operate on symbols with sufficient spacing so that the errors are more independent. A simple and common form of interleaving is periodic interleaving illustrated in Figure 7.4. At the receiver, the symbols are de-interleaved before decoding. The decoder therefore operates on symbols spaced m symbol periods apart as transmitted. The spacing can be in time, frequency, or both.

A simple form of block coding that is very widely used for error detection, not correction, is the cyclic redundancy check (CRC). Here a fixed number of check bits are appended to a block of arbitrary length. If the receiver detects any errors, a retransmission is requested. The encoding is based on an n-bit feedback shift register whose connection pattern is a primitive polynomial. The receiver employs a shift register with the same connection polynomial. This is shown in Figure 7.5.

The operations shown are modulo-2 addition (exclusive-or). At the encoder, the shift register is initialized to a given pattern, such as all ones. The input data is fed both to the channel and to the feedback

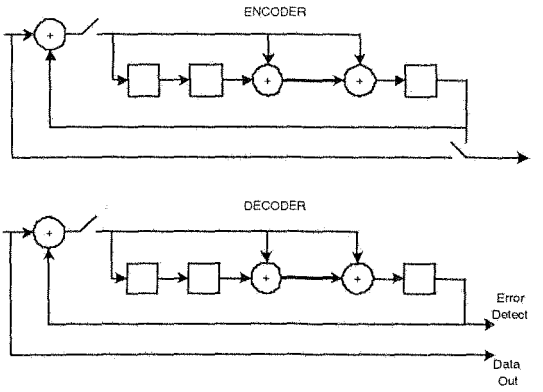


Figure 7.5: CRC implementation

shift register. After the input block is completed, the contents of the shift register are shifted out to the channel. At the receiver, the shift register is also initialized. The received data is simultaneously output and fed to the register. After completion of the un-coded bits, the contents of the register are examined. An all-zero condition of these bits indicates that no errors have occurred. A CRC of length n can detect any error pattern of length n or less. Any other error pattern is detected with probability $1 - 2^{-n}$.

7.3 Convolutional Encoding

Another very important form of coding is convolutional. Instead of operating on symbols arranged as blocks, the coding operates continuously on streams of symbols. At the encoder, the input is fed continually through a shift register of length m. The memory of the code is the "constraint length" $(m + 1)$, the number of output symbols affected by an input symbol. Each time a bit is read into the register,

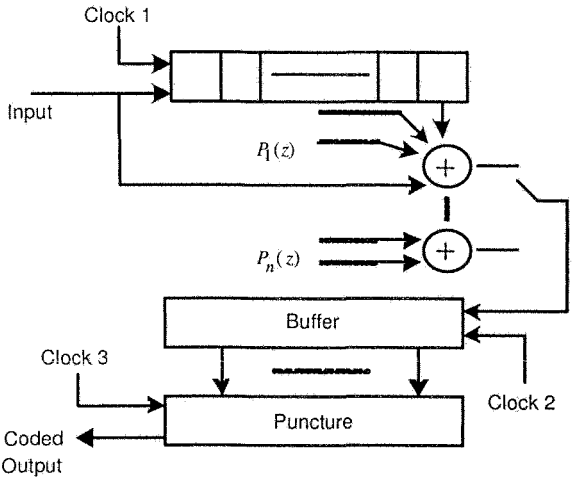


Figure 7.6: Generation of a convolutional code

several modulo-2 sums of the present and past bits are formed. The choice of which bits are operated on is designated as a polynomial $P(z)$ with binary coefficients. n such modulo-2 sums are formed and multiplexed to form the output of the "mother code". Since n bits are generated for each input bit, the rate of the code is $1/n$. In an elaboration not treated here, more than one bit at a time is read in to a corresponding number of shift registers.

The rate of the code can be increased by the process of puncturing. This involves deleting some of the bits generated by the mother code. The process of generating a convolutional code is shown in Figure 7.6. In addition, interleaving may be applied to a convolutional code in a manner similar to a block code.

A convolutional code can be described by a state diagram. For a binary convolutional code, the number of states is $s = 2^m$, where m is the number of shift register stages in the transmitter. In the state

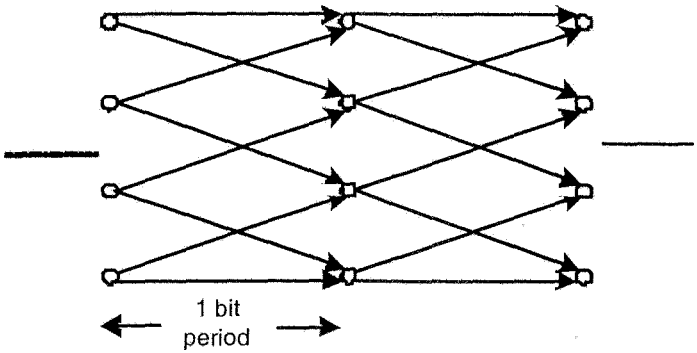


Figure 7.7: State diagram of a convolutional code

diagram shown in Figure 7.7, the transitions from a current state to a next state are determined by the current input bit, and each transition produces an output of n bits.

A particular advantage of convolutional codes is the relative ease of decoding. Sequence estimation techniques are applicable, in particular the Viterbi algorithm. The Viterbi algorithm has the desirable properties of not only very efficient computation, but also a fixed amount of computation per symbol and an orderly flow of computation. The low complexity of decoding permits soft decision decoding, where the minimum Euclidean distance between the received sequence and all allowed sequences, rather than Hamming distance, is used to form decisions.

The Viterbi algorithm is a general dynamic programming approach for solving problems such as finding shortest paths. For sequence estimation as in convolutional decoding, the algorithm finds the transmitted sequence whose Euclidean distance, or equivalently whose accumulated square distance, is closest to the received sequence. The operation consists of finding the path of states, where a state denotes

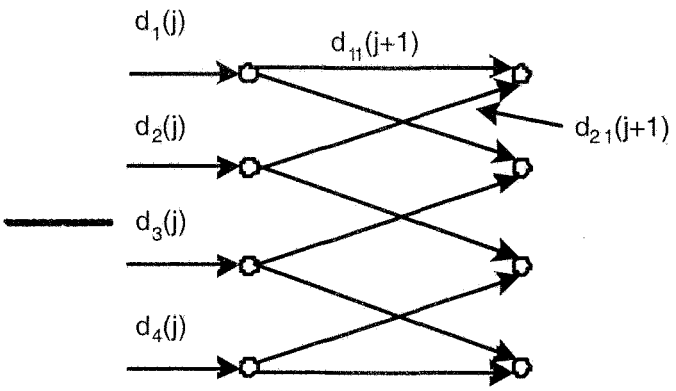


Figure 7.8: Viterbi algorithm applied to a simple convolutional code

the past history of the sequence over the constraint length as described above.

Figure 7.8 illustrates an example of a Viterbi algorithm applied to a simple code with $m = 2$, or $s = 4$ states. Because of the structure of the code, only some of the transitions between adjacent states are possible. These transitions are determined by the input bit sequence. The allowed sequences are described by the sequence diagram. Each transition corresponds to a received sequence of n bits, for which a tentative decision on one bit is performed.

The Viterbi algorithm at any instant keeps track of s survivor sequences up to that time, one terminating in each state, and an error metric associated with each such sequence. For a Gaussian channel, an optimum metric is the accumulated square of the distance between all of the bits in that sequence and the received sequence. One of those sequences is assumed to be correct, but which one is undecided.

When a group of n bits corresponding to one information bit is received, each survivor sequence is extended according to the added

metrics associated with the new input. Each possible sequence terminating in each new state is evaluated, and all sequences other than the one with minimum metric is discarded.

The figure shows how the survivor sequence for state 1 is chosen. The same process is performed for each state. For $r(k)$ the received signal over time k and $t(k)$ the transmitted symbols, then

$$d(j) = \sum_{i=0}^{j} |r(i) - t(i)|^2 \tag{7.1}$$

$$d_{pq}(j) = |r(j) - t_{pq}(j)|^2 \tag{7.2}$$

where $t_{pq}(j)$ is the symbol sequence associated with the transition from state p to state q. The values of r and t may be complex. The survivor sequence chosen for state 1 at time $j + 1$ is the one whose accumulated metric is

$$d_1(j + 1) = min\{d_p(j) + d_{pq}(j)\} \tag{7.3}$$

Only allowed paths pq are examined. Then at time $j + 1$ we again have one survivor sequence terminating in each state, each one being an extension of a previous survivor sequence. It is necessary to make a decision at each step. In almost all cases, if we project far enough into the past, all survivor sequences will have a common beginning. In practice going back $N = 6(m + 1)$ is sufficient. So at any time the stored survivor sequences are N information symbols long, one past information symbol is read out, and the delay through the decoder is N.

The squared error is optimum for a Gaussian channel, but a suboptimum metric for a fading channel. Nevertheless it is frequently used because of its simplicity. For optimum performance over a fading channel, the signal-to-noise ratio must somehow be measured for each received symbol. This measure is often referred to as the "channel state information". The squared error metric is then weighted for

each symbol by this measure. Note that this is same form of weighting that is used in maximum ratio combining when diversity is employed. An intermediate approach is similar to the use of erasure symbols described previously. In this case the squared error is used whenever the signal-to-noise is above a properly chosen threshold. When the measured signal-to-noise is lower than the threshold, zero weighting is applied to the symbol, so that the accumulated metrics are not increased. Most of the loss due to the original sub-optimum metric is recovered.

For optimum operation of Viterbi algorithm, the received inputs should be independent. This is approximated by sufficient interleaving, where the interleaving must effectively eliminate correlation in both time and frequency. For a sufficiently long constraint length and interleaving, a Rayleigh channel is converted to an approximately Gaussian one [12, 91]. The interleaving may be considered to be a form of diversity. Figure 7.9 illustrates the performance of an interleaved convolutional code over a Rayleigh channel. The interleaving in effect converts the channel to a Gaussian one, with a far lower error rate and steeper curve. The coding further improves and steepens the performance curve, in a manner previously sketched in Figure 7.3. This curve assumes ideal decoding using channel state information. For sub-optimum decoding, the curve will be less steep.

It is possible for the input and output of the convolutional encoder to be a continuous stream, with start-up only upon system initialization. However, it is more common for the code to restart and terminate over some interval. That interval could in fact correspond to an OFDM symbol.

In this case the encoder and decoder start in a given state and are forced to terminate in a given state. This allows the Viterbi decoder to eliminate all survivor states but one at the end and to read out the remainder of the stored detected sequence.

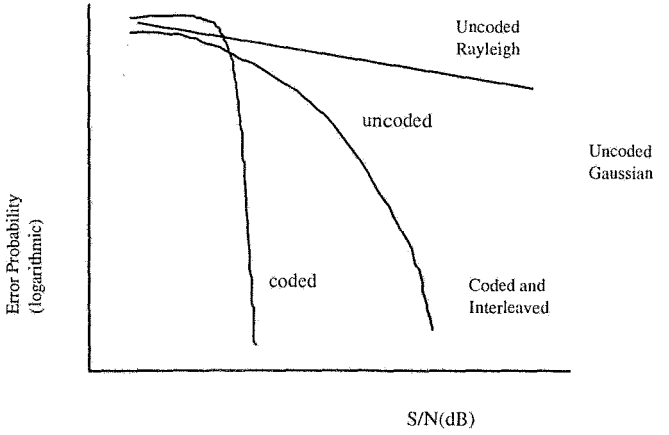


Figure 7.9: Performance of a convolutional code over a Rayleigh fading channel

7.4 Concatenated Coding

Combining convolutional and block codes in a concatenated code is a particularly powerful technique. The block code is the outer code, that is it is applied first at the transmitter and last at the receiver. The inner convolutional code is very effective at reducing the error probability, particularly when soft decision decoding is employed. This operation is shown in Figure 7.10. However when a convolutional code does make an error, it appears as a large burst. This occurs when the Viterbi algorithm chooses a wrong sequence. The outer block code, especially an interleaved Reed-Solomon code, is then very effective in correcting that burst error. For maximum effectiveness the two codes should be interleaved, with different interleaving patterns. Both interleaving patterns should account for both time and frequency correlation. Reference [12] evaluates the performance of several combinations of convolutional and Reed-Solomon codes for an OFDM system operating

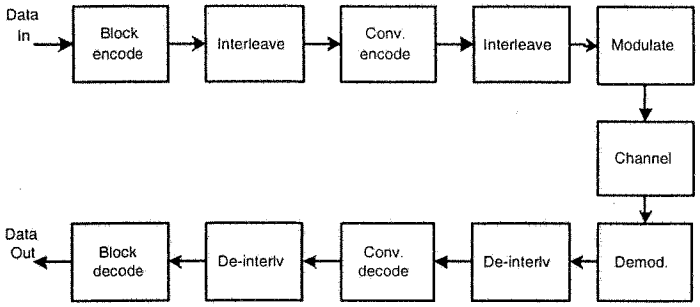


Figure 7.10: Concatenated coding with interleaving

over a fading channel.

7.5 Trellis Coding in OFDM

Trellis coding is similar in spirit to convolutional coding, but differs in the very important aspect that the coding is imbedded in the modulation rather than existing as a separable process [92, 93]. Soft decision decoding is based on minimum Euclidean distance of sequences, and is part of the demodulation procedure.

This is consistent with the basic philosophy of information theory. In Shannon's original paper, capacity is achieved by assigning a different waveform for each complete message, and the length of the message is allowed to approach infinity. Nowhere is the message described as a sequence of bits or symbols. The receiver chooses the message as the one closest to the received waveform. Trellis coding is a practical approximation to this theoretical scheme.

Rather than adding redundant bits or symbols prior to modulation, redundancy is introduced by using constellations with more points than would be required without coding. Usually the number of points

is doubled. The number of symbols per second is unchanged, therefore the bandwidth required is also unchanged. Since there are more possible points per symbol, it may appear that the error probability for a given signal-to-noise ratio would increase. However, as in convolutional coding, dependencies are introduced among the different symbols in that only certain sequences of constellation points are allowed. By properly making use of these constraints in the receiver, as in convolutional decoding, the error probability actually decreases. A measure of performance improvement is the "coding gain", which is the difference in S/N between a coded and an un-coded system of the same information rate that produces a given error probability.

The first step in designing a trellis code is to form an expanded constellation and to partition it into subsets. The points within each subset are made far apart in Euclidean distance, and will correspond to un-coded bits. The remaining, or coded bits, determine the choice of subset. Only certain sequences of subsets are permitted, those sequences being determined by a simple convolutional code. As in convolutional coding, the sequence is best described by a state transition diagram or "trellis". In order to keep allowed sequences far apart, the constellation subsets are chosen so that those corresponding to branching in and out of each state have maximum distance separation.

We will illustrate the procedure with a simple trellis code that uses a 16-point QAM constellation to carry 3 bits per symbol. The technique is also applicable to phase modulation, where the number of phases is increased. Shown in Figure 7.11 is a partition of the constellation into 4 subsets of 4 points each, where the labelling of points denotes subset membership.

Figure 7.12 shows the mapping of 3-bit input symbols into constellation points. Two of the input bits, I_2 and I_3, are un-coded and determine which member of a subset is used. The third bit I_1 is fed to a rate-1/2 convolutional encoder whose 2-bit output determines which subset is used.

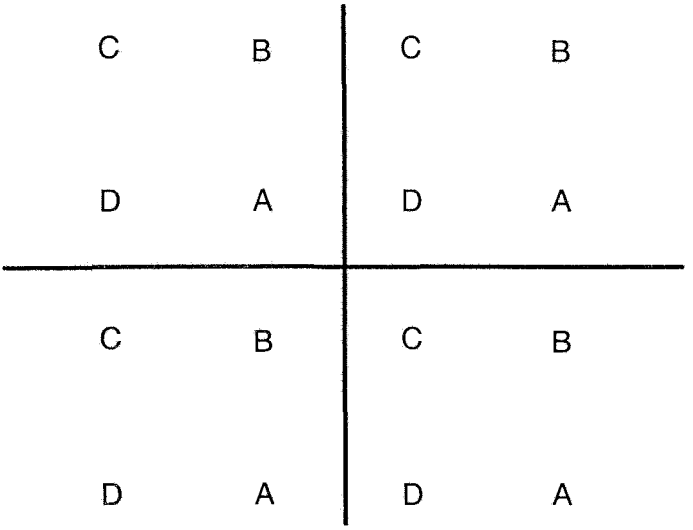


Figure 7.11: An expanded constellation partitioned for Trellis coding

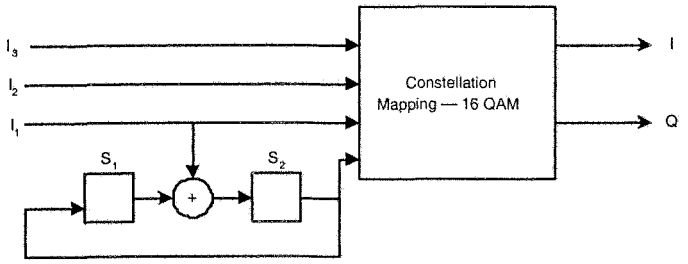


Figure 7.12: Implementation of a Trellis code

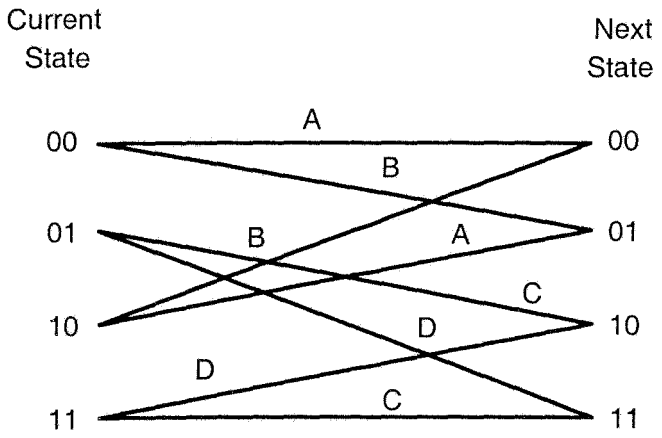


Figure 7.13: State diagram of the Trellis code

Figure 7.13 is the state diagram of the code. The states are labelled according to the contents of the shift register in the convolutional encoder. The labels of the transitions indicate which subset is used when that transition occurs.

At the receiver, a Viterbi algorithm is used for the combined demodulation and decoding. For each received symbol the distance to the nearest member of each subset is measured, as shown in Figure 7.14. The square of this value serves as the metric in extending the survivor states of the Viterbi algorithm. For each survivor state, not only must the coded bits and accumulated squared distance be stored, but also the un-coded bits corresponding to which member of the subset was the nearest point for each symbol.

The code described above operates on 2-dimensional symbols. Doubling the number of points reduces the distance between points by a factor of $1/\sqrt{2}$, so an initial loss of 3 dB (for a large constellation) must be more than compensated for. Performance can be improved by

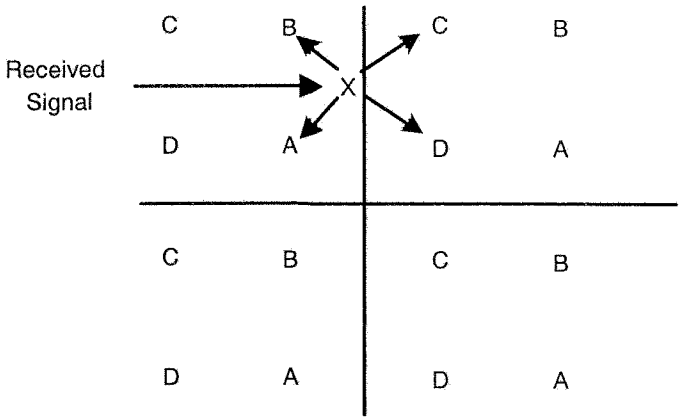


Figure 7.14: Metrics used for Viterbi decoding

operating on symbols of more than 2-dimensions. For example, pairs of QAM symbols may be treated as 4-dimensional symbols. By doubling the number of points in 4- dimensions, spacing between points now reduced by $1/\sqrt[4]{2}$, for an initial loss of 1.5 dB.

The error probability curve over a Gaussian channel is approximately parallel to the un-coded curve, shifted by the coding gain. Coding gains of up to 6 dB are theoretically possible, with up to 5 dB achieved by practical trellis codes. The very simple 4-state, 2-dimensional code described above provides approximately 3 dB coding gain. If for the same constellation we had used an 8-state code with 1 un-coded bit and 2 coded bits, then the coding gain would be increased to approximately 4 dB.

In OFDM it is the usual practice to perform trellis coding over the sub-carriers of a single OFDM symbol, rather than extending memory over a greater interval of time. At the beginning of each OFDM symbol, the code is started in a known state. The code is forced to a known state at the end of an OFDM symbol. Any null sub-carriers are

skipped. Variable size constellations are readily handled by using a fixed number of coded bits on each sub-carrier and a variable number of un-coded bits.

7.6 Turbo Coding in OFDM

Turbo coding has recently been used in many communication systems successfully. Random-like structure of turbo codes have resulted in outstanding performance by providing small error rates at information rates of close to theoretical channel capacity. In this section, we briefly review some fundamental features of turbo coding and explain why it is a good potential candidate for OFDM systems.

In general a turbo code comprises two or more concatenated or parallel codes. A typical turbo code includes parallel concatenated convolutional codes where the information bits are coded by two or more recursive systematic convolutional codes each applied to permutations of the information sequence as in Figure 7.15. However, the same iterative turbo decoding principle can be applied to serially and hybrid concatenated codes [94]. The high error correction power of turbo coders originates from random like coding achieved by random interleaving in conjunction with concatenated coding and iterative decoding using (almost) uncorrelated extrinsic information.

The structure and complexity of turbo encoder design is restricted by other system parameters. Here, we briefly review some of the critical issues which directly impact the code structure. *Decoding Delay* is critical for receiver design. Since the decoding structure of turbo codes is iterative and includes an interleaving /de-interleaving block for each iteration, the excess delay may interfere with overall system performance. For example, speech frames should be processed within a tight time frame in most applications. Another important system design issue is required *coding gain*. While the BER improvement of turbo coders is outstanding, a system designer should be aware of

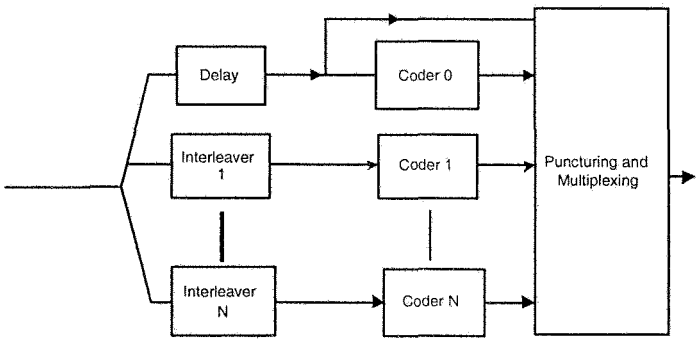


Figure 7.15: A typical turbo encoder

the consequences of low SNR system functionality. It is tempting to design the system for a low SNR condition assuming that a powerful turbo code can meet the target BER performance. However, many other receiver functions such as synchronization and adaptive algorithms require a minimum SNR which should be taken into account carefully.

The decoding process is iterative and in this overview we briefly explain two common iterative techniques: Maximum a *posteriori* Probability (MAP) technique and Soft input Soft Output Viterbi Algorithm (SOVA). Three different types of soft inputs are available for each decoder: the un-coded information symbols, the redundant information resulting from first Recursive Symmetric Code (RSC) and a *priori* (extrinsic) information. The output is a weighted version of un-coded information, a priori information and new extrinsic information.

In general, a symbol-by-symbol MAP algorithm is optimal for state estimation of a Markov process. MAP algorithms for turbo decoding calculate the logarithm of the ratio of a *posteriori* probability (APP) of each information bit being one to the a *posteriori* probability of the bit being zero. The MAP technique is complicated and

requires non-linear operations which makes it less attractive for practical purposes. In comparison, a simplification of MAP algorithm, namely, Soft Output Viterbi Algorithm (SOVA) leads to a practical sub-optimum technique. While performance is slightly inferior to an optimal MAP algorithm, the complexity is significantly less. MAP takes into account all paths by splitting them into two sets, namely, the path that has an information bit one at a particular step and paths which have bit zero at that step and returns the log likelihood ratio of the two. SOVA considers only the survivor path of the Viterbi algorithm. Therefore, only the survived competing path which joins the path chosen in the Viterbi algorithm is taken into account for reliability estimation.

While the encoders have a parallel structure, the decoders operate in serial which results in an asymmetric overall structure. For example, in the first iteration the first decoder has no a *priori* information [96].

As discussed before, coding schemes for OFDM systems take advantage of coding in both time and frequency in conjunction with proper interleaving. This provides extra protection against time and frequency selective fading. Use of turbo coding for OFDM systems is seriously being considered for several wireless standards. The depth of interleaving can be restricted to one OFDM block which provides correlation between symbols modulated by different sub-carriers and subjected to independent selective fading.

Another coding scheme supports inter-block interleaving which can be intelligently designed to provide time and frequency correlation. While the former technique does not impose extra delay at the receiver, the latter technique gives another degree of freedom to the system designer. The choice of proper coding scheme depends on system requirements and applications.

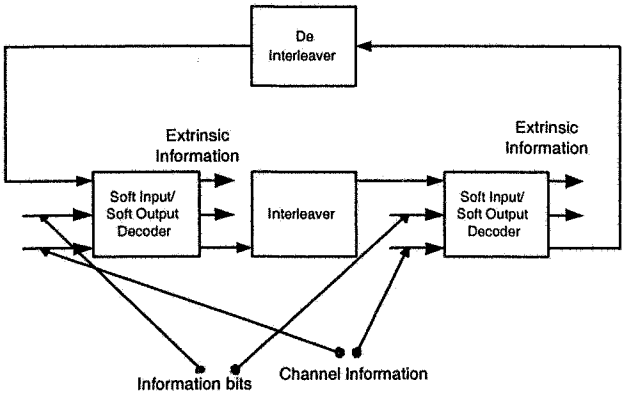


Figure 7.16: A turbo decoder structure

Chapter 8

ADSL

8.1 Wired Access to High Rate Digital Services

Ubiquitous communication channel is the subscriber line, or loop, consisting of an unshielded twisted pair of wires, connecting any home or office to a telephone company's central office. The overwhelming majority of these channels are used to carry analog voice conversations, which require a bandwidth of less than 4 KHz. It has long been recognized that most subscriber lines can support a much wider bandwidth, in particular, to carry high rate digital signals. The first such widespread use is for access to basic rate ISDN, in which the subscriber line carries 160 *kbps* simultaneously in both directions over a single pair.

More recently, higher rates have been introduced into numerous systems. Of particular interest here is ADSL which is primarily intended to provide access for residential applications [97, 98]. Most of such applications require a high data rate in the downstream direction (to the customer) and a much lower rate upstream (from the

customer). The primary applications of ADSL are the delivery of digitally encoded video, and access to digital services, particularly the Internet. ADSL meets these needs by providing a high rate digital downstream signal of over 1 $Mbps$, a moderate rate upstream signal, and a normal analog voice channel, all over a single wire-pair. Because virtually all customers have a wire-pair channel providing voice service, no additional channel need be installed to provide this new service. It only requires the installation of terminating equipment at the customers' premises and at the central offices. OFDM, typically referred to as DMT (Discrete Multi-Tone) in this application, has been adopted as the standard for transmission of the digital information. Several varieties of ADSL exist, differing primarily in the bit rates carried and the set of subscriber lines over which satisfactory performance is achieved.

8.2 Properties of the Wire-Pair Channel

Before discussing the transmission techniques, it is essential to first describe the channel - both its transmission properties and the types of interference to which it is subject. As with any communication channel, a combination of both factors must be considered in determining the performance that can be achieved.

The connection between the subscriber and the central office consists of several sections, each consisting of a wire-pair bundled in a cable with many other wire-pairs [99, 100, 101]. Insulation on modern cables is polyethylene, although older paper pulp cable is still present in much of the plant. The sections nearest the central office are referred to as feeder, where the cable size is large but the wire gauge may be fine. Closer to the customer, the distribution plant consists of smaller cables, but typically heavier gauge wire. Finally, the individual drop wire provides the final connection to the subscriber. Wiring on premises may also be significant. Typical wire gauge size through-

out is 24-gauge (0.5 mm) or 26-gauge (0.4 mm). The overall length of the circuit can be up to 5.5 km. If it is greater than this distance, loading coils are usually added which improve transmission in the voice-band but greatly restrict the bandwidth above 4 KHz. Loaded circuits therefore cannot be used to provide any high rate digital service. Fortunately, such circuits are rare and exist only in rural areas.

A trend is now taking place to introduce carrier systems and remote switches into subscriber plants. Such electronic equipment is connected to the central office digitally, often using optical fiber. ADSL in this case, need only to provide access from the subscriber to the remote electronic device rather than all the way to the central office. A plan, referred to as the Carrier Serving Area (CSA), envisions almost all future subscribers within 3.7 km of either a central office or of remote equipment if the cable used for the whole path is 0.5 mm or larger, or 2.7 km if 0.4 mm wire is used. An additional benefit of this plan is the elimination of loaded circuits.

The transmission properties of a wire-pair are determined by its primary constants, the resistance R, inductance L, capacitance C, and conductance G per unit length. At low frequencies, R is constant, but increases proportional $\sqrt{f}$ to at high frequencies (above about 200 KHz), due to skin effect. C is very constant with frequency, and L declines slightly with frequency. The conductance increases with frequency, but is small enough for modern cables that it may be neglected over the entire frequency range. Ignoring mismatched impedance effects, the transfer function of a line of length d is

$$H(f) = e^{-\gamma(f)d} \tag{8.1}$$

where the complex propagation constant γ is given by

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{8.2}$$

σ is the loss in nepers per unit length, and β the phase shift in radians per unit length. The phase distortion can, in general, be equalized, at least in principle. So, the loss has the primary effect on performance.

However, it should be noted that the phase distortion, that is the deviation from linear phase vs. frequency (flat delay), is very severe at low frequencies. At low frequencies

$$\alpha \approx \sqrt{\frac{\omega RC}{2}} \tag{8.3}$$

and is therefore proportional to $\sqrt{f}$. At high frequencies

$$\alpha \approx \frac{R}{2}\sqrt{\frac{C}{L}} \tag{8.4}$$

so the loss is also proportional to $\sqrt{f}$ because of the skin effect on R and the relative constancy of L.

In between, roughly in the range of 20 - 200 KHz, the loss increases more slowly with frequency. Figure 8.1 is a plot of loss vs. frequency for a typical cable pair. Note the slope of $\frac{1}{2}$ at the upper and lower ends, and the lower slope in the middle range.

A further consideration affecting the transfer function of a wire-pair is mismatched termination which results in reflection. Impedance matching over all frequencies is difficult, particularly at low frequencies where the characteristic impedance Z_0 of the line is highly variable, becoming infinite as $f \rightarrow 0$.

Another effect to be considered is the presence of bridged taps, which exist in many parts of the world, including the USA, but are not present in some other countries. Bridged taps are unused sections of line attached to the line of interest, resulting from the telephone company's loop installation and removal policies. A bridged tap has an open circuit at its end, and therefore has no effect at low frequency such as the voice band. However at high frequency, in particular where the length of the tap is near $\frac{1}{4}$ wavelength, the tap presents a low impedance shunt and reduces the transfer function. In addition, reflection from the end of the tap is a source of echo in both directions. Figure 8.2 illustrates a bridged tap and its effect.

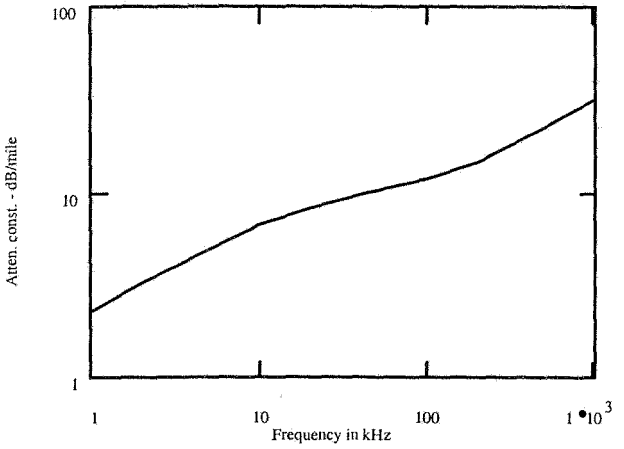


Figure 8.1: Attenuation constant of 24-gauge (0.5 mm) wire-pair

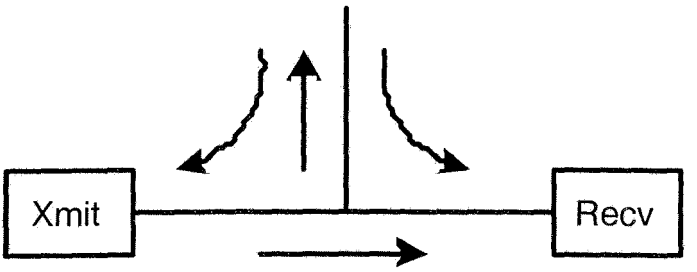


Figure 8.2: A bridged tap and its echoes

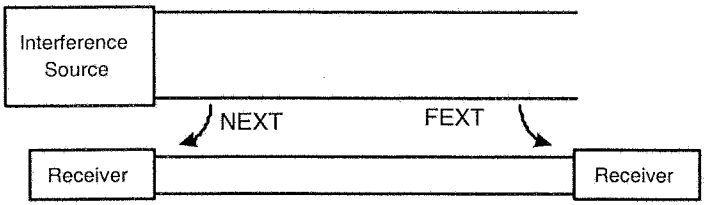


Figure 8.3: Crosstalk mechanism

Overall, unlike a radio or voice-band channel, the wire-pair channel is not sharply bandlimited but has a loss that increases gradually with frequency. This brings up the non-trivial question of the effective bandwidth of such a channel. It would be a mistake to consider only the loss in answering this question. The usable bandwidth depends equally on the noise present, because if the noise is low the usable bandwidth can be quite high, even though the loss at high frequency is high.

The primary source of noise is crosstalk from signals on other pairs in the cable [102, 103]. As illustrated in Figure 8.3, near-end crosstalk (NEXT) results from transmit sources co-located or close to the receiver in question, while far-end crosstalk (FEXT) is caused by sources at the other end of the channel. When both effects are present, NEXT is more severe because of the relative signal levels. NEXT may expected to be a more serious problem at the central office rather than at the customer end because many co-located signals are more likely to occur.

Crosstalk results both from capacitive and inductive coupling between wire-pairs in the same cable. Figure 8.4 illustrates the capacitive coupling mechanism.

Shown in Figure 8.4 is a differential length of cable in which a signal on the upper pair of crosstalk into the lower pair. The four capacitances between wires of one pair and wires of the other pair are

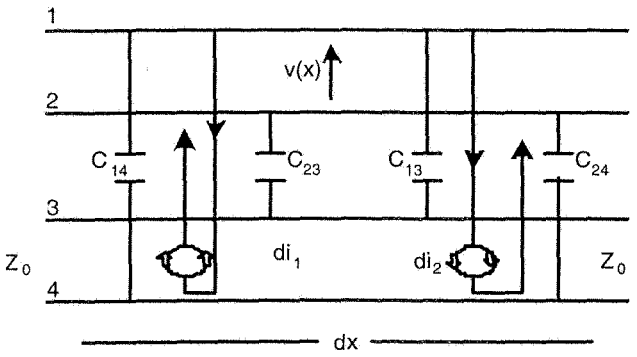


Figure 8.4: Capacitive crosstalk mechanism

shown. As shown, there are two current paths of opposite sign.

The net differential current is

$$\begin{aligned} di_c &= di_1 - di_2 \\ &= \frac{V(x)}{\frac{1}{j\omega C_{13}dx} + \frac{1}{j\omega C_{24}dx}} - \frac{V(x)}{\frac{1}{j\omega C_{14}dx} + \frac{1}{j\omega C_{23}dx}} \\ &= j\omega V(x)\left[\frac{C_{13}C_{24}}{C_{13}+C_{24}} - \frac{C_{14}C_{23}}{C_{14}+C_{23}}\right]dx \end{aligned} \tag{8.5}$$

If the differences between capacitances are much less than the capacitances themselves, as is almost always true, then the above quantity may be approximated by

$$di_C = \frac{j\omega C_u}{4} V(x)dx \tag{8.6}$$

where C_u is an unbalanced capacitance given by

$$C_u = C_{13} - C_{14} + C_{24} - C_{23} \tag{8.7}$$

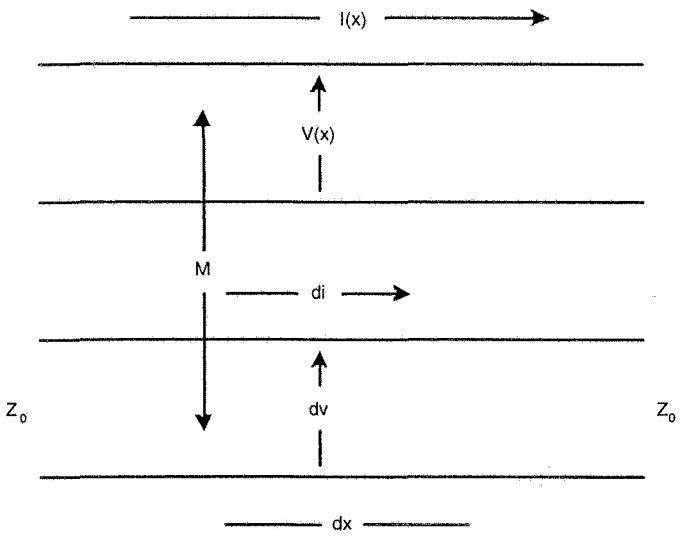


Figure 8.5: Inductive crosstalk mechanism

This quantity can be either positive or negative, and varies along the line, particularly as the relative twist between the two pairs changes. Over the length of the line, d, it will be zero mean. The induced incremental voltage is

$$d\nu_c = \frac{Z_0}{2} di_c = \frac{j\omega V(x) C_u Z_0}{8} dx \tag{8.8}$$

where it is assumed that the line is terminated in its characteristic impedance in both directions.

The inductive coupling may be treated similarly, as shown in Figure 8.5.

Over the differential length, the mutual inductance is $M dx$. Then

the differential induced voltage is

$$d\nu_m = j\omega I(x)Mdx, \text{ and } di_m = \frac{j\omega I(x)Mdx}{2Z_0} \tag{8.9}$$

The near-end and far-end currents are

$$di_n = di_c + di_m, \text{ and } di_f = di_c - di_m \tag{8.10}$$

These can be expressed as

$$di_n = j\omega Y_n dx, \text{ and } di_f = j\omega Y_f dx \tag{8.11}$$

where Y_n and Y_f are unbalanced functions given by

$$\frac{C_u Z_0}{8} \pm \frac{M}{2Z_0} \tag{8.12}$$

They are both zero mean with the same variance, and non-correlated over distances comparable to a twist length. Except at low frequency where Z_0 is highly variable, they are approximately constant with frequency.

We are now ready to evaluate overall crosstalk. Figure 8.6 illustrates the case of NEXT.

If the interfering source current is denoted by $I(0)$, then at a distance of x it has been attenuated to $I(0)e^{-\gamma x}$. The incremental induced crosstalk current is then

$$di_n(x) = j\omega Y_n(x)I(0)e^{-\gamma x}dx \tag{8.13}$$

and similarly, the incremental crosstalk voltage is

$$d\nu_n(x) = j\omega Y_n(x)V(0)e^{-\gamma x}dx. \tag{8.14}$$

When propagated back to the input, the contribution of this crosstalk component is

$$dx_n = d\nu_n(x)e^{-\gamma x}dx = j\omega Y_n(x)V(0)e^{-2\gamma x}dx \tag{8.15}$$

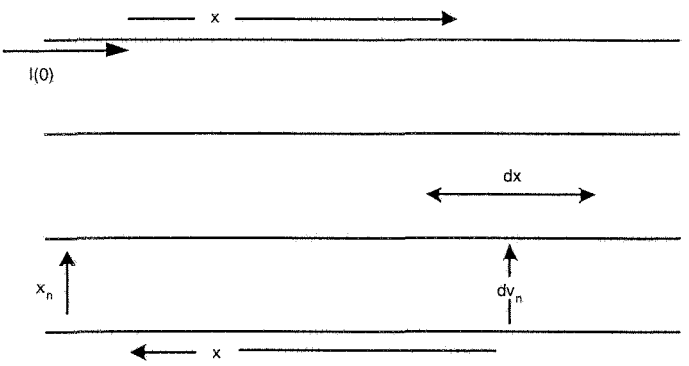


Figure 8.6: Near-end crosstalk generation

and the total NEXT voltage is

$$x_n = j\omega V(0) \int_0^d Y_n(x) e^{-2\gamma x} dx. \tag{8.16}$$

This quantity is zero mean, with variance

$$\bar{x_n^2} = \omega^2 V^2(0) \int_0^d \int_0^d Y_n(x) Y_n(y) e^{-2\gamma x} e^{-2\gamma^* x} dx dy \tag{8.17}$$

Because Y_n is non-correlated over short distances, we can approximate

$$< Y_n(x) Y_n(y) >= k_n \delta(x - y) \tag{8.18}$$

so that

$$\bar{x_n^2} = \omega^2 V^2(0) k_n \int_0^d e^{-4\alpha x} dx = \frac{\omega^2 V^2(0) k_n}{4\alpha} (1 - e^{-4\alpha d}) \tag{8.19}$$

For any but the shortest lines, $e^{-4\alpha d} << 1$. We will further approximate $\alpha \approx k\sqrt{f}$. Then, lumping together the various constants,

$$\bar{x_n^2} \approx V^2(0) \lambda_n f^{3/2}. \tag{8.20}$$

We now have the important result that the NEXT power transfer function,

$$T_n(f) = \frac{\bar{x_n^2}}{V^2(0)} \tag{8.21}$$

is proportional to $f^{3/2}$, and is independent of the line length except for very short lines. The level of the transfer function is determined by the constant λ_n. It can also be specified by giving its value at some frequency f_0. Then at any other frequency,

$$T_n(f) = T_n(f_0) + 15 \log_{10} \frac{f}{f_0} (in\ dB). \tag{8.22}$$

We can analyze the far-end crosstalk in a similar manner to obtain

$$T_f(f) = \frac{\bar{x_f^2}}{V^2(0)} = \lambda_f f^2 d e^{-2\alpha(f)d} \tag{8.23}$$

The exponential factor arises from the fact that all crosstalk components travel a total distance of d, partly on the interfering pair, and the remainder on the interfered pair. When both NEXT and FEXT occur, the former dominates at all frequencies of interest.

It is usual to model various system imperfections, such as finite precision digital implementations, as an added white Gaussian noise, with spectral density N_0. This is usually a reasonably accurate model. In addition, it eliminates singularities in signal-to-noise ratios in the presence of only crosstalk. It should be noted that at the relatively high receiver power levels in wire-pair systems, actual thermal noise is so far below the signal that it becomes completely negligible.

We are now ready to express the signal-to-noise ratio, which is the principal determinant of performance. Let the transmitted signal's power spectral density be given as $G(f)$, and each of k interferers as $G_k(f)$. The desired channel power transfer function is $|H(f)|^2$, and that of the crosstalk paths is $T_k(f)$. Then

$$SNR(f) = \frac{G(f)|H(f)|^2}{\sum_k G_k(f)T_k(f)} + N_0 \tag{8.24}$$

In an OFDM system the above quantity is an excellent approximation for the SNR of each sub-channel, where f is center frequency of the sub-channel.

The ideal capacity of the channel is

$$C = \int_0^\infty \log_2[1 + SNR(f)]df \tag{8.25}$$

A useful definition of the usable bandwidth of a wire-pair channel is the frequency range over which the SNR is greater than some value. For coded single carrier systems, a value of 0 dB is a reasonable one. This is also the range that typically includes over 95% of the ideal capacity. For an OFDM system, for each sub-carrier to carry at least a four-point constellation, the SNR should be at least approximately 10 dB.

8.3 ADSL Systems

Two classes of ADSL have been standardized recently, with many options in each. Full rate ADSL can carry up to approximately 8 Mbps downstream and 800 kbps upstream. A simpler class, commonly called "ADSL Lite", carries up to approximately 1.5 Mbps downstream and 500 kbps upstream. In both cases, data rates can be adjusted to any value in steps of 32 kbps. An analog voice channel is provided on the same pair. The target error probability is per bit, with some required margin [104, 105].

The two classes are somewhat compatible with each other. In both cases, sub-carriers are spaced 4312.5 Hz apart in both directions. After

every 68 frames of data, a synchronization frame is inserted. Because of this and the use of cyclic prefixes, the net useful number of data frames is 4000 per second in all cases. One of the sub-carriers of the frame is devoted to synchronization. Adaptive bit allocation over the sub-carriers is performed in all cases. This process is critical to ensure system performance.

In the full rate downstream direction, a block of 255 complex data symbols, including several of value zero, are assembled. These will correspond to sub-channels 1 to 255. The lower ones cannot be used because of the analog voice channels, nor can the 255th. Therefore, the highest frequency allowed sub-carrier is centered at 1.095 MHz. Sub-carriers which cannot support at least a 4-point constellation at the desired error probability will also be unused. Conjugate appending is performed on the block, followed by a 512-point DFT. This results in frame of 512 real values. A cyclic prefix of 32 samples is added, and the resultant 2.208 M samples per second transmitted over the line.

Upstream, 31 sub-channels are processed, although (again) the lower few and the 31st cannot be used. The same processing is performed with a cyclic prefix of 4 samples. The upstream and downstream sub-channels may overlap. This provides a larger data rate, but requires the use of echo cancellation.

The bit streams may be treated as several multiplexed data channels. Each such channel may be optionally Reed-Solomon coded, with a choice of code and interleaving depth. Other optional codes include a CRC error check, and a 16-state 4-dimensional trellis code. The trellis code, when present, operates over the non-zero sub-carriers of a block, and is forced to terminate at the end of each block.

"ADSL Lite" is intended as a simpler, lower cost system, with greater range of coverage because of the lower rate. One important difference is the elimination of filters at the customer's premises to separate voice and data channels. The upstream channel is created identically to that of the full rate system, except that the first 6 sub-

carriers must be at zero.

The downstream transmitted sampled rate is reduced by a factor of two, to 1.104 M samples per second. The IDFT is performed over an initial block of 127 complex numbers, of which the first 32 must be zero. The highest sub-carrier is now at 543 KHz. In this case, the upstream and downstream sub-carriers do not overlap. The signal is treated as a single bit stream. Reed-Solomon and CRC coding are again optional, but there is no trellis coding.

As of the time of publication of this text, ADSL systems of various rates are being installed around the world, with several million customers. The ultimate market for high rate digital residential services is not clear, nor are the relative advantages and disadvantages of providing these services over competitive cable television facilities.

Chapter 9

Wireless LAN Applications

9.1 Introduction

Wireless Local Area Networking (WLAN) has quickly become a significant solution for high speed connectivity and for the proliferation of mobile computers. WLAN replaces the last hop with a wireless link and connects the devices to the IP core. WLAN tries to offer mobility and flexibility in addition to the same features and benefits of traditional wired LANs. Much interest has been involved in WLAN due to its efficiency, lower cost, mobility, and flexibility.

9.1.1 Background

With the interest in WLAN products after 1990, IEEE (Institute of Electrical and Electronics Engineers) decided to standardize the efforts on WLAN. These efforts meant to replace the products based on proprietary technologies. In 1997, 802.11 is introduced as the first WLAN standard by IEEE. The three different technologies within that standard led to IEEE introduce two supplements: 802.11b and 802.11a

in 1999.

The European competitor of IEEE 802.11 is HIPERLAN/1 in the level of IEEE 802.11b. It was introduced in 1996 by European Telecommunication Standard Institute (ETSI) under the project named Broadband Radio Access Network (BRAN). ETSI's next standard standardized in 2000 is called HIPERLAN/2 to compete with IEEE 802.11a. Multimedia Mobile Access Communication (MMAC) System is the Japanese proposal for WLAN and MMAC introduced High Speed Wireless Access Network (HiSWAN) in 1999. The Table 9.1.1 presents a comparison among various WLAN standards.

The medium access control (MAC) layers show difference for 802.11 and HIPERLAN/2. IEEE 802.11a uses a Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) scheme to transmit packets, HIPERLAN/2, on the other hand, uses Time Division Multiple Access (TDMA) scheme.

CSMA/CA is based on contention. Each station contends for access in the same channel. The station attempts to send only after assuring that there is no station that is transmitting. If there is a transmission, stations wait until the channel is idle. The wait times are arbitrary which makes the protocol asynchronous.

TDMA, on the other hand, offers guaranteed medium access by assigning each station a time slot. This regular access allows to implement quality of service for voice and data. The quality of service is being implemented in 802.11e MAC protocol on the IEEE side.

9.1.2 Pros and Cons

WLANs are not a replacement for wired LANs but are a type of LAN that comes with mobility and lower cost because of the replacement of the cable with radio waves provides roaming and low cost deployment. Apparently, mobility, low cost, fast deployment, and scalability are the advantages. On the other hand, radio waves have the disadvan-

	Ratifi-cation	RF band	Maximum Data rate	Physical Layer	Range
IEEE 802.11	June 1997	2.4GHz	2Mbps	FHSS DSSS,IR	50-100m
IEEE 802.11b	Sept. 1999	2.4GHz	11Mbps	DSSS /CCK	50-100m
IEEE 802.11a	Sept. 1999	5GHz	54Mbps	OFDM	50-100m
IEEE 802.11g	June 2003	2.4GHz	54Mbps	OFDM PBCC	50-100m
ETSI HIPERLAN/1	early 1996	5GHz	23.5Mbps	GMSK	50m
ETSI HIPERLAN/2	Feb. 2000	5GHz	54Mbps	OFDM	50-300m
MMAC HiSWANa	April 1999	5GHz	27Mbps	OFDM	100-150m

Figure 9.1: Comparison of WLAN standards [124]

	Wireline	WWAN	WLAN
Connectivity	✓		
Performance	✓		✓
Mobility		✓	✓
Flexibility		✓	✓

Table 9.1: Comparison of LANs

tage in terms of being susceptible to interference and of having security deficiency. Table 9.1.1 gives a comparison of wireline, WLAN, and wireless wide area network (WWAN).

9.2 Topology

IEEE 802.11 has two types of network elements: stations (STAs) and access points (APs). A STA is defined as the device that is equipped with a component that can communicate via IEEE 802.11 protocol. A STA is called an AP if it is connected to a wired network and offers infrastructure service to mobile STAs. The networks composed of the combination of these actors are categorized as Independent Basic Service Set (IBSS) or Infrastructure Basic Service Set (Infrastructure BSS).

9.2.1 Independent BSS

This type of network is also referred as a ad-hoc networks. A network is an ad hoc network when there is no infrastructure. In an ad hoc network, stations communicate directly to each other and they spontaneously establish a distributed communication mechanism. As it can be seen in Figure 9.2, the IBSS is not connected to the wired domain. IBSS is initiated if a station is in the vicinity of another. One of the

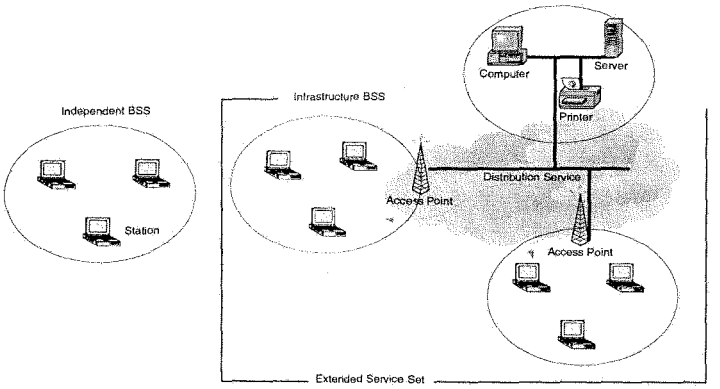


Figure 9.2: IEEE 802.11 networks

STA selects itself as the starter and sends Beacon signals indicating the service id of the IBSS (SSID).

It is rewarding to mention routing in IBSS. The routing algorithm in IBSS is handled by ad-hoc routing protocols which reside in the network layer (See Figure 9.4). In ad hoc networks, there is no fixed routers and each station is capable of functioning as a router which can discover and maintain routes to other nodes in the network.

Ad hoc routing protocols are classified as table driven and source-initiated on-demand routing protocols. A table driven routing protocol forms routes for each pair of stations and maintains up-to-date routing information from each node to every other node in the network with no regard to when and how frequently such routes are desired. The existing and famous table-driven protocols are Destination-Sequenced Distance-Vector Routing (DSDV), Clusterhead Gateway Switch Routing (CGSR), Wireless Routing Protocol (WRP). The differences arise in the number of necessary routing related tables and the methods for notification of changes.

Source-initiated on-demand routing, on the other hand, creates

routes only when it is necessary. When a station has a packet to send, it initiates route discovery protocol to establish the route to the destination. Once the route is constructed, it is maintained as long as it is no longer desired or the destination becomes inaccessible. Ad Hoc On-Demand Distance Vector Routing (AODV), Distance Source Routing (DSR), Temporally Ordered Routing Algorithm (TORA), Associativity Based Routing (ABR), Signal Stability Routing (SSR) are some of the on-demand routing protocols.

The table-driven ad hoc routing relies on the underlying routing table update mechanism since a route is always available regardless of whether or not it is needed. The limitation occurs in the route creation since it incurs significant control traffic and power consumption. In on-demand routing, on the other hand, this control traffic is minimized by initiating route discovery when needed but now the station has to wait until a route can be discovered.

9.2.2 Infrastructure BSS

Contrary to ad hoc networks, Infrastructure BSS contains a gateway to the wired domain and enables a station of a BSS to communicate to a station of other BSSs and other LANs as seen in Figure 9.2. The gateway is called access point (AP) in IEEE 802.11 jargon. APs communicate with each other via either cable or radio to interconnect BSSs. To enable roaming between interconnected BSSs and to create connections to wired network resources, the IEEE 802.11 standard specifies a distribution system (DS). The interconnected BSSs and DS allow to extend the IEEE 802.11 network as large as it is desired. This is called extended service set (ESS) in IEEE 802.11. There is a handover support in the MAC layer in order to provide seamless connectivity within the ESS. Each ESS is identified by its ESS id (ESSID). The DS resides above the MAC layer, which makes it homogeneous and be implemented by any other backbone network such as ethernet, token ring, optical or wireless network. The ESS communication in

backbone network is in layer 2.

It is useful to note the routing in infrastructure BSS. Basic operation requires all nodes to be within the vicinity of the access point. Transmission takes place between the access point and stations in a single hop. On the other hand, multihop routing, is defined as the direct communication between mobile stations and is left open in the standard. Although it will bring additional control mechanisms, multihop routing increases throughput along with a decrease in transmission power.

Let's look at the fact that decrease in range corresponds to increase in throughput intuitively. Define a disk of unit area D with n nodes. If $|X_i - X_j| < r(n)$ then station i transmits to station j successfully where $r(n)$ is defined as common range and X_i is the position of station i. For every other node k, if $|X_k - X_j| > (1+\triangle)r(n)$ then station k transmits simultaneously. The distance between receivers must be at least $\triangle r(n)$ from the triangle inequality.

Let's determine the maximum number of transmissions (Max_{Tx}). The disks of radius $\triangle r(n)/2$ centered at each receiver are disjoint and each receiver disk means one transmission, thereby the maximum number of transmission are reciprocal of the minimum area of a receiver disk which is (Min_{Area}).

Min_{Area} inside D is achieved when receivers are located at the boundary. Thus, we can approximate as follows;

$$Min_{Area} = \frac{1}{4}\frac{\pi\triangle^2 r(n)^2}{4} \tag{9.1}$$

and Max_{Tx} becomes $\frac{16}{\pi\triangle^2 r(n)^2}$. Total number of *bps* that should be served is T and

$$T = n\lambda(n)\frac{L}{r(n)W} \tag{9.2}$$

where L is average distance between the sender and receiver pair. $L/r(n)$ means number of hops. W and $\lambda(n)$ are transmission and data

generation rate in each station respectively. $\lambda(n)$ is supported if

$$T \leq Max_{Tx} \Rightarrow \lambda(n) \leq \frac{c}{nr(n)} \tag{9.3}$$

The effect of range on throughput can be observed from the equation (9.3) since the smaller the range, the higher the throughput [107].

Mobility

In WLAN, stations maintain their connections during their movement. There occur three types of transitions which are defined as follows;

- *No mobility or intra BSS mobility* occurs when the station stays inside the vicinity of one BSS.
- *Inter BSS mobility (intra ESS mobility)* occurs when the station moves from one BSS to another in the same ESS. Station compares the signal to noise ratio (SNR) of each AP and selects one of the APs. In this movement, the seamless roaming is handled by the DS and messages are sent to the new AP from the old one since in association, DS notifies the old AP about the new location of STA.
- *Inter ESS mobility* occurs when stations change their ESS. This transition is not supported by IEEE 802.11. Seamless Roaming can be handled by the upper layers. Mobile IP is one of the solutions for this type of movement. Mobile IP introduces two IP addresses: permanent and temporary. In basic Mobile IP designed based on IPv4, whenever a station changes its ESS, it gets a new temporary IP address and notifies its home agent (HA). HA is the agent located in the home network where the station's permanent IP address belongs. HA records the temporary IP address of the stations. Whenever a message is sent by a correspondent host (CH), any host in the Internet, to a mobile

host (MH). The packet is sent to the network where permanent address belongs. HA is responsible to capture the packets of STA if away and to send them in encapsulation to the foreign agent (FA) to where temporary IP address belongs. FA decapsulates the packet and sends in layer 2 to the MH. IPv6 brings improvement to Mobile IP with its large address space and eliminates the necessity to send the packets to the HA. Once associated, the packets can be sent directly to the MH from the CH.

9.2.3 Services

The IEEE 802.11 does not put constraints on how the DS should be implemented but defines the services. The IEEE 802.11 defines two services: station services (SS) and distribution system services (DSS).

The IEEE 802.11 has three types of messages: data, management, and control. Each of the services has its own special MAC frame type. Logical classification of services are as follows: distribution of messages in a DS, management of DS and reliable access of stations to DS.

1) DSS: Distribution of messages in a DS.

Distribution service defines how the message is distributed within the DS. DS should be capable of determining the destination. DS knows the attached stations by management services (association, reassociation, and disassociation).

Integration If the destination is LAN, integrated to DS. DS sends the message to *portal*. The portal is the gateway between DS and integrated LAN. The Integration function is defined as the procedure which handles the messages from or to the portal.

2) DSS: Management of DS.

Association In delivery of a message within DS, association is introduced in order for the DS to know which AP to access for the STA. A STA is allowed to associate only one AP to provide uniqueness for the destination. A STA initiates association after discovering the AP.

Reassociation A STA initiates reassociation if it moves to a new AP or if it wants to change the association attributes. This way inter BSS mobility is supported.

Disassociation Disassociation service is to terminate the existing association. Disassociation is invoked either by STA or AP and can not be refused.

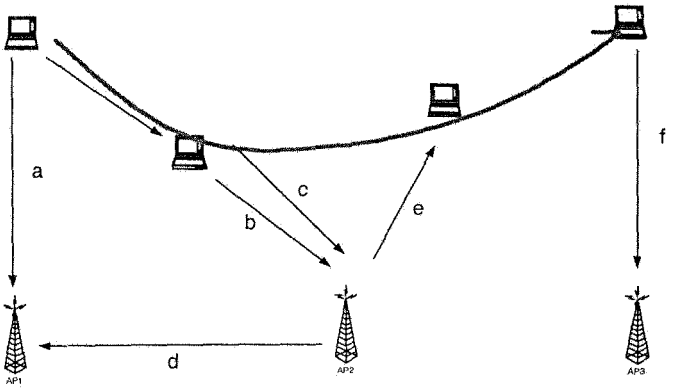
3) SS: Reliable access of stations to DS.

Authentication IEEE 802.11 introduces link-level authentication to bring the wireless link up to physical standard of its wired counterpart. This service is to bring the identities of source and destination reliability. There are two types of authentication schemes; open system authentication where there are no restrictions and any stations can become authenticated with any other, shared key authentication where identity is demonstrated by knowledge of a shared key. Unlike association, a STA can be authenticated with more than one STAs.

Preauthentication Authentication is required before an association and STAs can invoke preauthentication service to save time.

Deauthentication When there is need to terminate the existing authentication, deauthentication service is invoked. Deauthentication is a notification and shall be followed by disassociation.

Privacy Unlike the wired LAN where only physically connected stations hear the traffic, in wireless medium any station which is in the vicinity of the BSS can hear messages of others. This brings a security problem and standard proposes to overcome



(a) ---- The station finds AP1, it will authenticate and associate.

(b) ---- As the station moves, it may pre-authenticate with AP2.

(c) ---- When the association with AP1 is no longer desirable, it may reassociate with AP2.

(d) ---- AP2 notify AP1 of the new location of the station, terminates the previous association with AP1.

(e) ---- At some point, AP2 may be taken out of service. AP2 would disassociate the associated stations.

(f) ---- The station find another access point to authenticate and associate.

Figure 9.3: Relationship between IEEE 802.11 services © [108]

it by encryption. Wired equivalent privacy (WEP) option is to perform encryption of messages in order to prevent eavesdropping. WEP is being replaced by IEEE 802.1x due to its higher security standards.

An example realization of the services is shown in Figure 9.3.

9.3 Architecture

The significant element of the WLAN is its medium access control (MAC) protocol. MAC protocol controls the transmission and lies

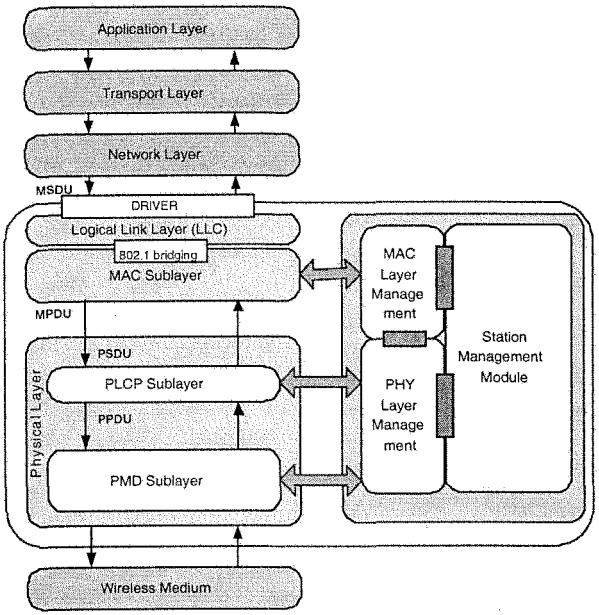


Figure 9.4: IEEE 802.11 architecture © IEEE

above the physical layer and compliant with the any physical layer introduced. The logical architecture is shown in Figure 9.4. The underlying mechanism in MAC layer is carrier sense multiple access (CSMA) scheme like ethernet. It is hard to detect the collision right away in wireless medium. As a result, IEEE 802.11 implements collision avoidance (CSMA/CA) unlike ethernet where collision detection is implemented (CSMA/CD). The physical layer handles the transmission from one station to the other in a reliable way.

A wireless medium introduces noise, multipath fading, and interference. Any transmission that is successfully terminated can be learned from the acknowledgement frame. In basic access mechanism, all transmitted frames are acknowledged to reduce the inherent error rate of the medium, at the expense of additional bandwidth con-

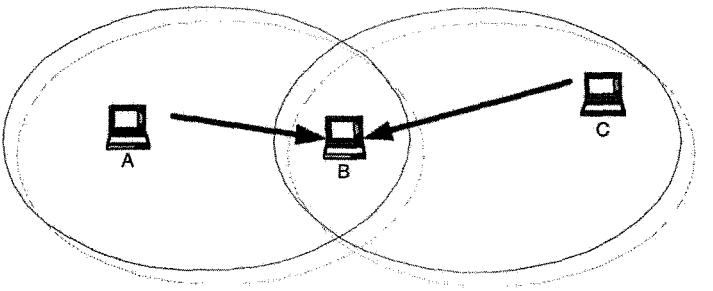


Figure 9.5: The hidden node problem

sumption. If the ACK frame is not received within an ACK time out period, the retransmission takes place in the transmitter. Dealing with this issue in a MAC layer is much more efficient since higher layer timeouts are larger and often measured in seconds.

The Hidden Node Problem

The hidden node problem that is unlikely to occur in a wired LAN is another challenge for WLANs. Because if two stations (A,C) are unreachable and if there is a station (B) in the middle of those two that is reachable to both, transmission from A to B can be interrupted by the transmission from C to B as illustrated in Figure 9.5. IEEE 802.11 introduces two additional frames to the basic access mechanism. The source sends Request to Send (RTS) and the destination replies with Clear to Send (CTS) packets to reserve the channel in advance. Stations that hear RTS delay their transmission until the CTS frame. The stations that hear CTS suspend transmission until they hear acknowledgement. If the stations that hear RTS does not hear CTS, they continue as if they did not hear RTS. Figure 9.6 shows the RTS/CTS frame exchange sequence.

The RTS/CTS transmission introduces a fair amount of capacity

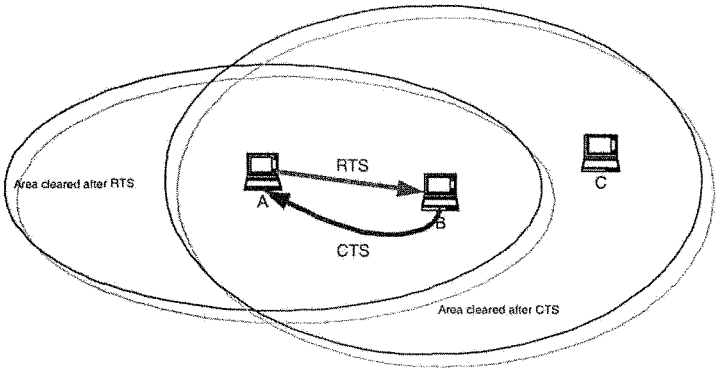


Figure 9.6: RTS and CTS solution

consumption. This makes the RTS/CTS usage decidable. RTS/CTS mechanism can be disabled by an attribute (*dot11RTSThreshold*) in the management information base (MIB).

RTS/CTS can be disabled when there is low demand for bandwidth correspondingly less contention, where all stations are able to hear the transmission of every station. When there is AP, the RTS/CTS is unnecessary in downlink.

The Exposed Terminal Problem

Depicted in Figure 9.7, the exposed terminal problem arises when station (C) attempts to transmit to station (D) while station (B) transmits to station (A). In this case, station (C) is unnecessarily delayed. If there is an RTS/CTS exchange then the CTS frame sent by station (A) will not propagate to station (C). Thus, station (C) knows that the transmission from station (B) to station (A) might not interfere with its transmission to station (D). As a result, the station (C) will initiate transmission to station (D).

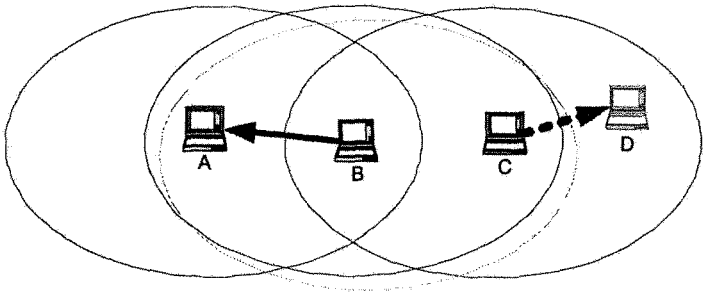


Figure 9.7: Exposed terminal

9.4 Medium Access Control

IEEE 802.11 introduces coordination functions. Coordination functions controls the access of a station to the medium. IEEE 802.11 has two different coordination functions based on contention or polling. The distributed coordination function (DCF) based on contention is the fundamental access method and is also known as CSMA/CA. DCF can be implemented in any network and basically a STA shall ensure that the medium is idle before any transmissions are attempted. Point coordination function (PCF) on the other hand, provides a contention free (CFP) period. Point coordinator (PC) located in AP controls the frame transmissions of the STAs by polling so as to eliminate contention. PCF alternates between contention free period (CFP) and contention period (CP) in time. CP is placed to give transmission right to late coming STAs or to DCF based STAs. PCF is only implemented in the infrastructure BSS.

Frames

The MAC supports three different types of frames:management, control, and data. Figure 9.8 illustrates the format of a typical MAC

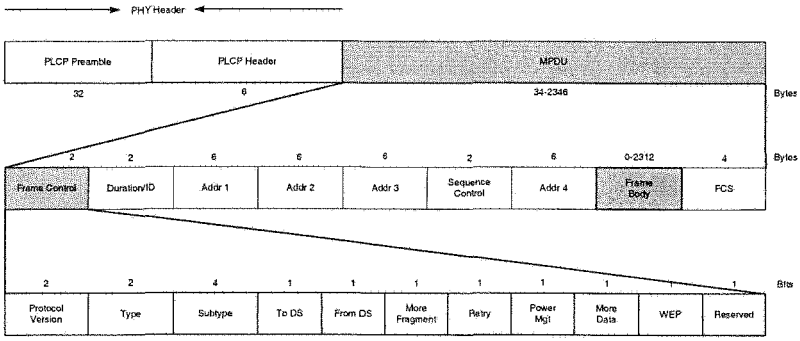


Figure 9.8: MAC frame

frame.

Data Frames Data frames are used to transfer data during CP and CFP. Several data frames are introduced specifically to CP or CFP. *Data* and *Null* frames are used within CP and *Data, Null, CF-Ack, CF-Poll, Data+CF-ACK, Data + CF-Poll, Data + CF-Ack + CF-Poll, CF-Ack + Poll, CF-End, CF-End + CF-ACK* exist for use within CFP.

Control Frames Control frames exist for reliable delivery of Data frames. They perform taking the transmission right in medium, channel acquisition, and acknowledgement of received data. *RTS, CTS, ACK, PS-Poll* are the control frames.

Management Frames WLAN implements management features to enable the station introduce itself to the infrastructure. Management frames perform the operations of joining or leaving a network and managing access points in handoff. This feature includes three steps; the locating of an infrastructure, authenticating to the AP to make the wireless channel reliable, associating to access services of the infrastructure. Types of management frames are *Beacon, Probe Request, Probe Response, IBSS*

announcement traffic indication map (ATIM), Disassociation, Deauthentication, Association Request, Reassociation Request, Association and Reassociation Response, Authentication.

Fragmentation

The packets are partitioned into small pieces in order to reduce the interference of non-STA objects such as Bluetooth devices, microwave ovens which operates also in 2.4-GHz ISM band. As a result, the interference only affects a small portion of the packet and can be recovered easily in small amount of time.

Fragmentation is only applied to packets that have unicast receiver address when the packet length exceeds the fragmentation threshold. Broadcast and multicast frames are not allowed to be fragmented. The fragments are sent in burst unless interrupted in wireless medium and acknowledged individually.

IFS Timing

Time intervals for control are called inter frame spaces (IFS). Starting from the shortest to the longest is as follows;

SIFS Short interframe space: It is the smallest time to give priority to completion of the frame exchange sequence since other STAs wait longer to seize the medium. SIFS[1] time is determined by the delay introduced in PHY and MAC layer.

Slot Time Time is quantized in slots. Slot[2] time is specific to PHY layers. In IEEE 802.11a, Slot time is shorter than SIFS. Backoff

[1]SIFS=Rx RF Delay + Rx PLCP Delay + MAC Processing Delay + Rx Tx Turnaround Time

[2]Slot Time=Channel Clear Assessment (CCA) Time + Air Propagation Time + SIFS

counter is decremented after each idle slot.

PIFS PCF interframe space: STAs operating in PCF seize the medium at least PIFS[3] time in contention free period (CFP) to gain access to the medium.

DIFS DCF interframe space: A STA using the DCF function is allowed to transmit, if its carrier sense mechanism determines that the medium is idle at least one DIFS[4] time. This also allows STAs that are in the RTS/CTS exchange to seize the medium immediately after SIFS.

EIFS Extended interframe space: EIFS period starts when PHY detects idle after an erroneous frame reception. EIFS is to give enough time for transmitter to acknowledge. EIFS[5] is derived from the SIFS, DIFS and the length of time it takes to transmit an ACK control frame with the lowest mandatory rate.

Backoff Time A STA having a packet to send initiates a transmission. STA invokes carrier sense mechanism in DCF to determine the state of the medium. If the medium is busy the STA defers until the medium is idle at least one DIFS period when the last frame detected on the medium received correctly or at least one EIFS when the last frame detected on the medium is erroneous. After DIFS or EIFS, the STA backs off for a randomly chosen integer period[6] over the interval [0,CW]. CW is restricted to lie between CWmin and CWmax which are specified by the PHY layer. Backoff level i is incremented up to CW_{max} or reset if the transmission is unsuccessful or successful respectively. The station attempts to retransmit the unsuccessful frame as long as its retry count reaches retry count limits (*dot11ShortRetryLimit* or *dot11LongRetryLimit*).

[3]PIFS=SIFS+Slot time

[4]DIFS=SIFS+2xSlot time

[5]EIFS=DIFS+ SIFS + ACK_duration(@ lowest mandatory rate)

[6]Backoff Time=Uniform[0, $CW_{min}2^i$-1] x Slot Time

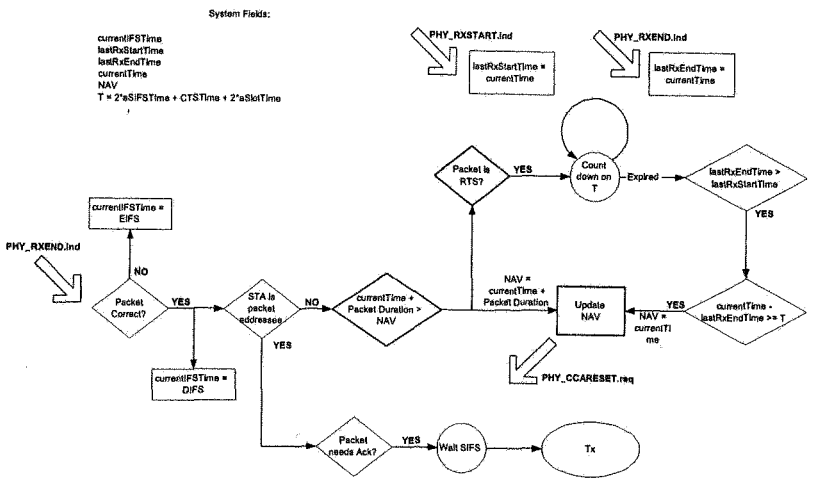


Figure 9.9: NAV Sub-module finite state machine

9.4.1 DCF Access

The CSMA/CA access method is the core of the DCF.

Carrier Sense Mechanism

IEEE 802.11 introduces Physical Carrier Sense (PCS) and Virtual Carrier sense mechanisms (VCS). PCS is a notification from PHY layer to the MAC layer whether medium is idle or not. PCS sets physical allocation vector (PAV). VCS is foreseeing that there is a transmission taking place. VCS sets network allocation vector (NAV) and updates with the value in the Duration/ID field of received packets only when the new NAV value is greater then the existing NAV and only when the STA is not the addressee.

Figure 9.9 indicates the NAV update procedure. When the packet is correctly received, *currentIFStime* attribute is set to DIFS otherwise

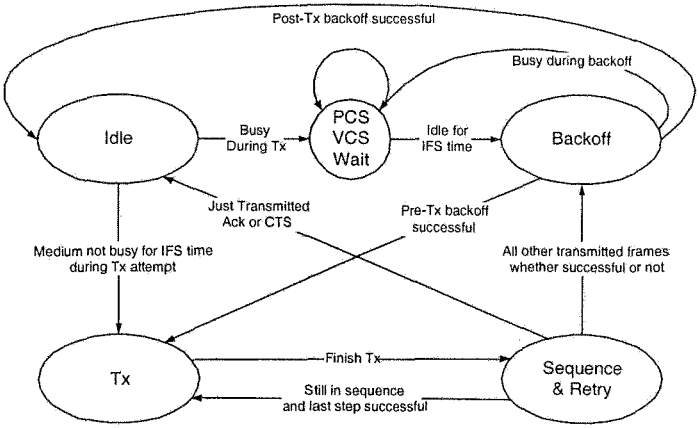


Figure 9.10: DCF finite state machine

to EIFS. If the packet is not sent to this station, the station updates its NAV value if the current NAV value is bigger than the existing NAV value and if the packet is not RTS frame. If the packet is RTS, NAV setting is permitted to reset if there is no start in receiving activity (*PHY-RXSTART*) for a duration of T after end time of receiving (*PHY-RXEND*). This is for the station to detect the erroneous CTS frame and reset its NAV accordingly.

Basic Access Mechanism

Basic access mechanism is illustrated in Figure 9.10. If the station has packet to send, the STA attempts to transmit. First, it decides for RTS if the packet size is bigger than equal *dot11RTSThreshold* and then checks its carrier sense mechanisms. If it detects that the medium is idle longer than DIFS the station transmits immediately as seen in Figure 9.11. If the medium is busy, the station needs to wait at least IFS time before backoff. If the previous frame was erroneous than IFS time is EIFS otherwise it is DIFS.

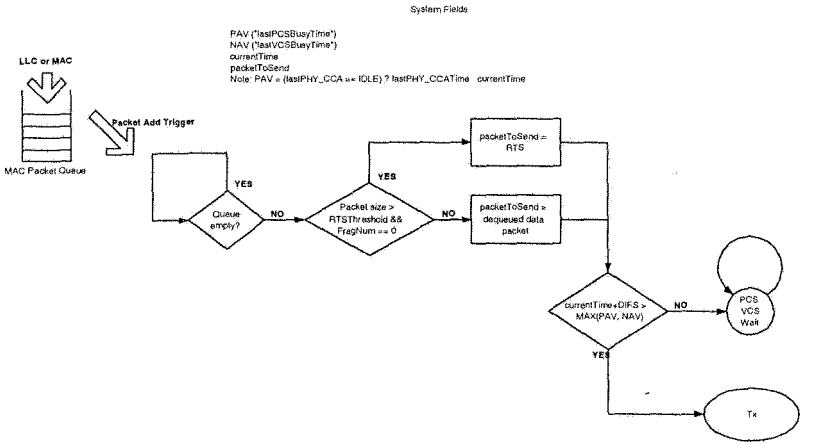


Figure 9.11: Idle sub-module finite state machine

In the beginning of the backoff procedure, the STA sets its Backoff timer to a random backoff time described in Section 9.4. If the previous frame that STA sent was correctly transmitted than backoff level i is set to 0 indicating that it is the initial attempt of transmission. After selecting a backoff level, the backoff counter is decremented for each idle slot. During backoff, if the medium is determined to be busy, decrementing of backoff timer is stopped. The procedure resumes after noting the medium is idle for the duration of a DIFS or EIFS period (as appropriate 9.4). The procedure is illustrated in Figure 9.12.

When the backoff timer reaches zero, the STA initiates transmission. The transmission precedes another backoff procedure whether the transmission is successful or not. If the transmission turns out to be erroneous, then STA waits EIFS time and selects a backoff value from $[0, CW_{min} \cdot 2^i\text{-}1]$ after incrementing its backoff level from $i-1$ to i. Otherwise, station resets its backoff level and sets its backoff timer from $[0, CW_{min}\text{-}1]$. If the transmitted frames are ACK or CTS frames then STA goes directly to idle state and does not backoff. By placing

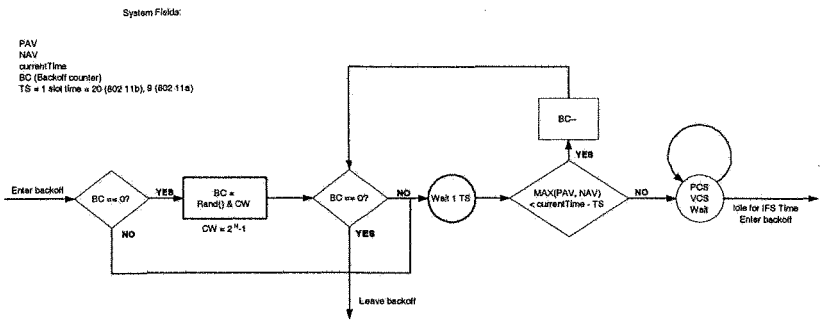


Figure 9.12: Backoff sub-module finite state machine

backoff between transmissions, the standard tries to implement fairness.

If the transmitted frame needs to be acknowledged then STA sets its ACK timer. If there is a timeout, STA increments one of the short retry counter (SRC) (or long retry counter (LRC)) if the packet size is smaller (or bigger) than *dot11RTSThreshold*. Following the Figure 9.13, if the valid ACK is received than SRC, LRC and backoff level is reset.

The station immediately continues with transmitting if it is in the middle of a sequence: RTS/CTS transmission sequence or fragmented packet stream. Otherwise, station invokes backoff procedure.

A realization of DCF access procedure is shown in Figure 9.14.

Features

Control SIFS interval plays a major role in control of delivery. Each packet is separated with a SIFS time interval. This SIFS time corresponds to the time that is introduced by the process in PHY and MAC layer. For example; After transmission, the STA switches from transmitting to receiving state and waits for ACK

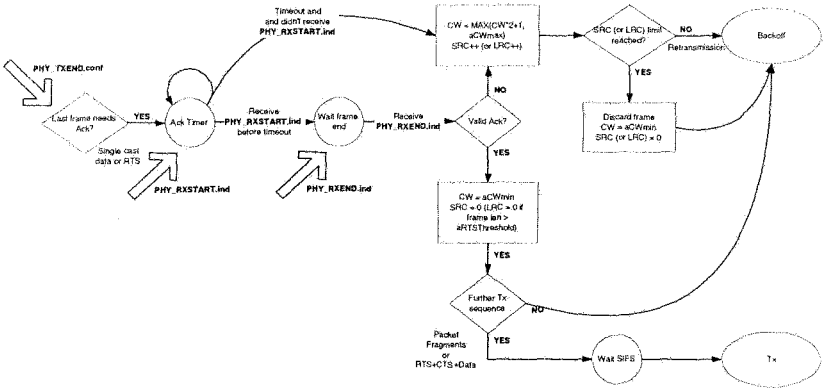


Figure 9.13: Frame sequence and retry sub-module finite state machine

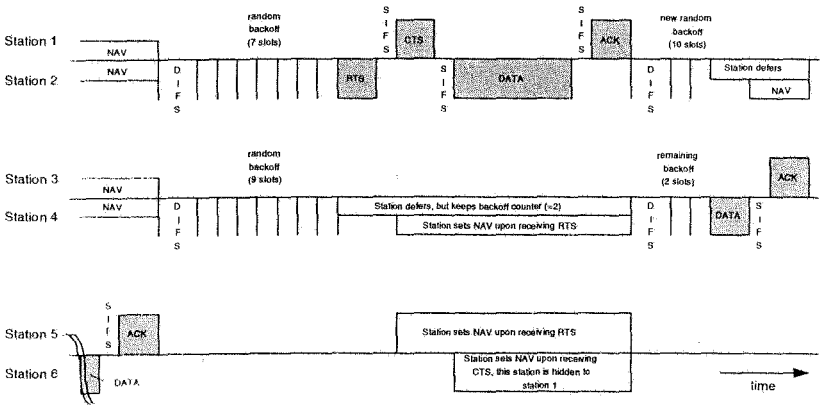


Figure 9.14: Timing of the 802.11 DCF: Note that station 6 cannot hear station 2 but station 1

or after receiving RTS (or ACK), the STA switches from receiving to transmitting and transmits CTS (or Fragment) frame.

Fragmentation RTS/CTS usage with fragmentation is shown with the following procedure. Each fragment is acknowledged. The RTS and CTS frames shall be used to set the NAV to indicate busy until the end of ACK 0 which is ACK of fragment 0. Both Fragment 0 and ACK 0 contain duration information for NAV until the end of ACK 1. This shall continue until the last fragment. Fragment and ACK play the role of RTS and CTS virtually.

CTS Procedure A STA who is the addressee of the RTS frame shall transmit a CTS frame after a idle SIFS period as shown in Figure 9.14. The duration field of the CTS frame is the subtraction of SIFS time and CTS time from the duration field of received RTS frame. After transmitting RTS frame, STA waits for a *CTS Timeout* interval to recognize a valid CTS frame. The recognition of any other frame is interpreted as RTS failure and STA invokes backoff procedure.

Transfer The RTS/CTS is used for unicast frames when the packet size is not less than *dot11RTSThreshold* attribute. In the case of broadcast or multicast packet transmission, only basic access methods are used (without RTS/CTS) and no ACK shall be transmitted by the any of the recipients of the frame. Reliability on broadcast or multicast frames is reduced relative to the unicast frames.

ACK Procedure Upon successful reception, recipient generates ACK frame (See Figure 9.14). The source STA waits *ACK Timeout* interval to conclude that transmission is failed if during which if there is not a valid ACK. Thus, STA invokes backoff procedure.

Duplicate Detection There is a possibility that a frame may be received more than one times due to acknowledgements and re-

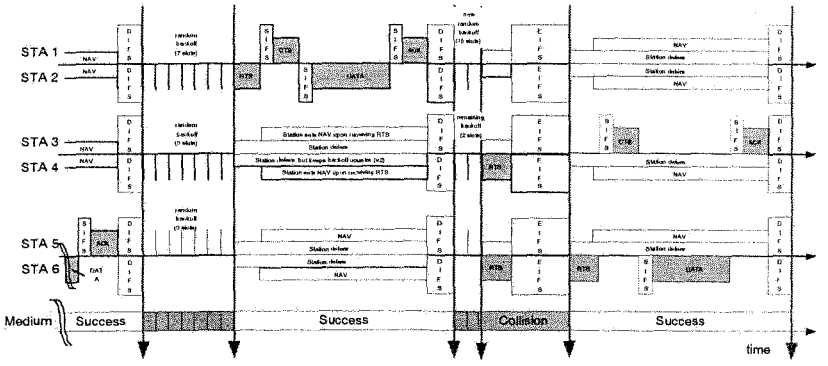


Figure 9.15: Time realization of DCF

transmissions. Duplicate data framing is filtered out by using the sequence number and fragment number fields assigned to each frame. The receiver keeps a cache of <Address, sequence number, fragment number> tuples and if there is duplicate frame that matches a tuple with the Retry bit set, the STA shall reject the frame. On the other hand, the destination STA should acknowledge all successfully received frames even if the frame is a duplicate.

9.4.2 Markov Model of DCF

Time realization of the DCF procedure can be observed as a combination of empty slots and transmissions as seen in Figure 9.15. DCF can be constructed as a two dimensional Markov model where state transmissions are triggered by either an empty slot or a transmission [128]. The transmission is done if there is a packet to send and if the channel is not busy. The station generates packet with probability λ. The derivation has two phases first obtaining the τ which is the probability that the station transmits a packet and then placing the τ in throughput calculation [115, 128].

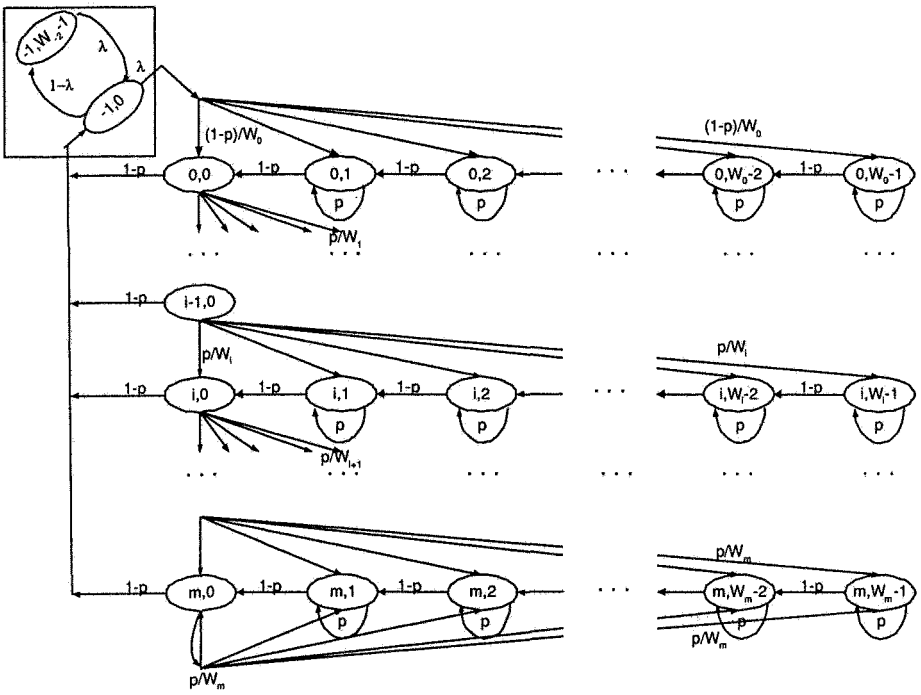


Figure 9.16: Markov Chain model for the IEEE 802.11 DCF Model in normal operating condition

Semi-discrete Markov model shown in Figure 9.16 has two dimensions which are $s(t)$ and $b(t)$ representing depth and slot number respectively. p represents the *conditional collision probability* and is independent and constant in each packet collision regardless of the number of retransmissions suffered. p also stands for probability of detecting the channel busy. As it can be seen in Figure 9.16 there is a self-loop to represent the froze in backoff states. If the stations hear each other, it is guaranteed that in IFS period after packet transmission, every station waits thereby collision probability is zero within the IFS waiting time.

The bidimensional process $\{s(t), b(t)\}$ is represented by the one step transition probabilities depicted in Figure 9.16. Having $b_{i,k}, i \in (-1, m), k \in (0, W_i - 1)$ as the stationary distribution we can solve this Markov chain easily by balance equations.

$$\begin{array}{ll} b_{i-1,0} \cdot p = b_{i,0}, & b_{i,0} = p^i b_{0,0} \quad i < m \\ b_{m-1,0} \cdot p = (1-p) b_{m,0}, & b_{m,0} = \frac{p^m}{1-p} b_{0,0} \\ b_{-1,0} . \lambda = b_{0,0} & \\ b_{-1,W_{-1}-1} = \frac{(1-\lambda)}{\lambda^2} b_{0,0}. & \end{array} \tag{9.4}$$

Following Equation 9.4, the stationary distribution can be represented as

$$b_{i,k} = \frac{W_i - k}{W_i(1-p)} \cdot \begin{cases} \lambda(1-p) \cdot b_{-2,0} & i = 0 \\ p \cdot b_{i-1,0} & 0 < i < m \\ p \cdot (b_{m-1,0} + b_{m,0}) & i = m \end{cases} \tag{9.5}$$

We can express $b_{i,k}$ for backoff section using (9.4) as follows;

$$b_{i,k} = \frac{W_i - k}{W_i(1-p)} b_{i,0} \;\; i \in (0, m), \; k \in (0, W_i - 1). \tag{9.6}$$

Sum of the probabilities of each state is equal to 1,

$$1 = \sum_{i=0}^{m} \sum_{k=0}^{W_i - 1} b_{i,k} + \sum_{k=0}^{W_{-1}-1} b_{-1,k}. \tag{9.7}$$

In order to find the τ, we need the stationary probability of $b_{0,0}$. As we can represent each state by $b_{0,0}$ by using Equations 9.4-9.6.

$$\begin{aligned} backoff &= \sum_{i=0}^{m} \sum_{k=0}^{W_i-1} b_{i,k} \\ &= \frac{b_{0,0}}{(1-p)} \left(\sum_{i=0}^{m} p^i \left(\frac{W_i+1}{2} \right) + \frac{p^{m+1}}{(1-p)} \left(\frac{W_m+1}{2} \right) \right) \end{aligned} \tag{9.8}$$

Stationary probability in idle state is introduced as

$$idle = \sum_{k=0}^{W_{-1}-1} b_{-1,k} - 1 = \frac{b_{0,0}}{\lambda^2} - 1. \tag{9.9}$$

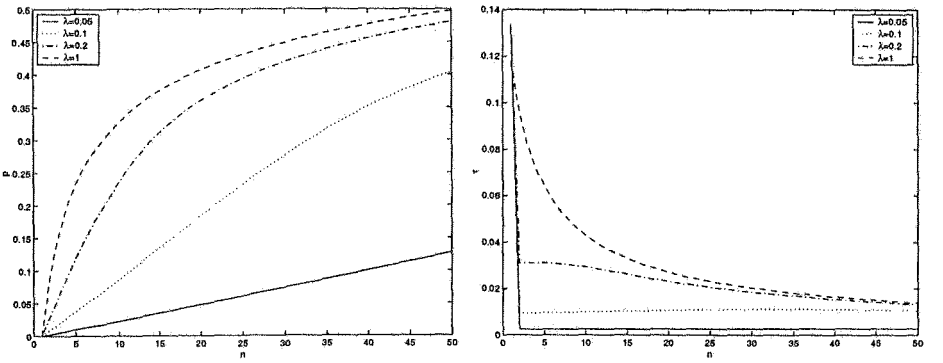
where $W_{-1} = 2$ and where (-1) in the equation is put to make the equation consistent with the saturation operating case which is when $\lambda = 1$. Since adding additional states introduces at least one slot time delay to the system. Having no (-1) will result in a situation where the station waits at least one slot time for transmission. Although this does not show significant change when λ is high, it is important to note the characteristic of modelling a finite state machine by a Markov model.

Packet is transmitted in states $b_{i,0}$ where $i \in (0, m)$. We can further simplify by adopting the contention window as $CW_{max} = 2^m CW_{min}$ which leads to $W_i = 2^i W$ where $W = CW_{min}$ and $i \in (0, m)$ and the transmission probability τ becomes

$$\begin{aligned} \tau &= \sum_{i=0}^{m} b_{i,0} \\ &= \frac{b_{0,0}}{(1-p)} \\ &= \frac{1}{\frac{(1-2p)(W+1)+pW(1-(2p)^m)}{2(1-2p)(1-p)} + (1-p)\left(\frac{1}{\lambda^2}-1\right)} \end{aligned} \tag{9.10}$$

τ is a function of $\tau(p, m, W, \lambda)$ where only p is unknown. When $n \to \infty$ then $p \to 1$ and $\tau \to 0$ and when $n \to 1$ then $p \to 0$ and τ is as follows;

$$lim_{p \to 0} \tau = \frac{1}{\frac{(W+1)}{2} + \left(\frac{1}{\lambda^2} - 1\right)} \tag{9.11}$$



(a) Probability of Collision (b) Probability of transmission

Figure 9.17: p and τ values versus n

Collision probability depends on the active number of stations n. Therefore p and τ in steady state are tied by

$$p = 1 - (1-\tau)^{n-1} \rightleftharpoons \tau(p) = 1 - (1-p)^{\frac{1}{(n-1)}} \tag{9.12}$$

assuming all stations are independent. Now, τ is a function of known variables $\tau(n, m, W, \lambda)$. Since $p \in (0,1)$, $\tau \in (0,1)$ and both equations (9.10) and (9.12) are continuous, the solution of τ is unique. The Figure 9.17 shows the behavior of p and τ versus number of stations n.

Definition. *The throughput is the fraction of time the channel is used to successfully transmit payload bits.* Let's first define the probability P_{tr} that there is at least one transmission in the considered slot time and the probability P_s that a transmission occurring on the channel is successful are as follows;

$$\begin{aligned} P_{tr} &= 1 - (1-\tau)^n \\ P_s &= n\tau(1-\tau)^{n-1}. \end{aligned} \tag{9.13}$$

Denote the throughput S as the ratio

$$S = \frac{P_s E[P]}{(1-P_{tr})\sigma + P_s T_s + (P_{tr} - P_s)T_c}. \tag{9.14}$$

where $E[P]$ is the average packet payload size, (P_s) is the probability for a successful transmission. The average length of empty slot time is with probability $(1 - P_{tr})$ and with probability $((P_{tr} - P_s))$ there is a collision. T_s is the average time the channel is sensed busy due to a successful transmission, and T_c is the average time the channel is sensed busy by each station during a collision. σ is the duration of an empty slot time. We can infer that individual throughputs are equal and $\frac{1}{n}S$.

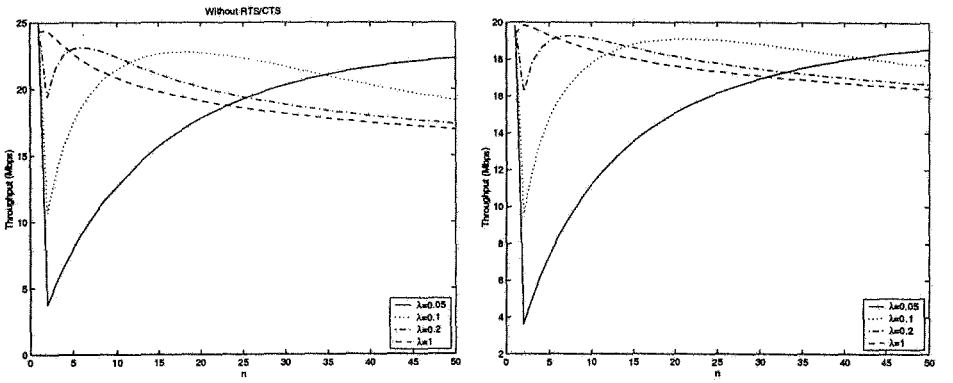
For two different access mechanism the time values are as follows;

$$\begin{aligned} T_s^{basic} &= T_{DATA} + SIFS + \delta + T_{ACK} + \delta + DIFS \\ T_c^{basic} &= T^*_{DATA} + \delta + EIFS \\ T_s^{rts} &= T_{RTS} + SIFS + \delta + T_{CTS} + SIFS \\ &\quad + \delta + T_{DATA} + SIFS + \delta + T_{ACK} + \delta + DIFS \\ T_c^{rts} &= T_{RTS} + \delta + EIFS. \end{aligned} \tag{9.15}$$

where T_{DATA} is the time that takes to send a packet with size $E[P]$ and T_{RTS}, T_{CTS}, T_{ACK} are the times that take to send the corresponding frames. T^*_{DATA} stands for average time that takes to send $E[P^*]$ which is the average length of the longest packet payload involved in a collision. When all packets have the same size, $E[P] = P = E[P^*]$. δ is the propagation delay. Unlike the "basic" access mechanism T_c^{rts} only contains T_{RTS} since only possibility to experience a collision occurs in the RTS frame transmission.

Timing values are obtained from the IEEE 802.11a OFDM physical layer parameters where the payload is considered constant and $E[P]$ is 1024 *bytes*.

Figure 9.18(a) shows that the throughput for the basic access scheme saturates as the number of active stations increases. In real life we can intuitively infer that backoff, deference and idle state introduces spaces in medium and throughput can increase if there is a station that can utilize these spaces efficiently. It is important to note that the network is saturated in different network sizes depending on the traffic intensity.



(a) Data Rate=54Mbps w/o RTS/CTS (b) Data Rate=54Mbps with RTS/CTS

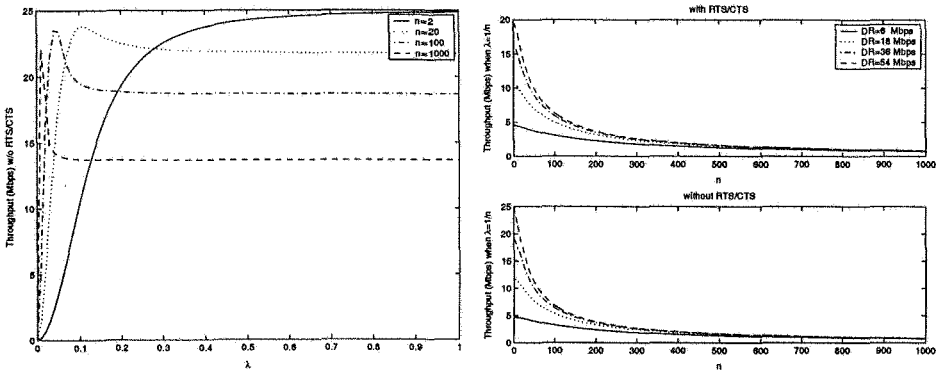
Figure 9.18: Throughput versus number of active nodes

Figure 9.18(b) reports the throughput for RTS/CTS access mechanism. In RTS/CTS access, the saturated behavior is not easily seen but as $n \rightarrow \infty$, the throughput decreases and goes to 0 slowly.

Figure 9.19(a) is to observe the traffic intensity over throughput when there is fixed number of stations. One can easily observe the same saturation behavior. Figure 9.19(b) is to show the behavior when the total load is constant. As n goes to infinity total throughput decays below 5 Mbps with whichever data rate is used.

9.4.3 PCF

The point coordinator function (PCF) introduces another access mechanism which can assist sessions that requires quality of service. PCF provides contention free period that alternate with the contention period. As opposite to DCF, PCF implements a centralized control where AP, wherein the point coordinator (PC) reside, is the central controller of the network. AP restricts the access to the medium and any station whether agreed to operate in PCF or not but associated can transmit



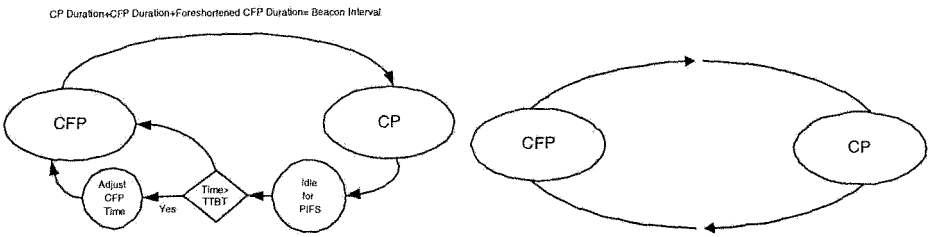
(a) Data Rate=54Mbps w/o RTS/CTS, fixed n (b) Fixed total load, $n \cdot \lambda = 1$

Figure 9.19: Throughput versus number of active nodes for fixed users and fixed total load

data as long as AP allows them to do so.

CFP Timing

PCF quantizes the time with superframe. Superframe is repetitive and composed of contention free period (CFP) and contention period (CP). Superframe starts with beacon as seen in Figure 9.20, AP and stations alternate in CP and CFP operation. CP is placed in order to give chance to new coming stations to introduce themselves to the network. Thus, CP must be long enough to transmit at least one frame sequence. *CFPRate* is defined as a number of DTIM intervals. *CFPRate* may be longer than a Beacon period as a result Beacon frames contain a *CFPDurRemaining* field in CF Parameter Set in order to indicate the remaining time to the end of CFP period. In CP, *CFPDurRemaining* is set to zero.



(a) Access Point (b) Station

Figure 9.20: Macro states in PCF

PCF Access

Figure 9.21 shows the timing diagram of a typical PCF operation. Beacon frame contains a CFP duration (*CFPMaxDuration*[7]) and have STAs set their NAV. Each frame transmission in CFP is interpolated by SIFS and PIFS which makes non-polled STAs lag behind since DIFS is bigger than PIFS and SIFS. When AP seizes the medium, AP polls STAs that agreed to be in *polling list* during association. In respond to receiving a contention-free poll (CF-Poll) frame from AP, STA is allowed to send only one frame. Each frame is separated by SIFS interval. if AP does not hear any respond to its CF-Poll frame for a PIFS period, AP polls the next station as the situation depicted in Figure 9.21.

Beacon Frame

Beacon and *Probe Response* frames contain the CF Parameter Set information element. *Probe Response* frame is sent in order to respond to the *Probe Request* frame transmitted by STAs. STA sends *Probe*

[7]CFPMaxDuration field is set to the minimum interval that may allow sufficient time for the AP to send one data frame to a STA and for the polled STA to respond with one data frame. CFPMaxDuration is set to the maximum interval that may allow to send at least one data frame during the CP.

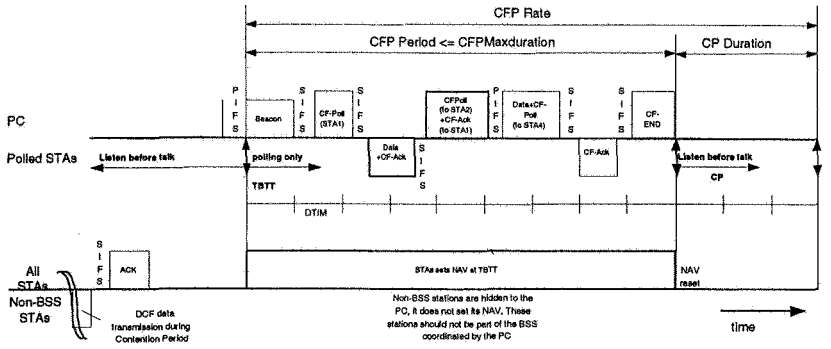


Figure 9.21: Timing diagram for PCF

Request frame to find and initiate association to a network before waiting to hear a *Beacon*. CF Parameter Set contains attributes that delineate the outline of the PCF operation.

CFP Period Length of CF Period (indicated in DTIM intervals).

CFP Count Time that is between Beacon time and start of the next CF Period (indicated in DTIM intervals).

CFP MaxDuration Maximum duration of CFP period (indicated in time units (TUs)).

CFP DurRemaining From Beacon time to the end of CFP period (indicated in TUs).

Piggybacking

In order to save bandwidth, piggybacking is introduced. Piggybacking is the functionality of expressing the presence of frames in a single frame. For example *Data+CF-Ack+CF-Poll* brings data transmission, polling feature, and acknowledgement in one frame.

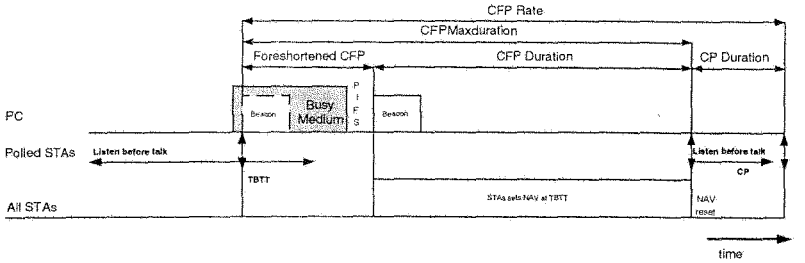


Figure 9.22: Contention free period determination

CFP Duration

Figure 9.22 shows the situation when there is a delay in Beacon transmission. AP tries to keep the Beacon transmission periodic by shortening the contention free period if there is a delay in Beacon transmission due to CP interval. Contention free period[8] is adjusted according to Target Beacon Transmission Time (TBTT). The PC also ends the CFP period at any time by *CF-End* frame depending on the polling list, traffic load, or any other reason.

NAV Operation

NAV operation facilitates the PCF in the case of overlapping BSSs. Each STA in BSS sets its NAV to *CFPMaxDuration* with the beacon, at each target beacon transmission time (TBTT) and updates its NAV by *CP Period* field with every other beacon frame. This operation prevents STAs from seizing the medium.

A new coming STA updates its NAV with the information in the *CFPDurRemaining* field of any received Beacon and waits to hear a *CF-End* or *CF-End+ACK* frames in order to reset its NAV.

[8]CFP Period=*CFPMaxDuration*-(Actual CFP Start-TBTT)

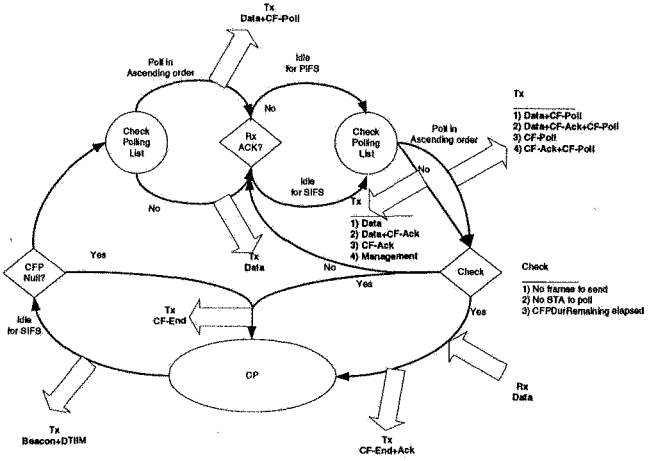


Figure 9.23: Access point CFP finite state machine

PCF Transfer Procedure

According to the direction of the transmission whether from AP/PC or to AP/PC, *From DS* or *To DS* fields are set in the frame structure (See Figure 9.8) respectively.

A STA notifies AP about its decision about being included to the polling list (*CF-Pollable*) in association. A STA wanting to change its condition may initiate retransmission for being included (or secluded) to (from) polling list. When included, STA alternates between CP and CFP as AP. The CFP state machines both for AP and STA are showed in Figures 9.23 and 9.24 respectively. One can follow the figures easily to get the operation and packet exchange of AP and STAs.

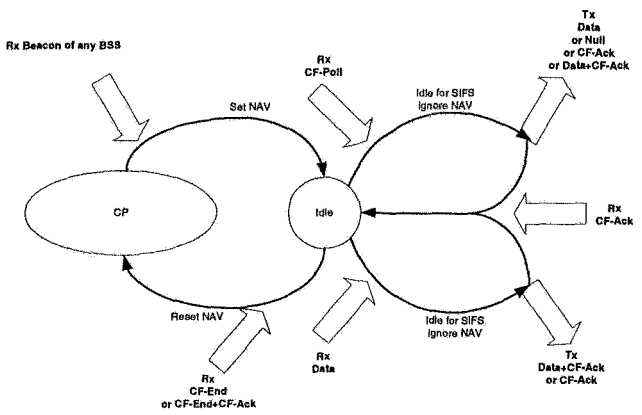


Figure 9.24: Station CFP finite state machine

9.5 Management

The IEEE 802.11 has management features to compete with the wireless medium. Wireless medium is a kind of a medium where STA's crossing the physical boundaries are prevalent and unavoidable. These make the medium unreliable and easily accessible. Unattached STAs drain the battery power easily, the power consumption is another concern and can be reduced by reducing the operating time of radio without any degradation to the performance.

9.5.1 Synchronization

Synchronization plays a crucial role for all STAs. Channel access mechanism and physical layer needs precise timing. All STAs within a single BSS are synchronized to a common clock. Timing synchronization function (TSF) is introduced to keep the timers of all STAs within the same BSS synchronized to a local TSF timer. Timing values are put in beacon and probe response frames. Beacon frames are

periodic and supposedly transmitted at end of beacon interval which is called target beacon transmission time (TBTT).

TSF for infrastructure BSS

AP is the master to regulate the TSF. The AP periodically transmits beacon frames that contain the TSF timer. A receiving STA accept the timing information in beacons sent from its AP.

TSF for independent BSS

A distributed algorithm is implemented to apply TSF in IBSS. Each station is capable of transmitting beacon at each TBTT. Each station randomly picks a backoff interval from 0 to twice the CWmin. The one who picks the smallest backoff interval sends the beacon and this makes the others update their clock according to the time indication in beacon and cancel their beacon transmissions.

The synchronization mechanism in IBSS does not show a precise timing but adjusts the time to the time of STA who advances in time.

9.5.2 Scanning

A wireless medium where there is no physical attachment requires additional functionality to locate a network. Thus, scanning is a functionality implemented in STAs to help them to find and identify BSSs. Each BSS is distinguished by its type (independent or infrastructure), BSSID (BSS id), and SSID (ESS id).

Passive Scanning

Passive scanning implements a power efficient mechanism. The mechanism is to listen each channel for a definite time interval. In each channel, a STA seeks for a beacon frame which already contains the information that is necessary to identify a BSS. It reports those findings after going through scanning channels.

Active Scanning

Active scanning involves finding the network rather than listening. STA sends *probe request* to receive *probe response* frame from AP or STA which contains the same information as beacon frame. A STA moves to a channel and solicit a *probe request* frame. If there is a network in the vicinity, the station who sends the last beacon frame responds to the soliciting station with *probe response* frame.

After locating APs, station shall do authentication with different APs. This procedure is called pre-authentication and saves time if STA frequently changes APs. The station willing to join a BSS first chooses a BSS and initiates authentication with the AP or its peer in IBSS. The station joining an infrastructure BSS need to perform association to gain access to the infrastructure.

9.5.3 Power Management

Power is depleted easily in mobile stations. Major power consumption occurs in transmission since as opposed to a wired domain, continuous connectivity depends on continuous signalling. The power management feature implements a power saving mechanisms for IEEE 802.11 nodes to save power without losing connectivity.

Power Management in infrastructure BSS

AP which normally has continuous power due to its connectivity to the wired domain pursues the power management. AP has buffer to save the packets of those who are in power save (PS) mode. There are two information field: traffic indication map (TIM) for unicast packets and delivery traffic indication map (DTIM) for multicast and broadcast packets. These are virtual bitmaps and have an entry if the STA, having a unique association (AID) has a packet in buffer. The multicast and broadcast packets are paired with the AID equals to 0. TIM is conveyed with each beacon frame and DTIM is conveyed in beacon frames at fixed number of beacon intervals. Each STA is supposed to wake up to listen beacon frames. They send PS-Poll frame to request the buffered frames. PS-Poll transmission also obeys the rules in CSMA/CA scheme and each station selects a backoff interval before transmitting to avoid collision.

Power Management in independent BSS

The IBSS implements a distributed power management scheme. Each station is capable of leading the power management. Each STA must be awake after each beacon for announcement traffic indication message (ATIM) window. If the beacon arrival is later than TBTT than ATIM window shrinks. During ATIM window, STA waits for beacon, RTS, CTS or ATIM frame. The STA who has packet to send notifies the destination by ATIM frame. The STA who sends the beacon is not allowed to enter power save mode until the end of the next beacon.

9.5.4 Security

A wireless medium requires a privacy mechanism which protects the data frames from potential eavesdroppers since publicly available radio signals are vulnerable to any kind of attacks. IEEE 802.11 stan-

dard defines Wired Equivalent Privacy (WEP) mechanism to provide a protection to data frames. WEP is basically an encryption mechanism that uses an encryption algorithm to convert the frame body of MAC packet to a uniquely decodable encrypted one. WEP focuses more on secure WLAN connectivity than on verifying user or station identity. For enterprise or campus WLANs, an efficient and scalable authentication mechanism is needed to manage hundreds or thousands of users. IEEE 802.1x is an authentication framework for WLANs and leverage an existing protocol known as the Extensible Authentication Protocol (EAP [RFC 2284]). Following WEP we explain IEEE 802.1x.

WEP

WEP uses a RC4 cipher. RC4 is a variable length symmetric chipper. It is symmetric since the same key and algorithm are used for both encryption and decryption. There is a secret key known by the pairs. The secret key is concatenated by initialization vector (IV) and used as a seed in pseudo-random number generator (PRNG). The resulting key sequence and plain text are XORed. Enciphered message and IV is the output message. The receiver does the reverse operation by knowing IV, secret key and cipher text. The secret key distribution is not defined in the standard and it is generally handled manually.

WEP is not capable of providing full secure solution. It is easily breakable and key management is a serious problem. Once the key is widely shared its secrecy ceases. Manual key distribution is another problem and it is not scalable. WEP does not protect users to each other if they have the same key.

IEEE 802.1x

Evolution of 802.1x starts from Point to Point protocol (PPP) and EAP. PPP is most commonly used in dial-up access in order to iden-

tify the user. PPP evolved into EAP where there is more security features besides simply employing username and passwords for access. EAP brings the feature that simplifies the authentication mechanism and leaves it to only client and authentication server. Thus, middle actors get out of business which makes the system inter-operable and compatible.

IEEE 802.1x is basically EAP over LANs (EAPOL). The EAP information is passed through by Ethernet frames and the overhead of PPP is eliminated. There are three actors involved in the IEEE 802.1x process. The client called supplicant, the authentication server typically a RADIUS server and the middle actor called authenticator which is usually the access point. The significant point is that there is no need for intelligence in authenticator.

When authenticator locates an active link, it sends EAP - Request / Identity message to make the supplicant aware of the authentication mechanism. The authenticator plays a relay mechanism and forwards the messages from or to a supplicant or authentication server. The handoff mechanism which is moving an authentication relationship from one access point to another is not in IEEE 802.1x because it is designed initially for wired domain. This open area shall be addressed under IEEE 802.11i protocol.

After authentication EAP-Key frame is sent from access point to supplicant. This feature allows to change keys dynamically between access point and client on run time. This is very crucial since as we pointed out, long lifetime key is a easily hacked deficiency of WEP mechanism.

9.6 IEEE 802.11 and 802.11b Physical Layer

The physical layer (PHY) is the layer 1 element of OSI protocol stack. It is the responsible module to relieve the bits to the air. IEEE 802.11

introduced Frequency Hopping Spread Spectrum (FHSS), Direct Sequence Spread Spectrum (DSSS), Infrared Light (IR) PHY standards in 1997 and 802.11b: High-Rate Direct Sequence (HR/DSSS), and 802.11a: Orthogonal Frequency Division Multiplexing (OFDM) PHY standards in 1999.

The PHY layer is composed of physical layer convergence (PLCP) and physical medium dependent (PMD) layers as seen in Figure 9.4. PLCP is an interface to MAC layer and PMD is equipped with transmission interface to send and receive files over the air. PLCP layer prefixes preamble and header to the MAC layer data frame as seen in Figure 9.8. The PLCP preamble is to provide enough time for receiver to apply its functions such as antenna diversity, clock, data recovery, and field delineation of the rest of the packet. There is a synchronization field in preamble to detect the signal and there is a Start Frame Delimiter (SFD) field for receiver to lock the start of the frame. The PLCP header is used to provide enough information for the receiver to process the packet.

Allocated spectra for wireless LAN applications are typically in 2.4 GHz and 5 GHz range where the bandwidth is scarce and much in demand. WLAN system at 2.4 GHz must meet ISM band requirements. The IEEE 802.11 standard supports both FHSS and DSSS techniques for this band. Spread spectrum techniques shows robustness against interference and do not require adaptive equalization. However, for higher data rates the synchronization requirements of spread spectrum techniques are more restrictive and complex. For data rates of above 10 Mbps, OFDM system shows better performance. It is rewarding to cover spread spectrum to observe the motivation towards OFDM based physical layer.

9.6.1 Spread Spectrum

Initiated from military technology, spread spectrum technology became the vital component ranging from cellular to cordless to WLAN

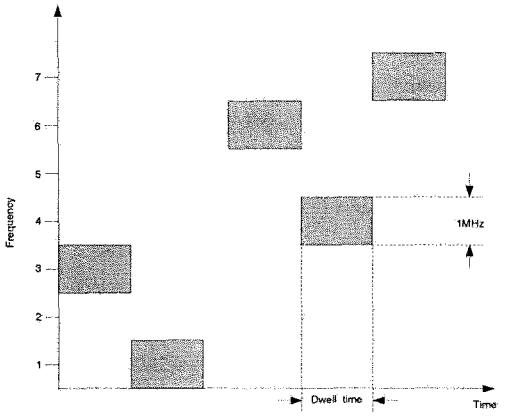


Figure 9.25: Frequency hopping

systems. Spread spectrum uses wideband, noise-like signals that are hard to detect. Spread spectrum technique is spreading the narrow band signal over wider band by pre-known or pseudo-noise spreading code in both sides in order to increase the bandwidth of signal to be transmitted. The main parameter is processing gain which is defined as the ratio of transmission bandwidth and information bandwidth. The receiver performs the inverse operation and consequently the spread signal is reconstructed in a narrow band along with narrowed noise. Spreading the signal over wider frequency makes the signal more noise like and secure. Having a different code in each receiver diminishes the interference considerably but not totally. It also allows access at any arbitrary time.

Processing gain determines the number of users that can be allowed, the amount of multi-path effect reduction and the difficulty to jam or detect a signal etc. Its advantageous to have processing gain as high as possible.

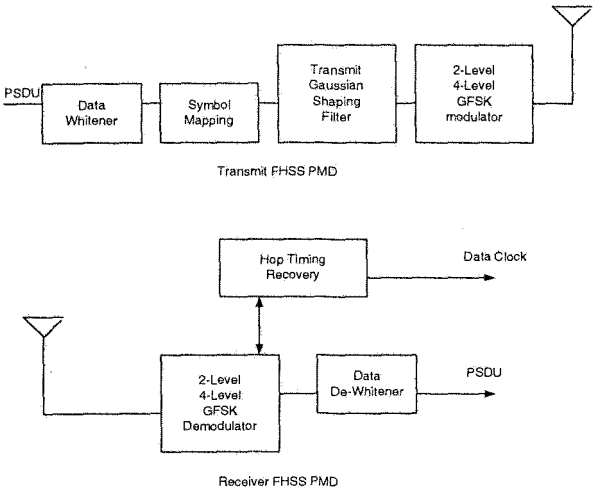


Figure 9.26: Transmit and Receive FHSS © IEEE

9.6.2 FHSS Physical Layer

Frequency hopping is simply changing the frequency from one to another in each timely slotted intervals. The hopping pattern is unique and available in the transmitter and the receiver. In this case bandwidth is increased by a factor length of the sequence. Time spent in one frequency is called *dwell time*. There is a synchronization mechanism that helps the receiver and the transmitter pair operate in the same frequency and in the same dwell time. Beacon frames include information about the pattern and time. Thus a receiver can know the hopping sequence number, hopping index and timestamp. Hopping occurs when timestamp equals to the multiples of dwell time (See Table 9.2).

Channel length	1 MHz
Channel number	79 from 2.402 GHz
Dwell Time	0.4secs
Hoping Sequence Size	26
Data rate	1-2 Mbps

Table 9.2: FCC rules for IEEE 802.11 FHSS

Gaussian Frequency Shift Keying

FHSS uses Gaussian Frequency Shift Keying (GFSK): GFSK uses frequency modulation to get rid of unwanted changes in the amplitude. A narrow band signal is filtered with low pass Gaussian filter (500KHz bandwidth in 3dB) and then the signal is FM modulated in GFSK. It is basically deviating the frequency of either side of the carrier hop frequency depending on if the value of the binary symbol being transmitted equals to 1 or 0. 1 Mbps and 2 Mbps data rate is achieved by 2-level GFSK and 4-level GFSK respectively since 4GFSK uses four symbols instead of two as in 2GFSK and also brings a complex and high-cost transceiver pair.

FHSS PLCP and PMD Sublayer

The PLCP state machine contains transmit, receive and carrier sense / clear channel assessment (CS/CCA) states. CS/CCA is to support CSMA/CA mechanism of MAC layer. CS/CCA informs the MAC layer about the medium whether it is busy or idle.

The PLCP Header contains PSDU Length Word (PLW), PLCP Signalling Field (PSF) and Header Error Check (HEC) fields. They are for the length of the MAC frame, data rate, frame check respectively. The PLCP Header is sent always with 1 Mbps. Data whitening is applied to the PSDU before transmission to minimize DC bias. The PMD layer is responsible for controlling the hopping sequence and

transmitting the whitened PSDU by hopping from channel to channel in a pseudo-random fashion. The transmitter and receiver are illustrated in Figure 9.26.

Pros and Cons

The wireless environment introduces frequency selective fading and fades are correlated in adjacent frequencies. The hopping sequence determination should try to avoid selection of the adjacent channels. FHSS shows resistance to jamming unless the jammer jams all frequencies. Collision in one or two frequencies can be recoverable. Time spent for changing frequencies delays transmission time. The FHSS systems are very cheap. At one time, they were very important in the history of communication. However, higher rate systems have overtaken FHSS technology due to their ability to achieve higher processing gain.

9.6.3 DSSS Physical Layer

Direct Sequence is the best known Spread Spectrum Technique. A data signal at the point of transmission is multiplied with a higher data-rate bit sequence (also known as a chipping code) that divides the data according to a spreading ratio and spreads the spectrum out while dropping the power spectral density. This chipping code introduces redundancy which enables correct recovery of the symbol, even if some of it is damaged. Increasing the rate of the chipping signal increases the throughput and consequently requires higher bandwidth and complex transmitter receiver pair. A good chipping sequence is called pseudo noise (PN) sequence since a given PN sequence is orthogonal with all the other PN sequences of the same length and a given PN sequence is almost uncorrelated to a shifted version of the given sequence which is very important in synchronization.

Channels	1 to 11 (2.412-2.462 GHz)
Channel bandwidth	5 MHz
Channel Spread	25 MHz
Chipping Sequence	Barker Sequence
Data rate	1-2 Mbps
Min. Processing Gain	10dB

Table 9.3: FCC rules for IEEE 802.11 DSSS

A receiver begins with de-spreading the signal. This is done with the help of the same PN sequence. This collapses the desired signal to its original bandwidth due to the correlation of PN sequence and the transmitted signal. The signal power is also increased by the amount of the processing gain. DSSS allows multiple access to the medium since each transmission has its PN sequence and each DS receiver collapses only correlated signals to the data. The block diagram is illustrated in Figure 9.27

DSSS PLCP and PMD Physical Layer

The PLCP header has signal, service, length and CRC fields. Signal fields are to notify the receiver about what modulation is applied. Service field is reserved for future use and the length field is to indicate the length of MDPU. The CRC checks whether the frame is received correctly or not. PMD scrambles all the bits transmitted in order to randomize the data in the SYNC field and data patterns where there may be long strings of repeated 1s or 0s. PLPC Preamble and PLCP Header are modulated using differential binary phase shift keying (DBPSK). The MPDU is transmitted with either 1 Mbps DBPSK or 2 Mbps DQPSK (Differential quadrature phase shift keying).

In the transmitter, the 11-bit Barker word is applied to each of the transmitted bits via modulo-2 adder. The information bits modulated with a rate of 1 Mbps are extended to data rate 10x higher than the

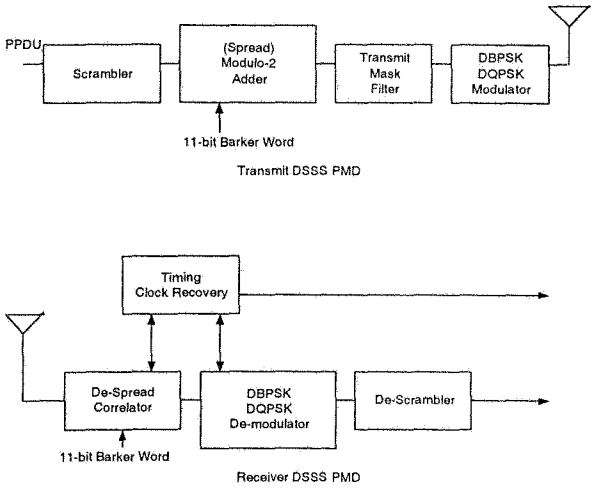


Figure 9.27: Transmit and Receive DSSS © IEEE

information rate when 11-Barker word is at 11 Mbps. This spreading ratio is also called processing gain.

Pros and Cons

The main problem with DSSS is the Near-far effect. This problem is present when an interfering transmitter is much closer to the receiver than the intended transmitter. The correlation of received signal from interfering transmitter and the code can be higher than the correlation of received signal from the intended transmitter and the code. The DSSS system also consumes more power than the FSSS system since it is more complicated. On the other hand, the DSSS has more resistance to interference than FHSS since unless the interference level passes the noise floor, the signal can be recoverable. Otherwise, nothing can be recovered. Transmission time in DSSS is shorter than FHSS since it does not need to wait to change frequencies.

9.6.4 IR Physical Layer

The infrared (IR) based physical layer is another specification introduced by IEEE 802.11. Infrared transmission requires line of sight and is very susceptible to reflection. As a result it has very low range and can be confined to a room since infrared signals can not pass through the walls. IR uses pulse position modulation (PPM) where it keeps the amplitude, pulse width constant, and varies the position of the pulse in time. The position in time represents a symbol. The IR transmits binary data with 1 Mbps and 2 Mbps with 16-PPM and 4-PPM respectively. IR PHY have never pulled interest and no product is developed.

9.6.5 HR/DSSS Physical Layer

IEEE 802.11b, introduced in 1999, implements high rate (HR) extension of the physical layer of Direct Sequence Spread Spectrum (DSSS). HR/DSSS provides two functions. First, the HR/DSSS builds 5.5 and 11 Mbps data rate capabilities in addition to the 1 and 2 Mbps rates. Secondly, the HR/DSSS provides a rate shift mechanism to introduce the functionality of falling back to 1 and 2 Mbps to interoperate with the legacy IEEE 802.11. IEEE 802.11b operates in 2.4GHz ISM band along with cordless phones, microwave ovens, other adjacent WLAN networks, and personal area networks (PANs). Being susceptible to interference is the primary concern for IEEE 802.11 products. The 802.11b uses the same configuration as in DSSS except the modulation and chipping sequence.

Complementary Code Keying (CCK) Modulation

IEEE 802.11b uses eight-chip complementary code keying (CCK) as the modulation scheme to achieve the higher data rates. DSSS system uses Barker sequence which is static but HR/DSSS uses a code word

that is derived partially from data consequently carries information as well as spreading the signal. CCK has mathematical properties that allows them to be correctly identified from one another in the presence of noise and interference.

The spreading code length is 8 and is based on complementary codes. The chipping rates is 11Mchip/s. 1-2 Mbps is achieved with DQPSK same as in DSSS. Transmission at 5.5 Mbps with CCK is transmitting 4 bits per symbol. Two of them are carried by DQPSK and the other two are injected to the code words. In the same way for 11 Mbps transmission, two of them is carried in DQPSK and the other 6 are injected to the code words.

9.6.6 RF Interference

The presence of unavoidable interfering RF signals disturbs the IEEE 802.11 operation. The power of interferer when significant may cause the STA inactive for an indefinite periods until the interferer disappears.

An interfering signal is basically cause an energy jump in radio and may start abruptly in the middle of a transmission causing a collision since CS/CCA mechanism sends busy signal to MAC layer through carrier sense mechanism. If the interference persists, the STA either automatically switches to lower rate or waits until the medium is clear for as long as the interference is there.

Unlicensed 2.4 GHz bandwidth occupies several technologies that can interfere IEEE 802.11 or 802.11b network: wireless phones, microwave ovens, Bluetooth equipped devices. For a WLAN, another WLAN that operates in the same channel is also an interferer. This situation is likely occur in IEEE 802.11 and 802.11b since there is only three non-overlapping channels. The IEEE 802.11a is a solution to avoid interference for the near future since most potential for RF interference today is in the 2.4GHz and IEEE 802.11a has twelve

non-overlapping channels and operates in 5 GHz.

9.7 IEEE 802.11a Physical Layer

IEEE 802.11a is developed to provide high data rate service at 5GHz U-NII bands. IEEE 802.11a selects multi-carrier modulation. Multi carrier modulation is a strong candidate for packet switched wireless applications and offers several advantages over single carrier approaches. OFDM system is a viable solution to accommodate 6-54 Mbps data rates for the following reasons:

Robustness against delay spread: Data transmission in wireless environment experiences delay spreads of up to 800 ns which covers several symbols at baud rates of 10 Mbps and higher. In a single carrier system an equalizer handles detrimental effects of delay spread. When delay spread is beyond 4 symbols, use of maximum likelihood sequence estimator structure is not practical due to its exponentially increasing complexity. Linear equalizers are not suitable for this application either since in a frequency selective channel it amounts to significant noise enhancement. Hence, other equalizer structures such as decision feedback equalizers are used. The number of taps of the equalizer should be enough to cancel the effect of inter-symbol interference and perform as a matched filter too. In addition, equalizer coefficients should be trained for every packet, as the channel characteristics are different for each packet. A large header is usually needed to guarantee the convergence of adaptive training techniques. A multi-carrier system is robust against delay spread and does not need a training sequence. Channel estimation is required however.

Fall-back mode: Depending on the delay spread for different applications a different number of carriers is required to null the effect of delay spread. The structure of FFT lends itself to simple

Parameters	Slot Time	SIFS	PIFS	DIFS	CWmin	CWmax
802.11a	9μs	16μs	25μs	34μs	15	1023

Table 9.4: IEEE 802.11a OFDM PHY characteristics

fallback mode by using the butterfly structure of FFT as shown in Chapter 2.

Computational efficiency: Use of FFT structure at the receiver reduces the complexity. As the number of carriers grows, the higher efficiency can be achieved.

Fast synchronization: OFDM receivers are less sensitive to timing jitter compared to spread spectrum techniques.

Near-far effect Unlike spread spectrum techniques where the signal is detected by correlation, near-far problem does not apply to OFDM receivers since the detection is not based on correlation.

There are 12 channels where each four are in U-NII lower band (5.15-5.25 GHz), mid-band (5.25-5.35 GHz), and upper band (5.725-5.825 GHz) with a channel spacing of 20 MHz. IEEE 802.11a parameters are seen in Table 9.4.

9.7.1 OFDM Architecture

A typical OFDM wireless LAN transceiver is shown in Figure 9.28. We explain the process of IEEE 802.11a transceiver starting from the binary input to the binary output. First, an OFDM symbol is constructed then an OFDM frame in PLCP layer is put into together with the OFDM symbols. PLCP/PMD layer sends the frame into air. On the receiver side, first, the frame is detected and then receiver converts the radio signal into OFDM symbols. OFDM symbols are demodulated and binary data is extracted.

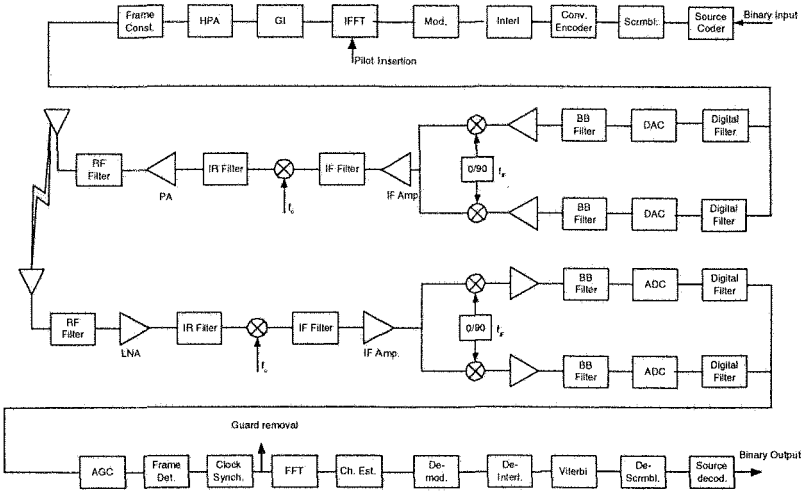


Figure 9.28: Wireless LAN system architecture

9.7.2 Transmitter

Logical block diagram is represented in Figure 9.29. Each block is set of modules that states the changes of binary input up to binary output.

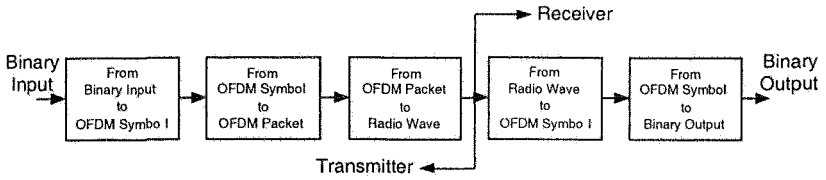


Figure 9.29: Logical block diagram of OFDM architecture

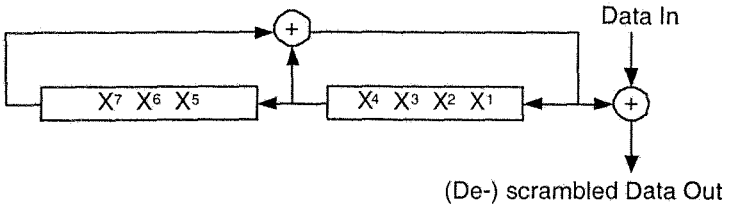


Figure 9.30: Scrambler/Descrambler

(1) From Binary Input to OFDM Symbol

Source Coding

A number can be represented in binary with 1s and 0s. Binary data is converted into an efficient representation where the redundancy of the binary data is reduced by leveraging the correlation between the numbers. Information theory defines entropy as the smallest achievable average length after removing redundancy. Section 2 covers the topic partially and interested reader shall refer to *Elements of Information Theory* for detailed information.

Scrambler

Data may contain long sequence of 1s or 0s. This degrades the performance of receiver when resolving the data. The scrambler helps to change the bit pattern with a 127 pseudo random sequence. The same scrambler are used to descramble the data. When transmitting, the initial state of the scrambler will be set to a pseudo random non-zero state. The 7 least significant bits of the SERVICE field are set to all zeros prior to scrambling to enable estimation of the initial state of the scrambler in the receiver (See Figure 9.30).

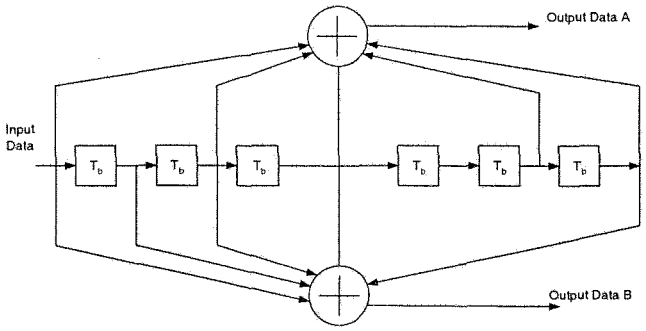


Figure 9.31: Convolutional encoder (k=7) © IEEE

Convolutional Encoder & Puncturing

The channel coding is the crucial part of any communication system. Coding brings reliability at the expense of adding redundancy. Same bit error rate can achievable with a lower signal to noise ratio.

There are two types of coding: block and convolutional. IEEE 802.11a and HIPERLAN/2 WLAN use convolutional error coding. The convolutional encoder is a technique which is suitable for continuous streams. Variable data rates by using different modulation schemes requires variable coding gain for each data rate. Different coding rates for convolutional coding are proposed depending on required protection. Obviously, more crowded constellations need higher coding protection. A combination of error detection, such as CRC, and error correction, such as convolutional coding, in conjunction with interleaving are used to provide better coding gain in the presence of frequency selective fading.

The convolutional encoder is a kind of finite state machine and has registers to represent the state of the system. The coding rate ($\frac{k}{n}$) is defined as the number of registers (k) over the number of output (n). The length of the shift registers makes the system more reliable, but on the other hand, it increases the decoding complexity of Viterbi

Mode	Modulation	Code Rate	Data Rate	BpS
1	BPSK	1/2	6 Mbps	3
2	BPSK	3/4	9 Mbps	4.5
3	QPSK	1/2	12 Mbps	6
4	QPSK	3/4	18 Mbps	9
5	16-QAM	1/2	24 Mbps	12
6	16-QAM	3/4	36 Mbps	18
7	64-QAM	2/3	48 Mbps	24
8	64-QAM	3/4	54 Mbps	27

Table 9.5: Eight PHY Modes of the IEEE 802.11a

algorithm.

For a 1/2 rate coder, each input bit is processed as two coded output bits. Thus, the convolutional coding increases the length of the original message. Figure 9.31 shows the coder used in IEEE 802.11a.

Puncturing is a method to achieve high data rates by just omitting some of the encoded output bits in the transmitter. Higher coding ratios of 2/3 and 3/4 are obtained by puncturing the 1/2 rate code. When 2 out of 6 bits are omitted the resulting rate becomes 3/4 and when 1 out of 4 bits is omitted the result gives a code with a rate 2/3. Viterbi algorithm shows the best possible choice for decoding since algorithm tries the best by maximum likelihood code word estimation. The data is coded with a convolutional encoder of coding rate R, corresponding to the desired data rate as seen in Table 9.5.

Interleaving

Performance of codes designed for AWGN channel rapidly degrades in frequency selective channels with correlated channels. Interleaving is a method to separate bits in time and frequency. The block interleaving is used for a length of one OFDM symbol since the channel is

basically be quasi-static and assumed to stay the same for duration of a symbol. OFDM systems leverage frequency diversity since they are wide-band systems. Block interleaving basically is done by a matrix where the input is written in columns and the output is read in rows. For 6×8 Matrix, the interleaving depth is 48.

Modulation

Instead of sending one bit in one symbol, multiple bits can be sent by converting more than one bits into a complex number bigger than 1 or 0. Modulation schemes have constellation diagrams that maps multiple bits into complex numbers. The type of constellation diagram affects the bit error rate (BER), peak to average power ratio (PAPR) and RF spectrum shape. IEEE 802.11a uses coherent modulation where there is phase lock between transmitter and receiver. This helps receiver to convert complex numbers into bits easily with a complex receiver architecture. Number of data bits per OFDM symbol is calculated from desired data rate and coding rate. For instance for a rate of 54 Mbps, 64-QAM is used with a rate of 3/4. This corresponds to 6 bits en each subcarrier and 6x48x3/4 (27bytes) for an OFDM symbol when number of data sub-carriers is 48. Refer to Table 9.5 for other data rates.

OFDM Symbol

In an OFDM symbol, IEEE 802.11a uses sub-carriers of an IFFT block of length 64. 48 sub-carriers are occupied with modulated symbols and 4 sub-carriers are used for pilot symbols to make the coherent detection overcome frequency offsets and phase noise problems. The pilots are BPSK modulated and inserted in -22, -7, 7, 22 sub-carriers and center subcarrier at 0 is not used as seen in Figure 9.33. The remaining 12 sub-carriers are not used due to the bandwidth limitations.

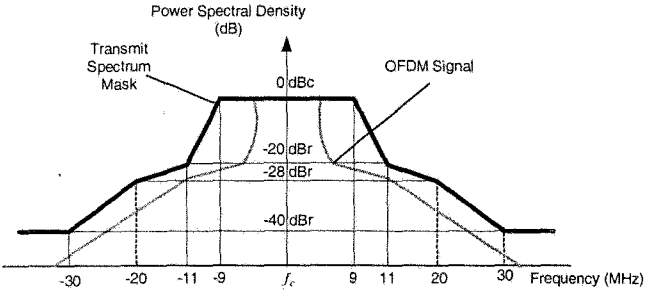


Figure 9.32: Transmitter spectrum mask

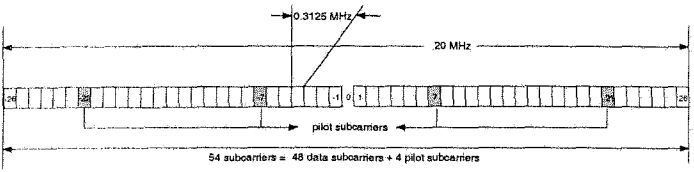


Figure 9.33: OFDM subcarrier allocation

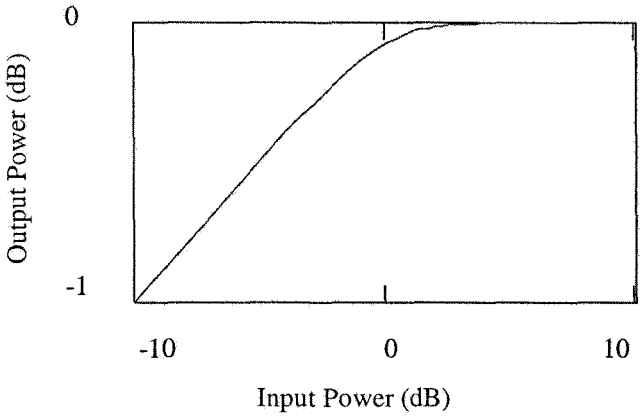


Figure 9.34: Nonlinear HPA model

In the data stream, each 48 modulated symbol is used to make a OFDM symbol. The output of IFFT is prepended with a circular extension to create guard interval (GI). The guard interval is used to avoid ISI from the previous frame. The total OFDM symbol is considered guard interval plus IFFT block. Guard interval is one fourth of FFT length in IEEE 802.11a. Table 9.6 shows the key parameters of IEEE 802.11a. Also, to make the data in multiple of OFDM symbols adequate number of pad bits are appended.

Pulse shaping has an important role in robustness of the system against phase noise and frequency offset as explained in Chapter 5. In addition, pulse shaping provides a smoother transmit signal spectrum which is critical in an OFDM system due to higher peak-to-average ratio. Transmitter output has to meet strict criteria of out-of-band leakage. A spectrum mask, such as the one shown in Figure 9.32, is usually defined to control the effect of bandwidth regrowth caused by amplifier clipping and makes the signal band limited to 30 MHz. A typical amplifier non-linearity model used in simulation and analysis

Parameter	Value
N_{SD}:Number of data subcarriers	48
ND_{SP}: Number of pilot sub-carriers	4
N_{ST}: Number of sub-carriers, total	52
Δ_F: Subcarrier frequency spacing	0.3125 MHz (20 MHz/64)
T_{FFT}: IFFT/FFT period	3.2μs ($1/\Delta_F$)
$T_{PREAMBLE}$: PLCP preamble duration	16 μs (T_{SHORT}+T_{LONG})
T_{SIGNAL}: Duration of the SIGNAL	4.0 μs (T_{GI}+ T_{FFT})
T_{GI}: GI duration	0.8 μs (T_{FFT}/4)
T_{GI2}: Training symbol GI duration	1.6 μs (T_{FFT}/2)
T_{SYM}: Symbol interval	4 μs (T_{GI}+T_{FFT})
T_{SHORT}: Short training sequence duration	8 μs (10xT_{FFT}/4)
T_{LONG}: Long training sequence duration	8 μs (T_{GI2}+2xT_{FFT})
W:Signal Bandwidth	16.66 MHz

Table 9.6: Key parameters of the IEEE 802.11a

is:

$$\begin{aligned} v_o &= \frac{v_i}{(1+(\frac{|v_i|}{v_s})^{2p})^{1/2p}} \\ back-off &= -10log\frac{Avg(|v_i|^2)}{|v_s|^2} \end{aligned} \tag{9.16}$$

where p is smoothness factor and v_s is saturation voltage.

Higher power amplifier (HPA) is used in wireless systems. HPA is modelled by a memoryless nonlinear subsystem, a nonlinear HPA model is shown in Figure 9.34.

(2) From OFDM Symbol to OFDM Packet

The resulting OFDM symbols are appended one after another and data field of a frame is constructed. A typical frame structure is shown in Figure 9.36 and Figure 9.35 shows the frame in OFDM symbols. Framing occurs in PLCP layer and frame structure includes PLCP

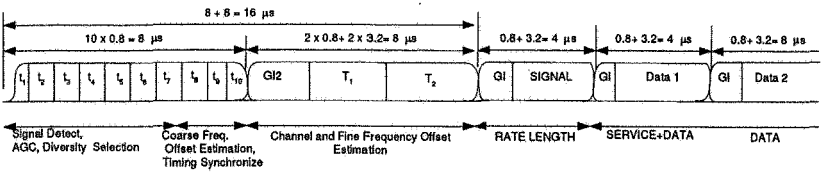


Figure 9.35: Format of an OFDM frame © IEEE

preamble, PLCP header, PLCP data (PSDU and tail and pad bits).

PLCP Preamble

PLCP preamble contains 10 repetitions of a "short training sequence (STS)" and two repetitions of a " long training sequence (LTS)". STS is used for automatic gain control (AGC) convergence, diversity selection, timing acquisition, and coarse frequency acquisition. LTS is used for channel estimation and fine frequency acquisition.

STS consists of 12 sub-carriers as follows:

$$S_{-26,26} = \sqrt{52/(2 \cdot 12)} \times \{ \begin{array}{l} 0,0,1+j,0,0,0,-1-j,0,0,0,1+j, \\ 0,0,0,-1-j,0,0,0,-1-j,0,0,0, \\ 1+j,0,0,0,0,0,0,0,-1-j,0,0, \\ 0,-1-j,0,0,0,1+j,0,0,0,1+j, \\ 0,0,0,1+j,0,0,0,1+j,0,0\} \end{array} \quad (9.17)$$

where $\sqrt{52/(2 \cdot 12))}$ is used for energy normalization. Periodicity of $T_{FFT}/4 = 0.8\mu s$ in the time domain is achieved after taking a 64-point IFFT of the sequence $S_{-26,26}$.

LTS consists of 53 sub-carriers and is generated directly by apply-

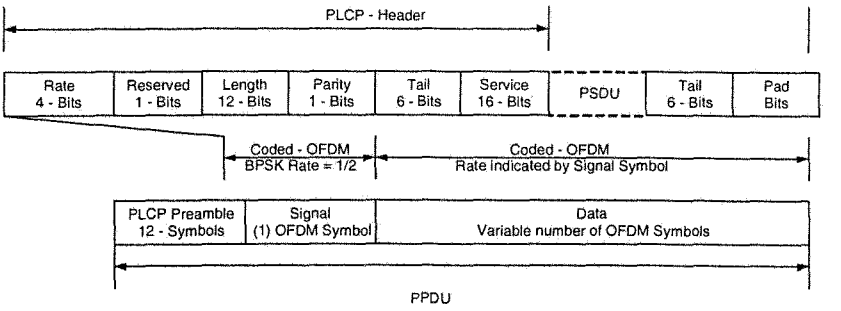


Figure 9.36: Logical representation of an OFDM frame © IEEE

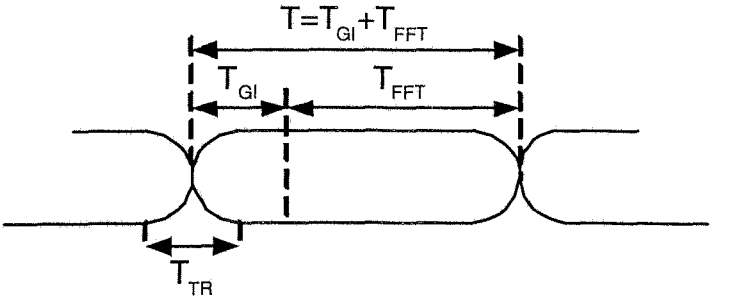


Figure 9.37: Format of an OFDM frame © IEEE

ing the IFFT to the following training sequence:

$$\begin{aligned} L_{-26,26} = \{ \quad & 1,1,-1,-1,1,1,-1,1,-1,1,1,1,1,1, \\ & 1,1,-1,-1,1,1,-1,1,-1,1,1,1,1,0, \\ & 1,-1,-1,1,1,-1,1,-1,1,-1,1,-1, \\ & -1,-1,1,1,-1,1,-1,1,-1,1,1,1,1,1\} \end{aligned} \tag{9.18}$$

Time-windowing function, $w_T(t)$ is applied to provide smoothed

transition. $w_T(t)$ is given by

$$w_T(t) = \begin{cases} sin^2(\frac{\pi}{2}(0.5 + \frac{t}{T_{TR}})) & -\frac{T_{TR}}{2} < t < \frac{T_{TR}}{2} \\ 1 & \frac{T_{TR}}{2} < t < T - \frac{T_{TR}}{2} \\ sin^2(\frac{\pi}{2}(0.5 - \frac{(t-T)}{T_{TR}})) & T - \frac{T_{TR}}{2} < t < T + \frac{T_{TR}}{2} \end{cases} \tag{9.19}$$

where T_{TR} is illustrated in Figure 9.37 and continuous time short OFDM training symbol is defined as

$$r_{short}(t) = w_T(t) \sum_{k=-N_{ST}/2}^{N_{ST}/2} S_k \exp(j2\pi k \Delta_F t) \tag{9.20}$$

PLCP Header

PLCP header field contains fields about rate, length, parity, tail, service. Except the service field PLCP header is sent in one OFDM symbol called SIGNAL and coded with BPSK with a coding rate of R=1/2. Rate and length fields help CCA to determine the busy time of the station even if the station is not the addressee.

PLCP Data

Service field and PSDU (with tail and pad bits) are sent in multiple OFDM symbols at the data rate described in the Rate field.

Finite state machine of the PLCP/PMD structure is shown in Figure 9.38. Following the states, MAC layer notifies PLCP/PMD about a packet transmission. PLCP/PMD layer confirms the transmission and MAC layer sends the data and PLCP/PMD layer notifies the MAC layer after finishing transmission. This ending shifts the receiver from transmit state to the receive state.

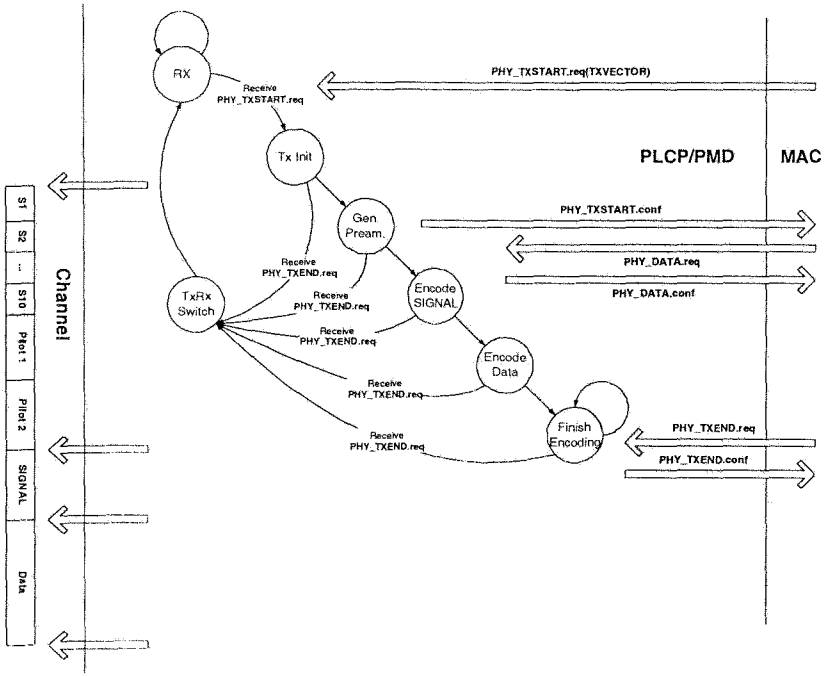


Figure 9.38: IEEE 802.11a PLCP/PMD transmitter state machine

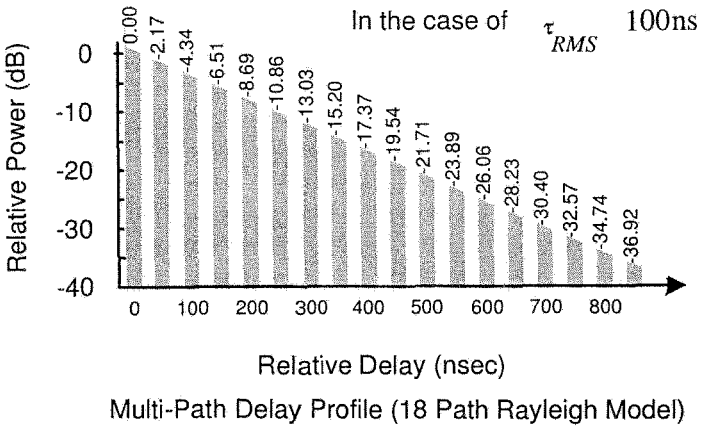


Figure 9.39: Channel impulse response for typical wireless LAN medium

(3) From OFDM Packet to Radio Waveform

Baseband signal is converted into bandpass signal in RF system. The digital data stream is split into two and each stream is converted to analog with a digital to analog converter (DAC) and a baseband (BB) filter is applied. Analog streams are multiplied by sine and cosine waveforms to create the modulated intermediate frequency (IF) signal in-phase and quadrature (IQ) modulator. The IF signal is brought up to RF signal by a local oscillator of f_c frequency and a multiplier. After passing the signal through RF filter to make the signal reside only in the approved bandwidth, signal is relieved into air. The wireless channel model considered for WLAN system is illustrated in Figure 9.39.

9.7.3 Receiver

(4) From Radio Waveform to OFDM Symbol

Frame Detection

Initially, the receiver must detect the frame, adjust AGC to proper level, utilize any diversity capability, and adjust for coarse frequency offset. Fast synchronization is critical in this application. High phase noise and frequency offset of low cost receiver oscillators (with about ± 10 ppm frequency stability) interfere with the synchronization process and require robust algorithm design.

A typical AGC block consists of a correlator followed by a confirmation block. Since received signal level could vary significantly due to shadowing, a low threshold should be used followed by confirmation block, which utilizes the repetitive pattern of preamble sequence to reduce false alarm probability.

Probability of false alarm and detection is a function of threshold and signal to noise ratio. Probability of detection versus false alarm for different thresholds is referred to as Receiver Operation Characteristics (ROC).

If more than one training sequence is used for synchronization the probability of false alarm decreases while probability of detection improves. The overall P_F^t and P_D^t will be

$$\begin{aligned} P_F^t &= \sum_{i=0}^{t} \binom{N}{i} p_f^i (1-p_f)^{N-i} \\ P_D^t &= \sum_{i=t+1}^{N} \binom{N}{i} p_d^i p_m^{N-i}, \end{aligned} \tag{9.21}$$

where N is the number of correlations and t is the threshold determined by required probabilities of false alarm and detection.

Preamble field of OFDM packet is dedicated to facilitate packet detection. The double sliding window algorithm is used to leverage

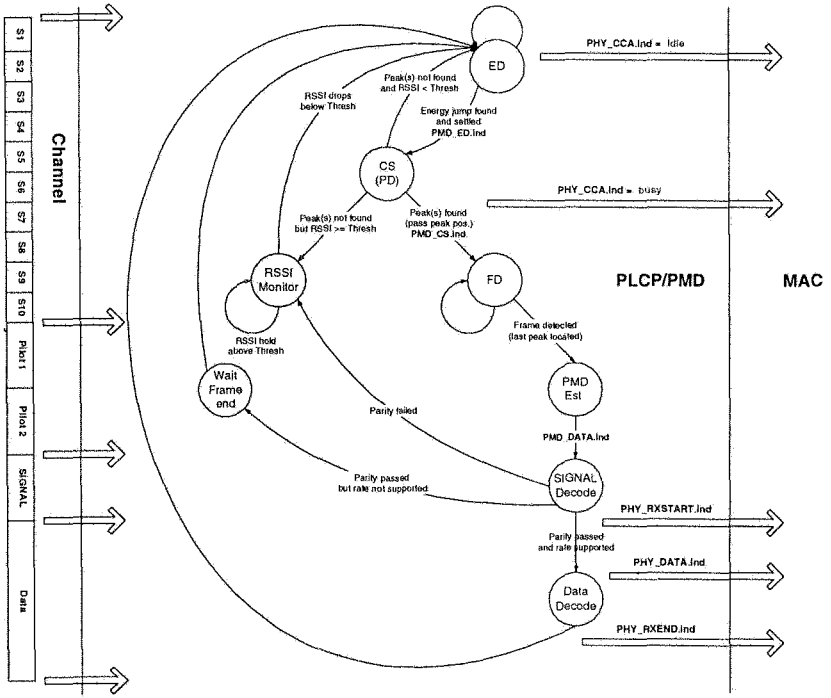


Figure 9.40: IEEE 802.11a PLCP/PMD receiver state machine

the periodicity of short training symbols. The receiver keeps two windows and observes energy. The first window is used to keep the cross-correlation between the signal and its delayed version and the second window is used to keep the power level of the received signal. When there is a packet ratio of first to second shows a jump, otherwise it is approximately equal to zero since the cross-correlation of noise is 0.

Figure 9.40 shows the finite state machine of PLCP/PMD receiver. Energy detection (ED) seeks for a peak and goes to carrier sense (CS) field. If the observed energy jump is higher than the threshold the PLCP/PMD layer sends notification to MAC layer to inform that the channel is busy. If the CS field continues to observe peaks than it con-

cludes that this is a frame and finite state machine passes to frame detection (FD) state. Otherwise, it goes to RSSI state wherein the signal strength of the received signal is observed until the peak goes below threshold. The PLCP/PMD layer starts decoding the signal and sends an ending notification to MAC layer when the process is done.

Symbol Detection

Symbol timing synchronization has the task of finding the precise moment of the OFDM symbol after the packet detector has provided an estimate of the start edge of the packet. Correlation is further exploited to achieve precise OFDM symbol detection. Since the receiver knows the training symbols, any high correlation can be expected as the start of the first OFDM symbol.

Clock Synchronization

A mismatch in sampling instances of the digital to analog converter (DAC) and the analog to digital converter (ADC) causes the detection of slightly shifted versions of the symbols which degrades performance. Pilot carriers inserted symmetrically into the FFT block are used to detect the sampling frequency error and the necessary adjustment is done in the receiver structure.

There are two approaches to do frequency offset estimation: the first one is using the short and long training symbols and the second one is using cyclic prefix. Frequency offset estimation can not perfectly eliminate the offset there remains some residual frequency error resulting in constellation rotation. The main purpose of inserting four pilots is to help the receiver to track the carrier phase.

Frequency offset estimator has an operating range which determines how large frequency offset can be estimated. The limit can be

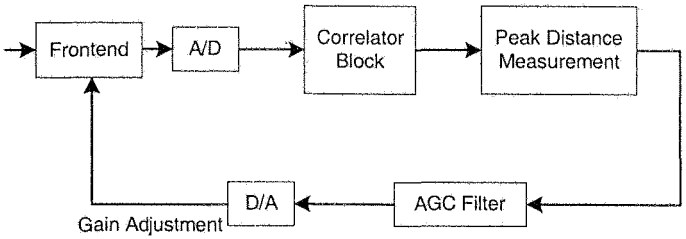


Figure 9.41: Frame synchronization and AGC

defined as follows,

$$|f_\Delta| \leq \frac{N}{2D}\Delta_F \tag{9.22}$$

where $\Delta_F = 0.3125MHz$ is the sub-carrier frequency spacing and D is the delay between the identical samples of two repeated symbols. For the short training symbols (D=16), the limit is 625 KHz (coarse) and for the long training symbols (D=64), the limit is 156.25 KHz (fine). In IEEE 802.11a, the standard specifies a maximum oscillator error of 20 parts per million (ppm) and if the channel is 5.3 GHz then corresponding frequency offset is

$$|f_\Delta| = 40 \times 10^{-6} \times 5.3 \times 10^9 = 212KHz \tag{9.23}$$

which is 156.25 KHz (fine) $<$ 212 KHz $<$ 625 KHz (coarse). If initial frequency offset is very high it may reduce the magnitude of peak and increase probability of false alarm. A frequency offset of f_Δ amounts to peak reduction of

$$\left|\sum_{n=1}^{N} e^{j2\pi nTf_\Delta}\right| = \frac{sin(\pi NTf_\Delta)}{sin(\pi Tf_\Delta)} \tag{9.24}$$

In that case a peak detector can be implemented as a bank of correlators tuned to different frequency offset values. Therefore, frame synchronization, coarse frequency offset estimation and AGC can be implemented using a common structure of Figure 9.41.

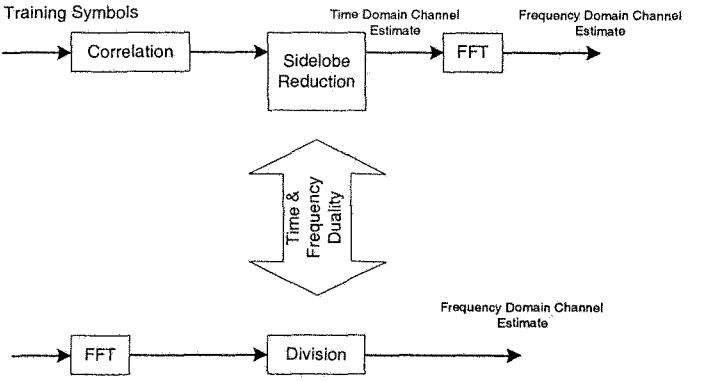


Figure 9.42: Channel estimation block

(5) From OFDM Symbol to Binary Output

Once initial acquisition is accomplished, the detection process begins. The detection process requires channel estimation, prefix removal, frequency domain equalization and soft decoding.

Channel Estimation

In WLAN, channel is assumed to be stationary for the duration of the packet. Long training symbols (LTS) in preamble are used to estimate the channel. This type of estimation is called block type channel estimation where all sub-carriers are pilot symbols. The channel for each sub-carrier can be estimated easily and since there are two LTS, the average of those two gives the possible applicable channel to the frame. Typically, the channel estimation block has the structure of Figure 9.42.

After channel estimation and fine synchronization, guard intervals should be removed. The position of the OFDM block should be adjusted carefully to make sure that the effect of pulse shaping is pre-

served and interference from adjacent blocks are minimum.

Using a reliability index concept enables the performance of soft decoding of the demodulated signal. Due to the presence of an interleaver block, a classic trellis decoding is not possible. We correspond a reliability parameter to each bit and use that in decoding. The reliability index for each bit is derived from the difference in the metric value caused by bit flipping.

After de-interleaving and Viterbi decoding, descrambling and source decoding is applied. This sequence of processes gives the binary output.

9.7.4 Degradation Factors

Some typical degradation factors at the transmitter and receiver are briefly discussed here and performance results are shown in the next section. Quadrature phase error and phase imbalance of transmitter modulation result in a C/N degradation and higher packet error rate. Another source of degradation is the fixed-point error caused by truncating the output of IFFT butterfly. After each stage of IFFT (or FFT) butterfly, the required number of bits grows. Truncation of bits can also increase the packet error rate. Maximum tolerable transmitter degradation is restricted by vector error requirements set forth in wireless LAN standards.

Typical degradation sources in the receiver are phase noise of the local oscillator, which was thoroughly discussed in previous chapters, and D/A quantization error. Overall, transmitter power amplifier nonlinearity and receiver phase noise are dominant factors for total vector error.

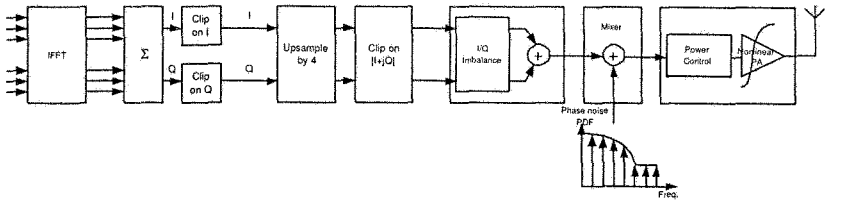


Figure 9.43: Simplified schematic of the IEEE 802.11a simulation model for the transmitter [132]

9.8 Performance of 802.11a Transceivers

IEEE 802.11a is highly sensitive to the impairments that are introduced by front-end components and channel. The components that cause distortion are phase noise, channel, frequency offset, IQ imbalance, quantization and clipping, power amplifier, narrow-band systems, co-channels, ultra-wide band systems. These impairments can lead to an irreducible error floor. They can be investigated with a MATLAB simulator provided by IMEC [134]. The simulator provides a generic front-end digital baseband and a set of multi-path channels corresponding to an indoor environment. Figure 9.43 shows a simplified schematic of the transmitter simulation model and a description of the simulator obtained from [132] is described as follows: " Signals coming out of the IFFT is clipped to reduce large crest factors inherent to DMT modulation techniques. The spectrum of the OFDM signal occupies 16.8 MHz in a 20 MHz channel and the digital modem produces I,Q signals sampled at 20 MHz [132]. A typical power spectrum obtained from the simulator is shown in Figure 9.44 and a FIR filter that is designed to provide 30dB power at the receiver is shown in Figure 9.45. To avoid high sample rates when simulating this bandpass signal at RF frequency of 5.2GHz, simulations are conducted at baseband (BB) using its complex low-pass equivalent representation. Over-sampling by 4 is added in the interface between the modem and the front-end so that nonlinear distortion can be evidenced. In the final

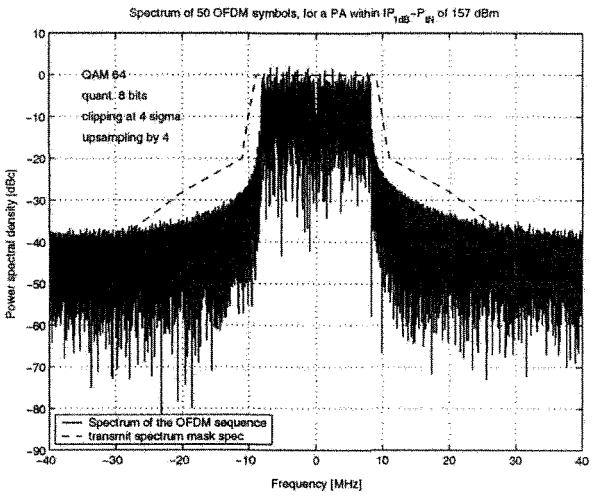


Figure 9.44: An example of power spectrum

steps, analog front-end model consists of three blocks: an I/Q modulator adding I/Q imbalance, a mixer with a local oscillator signal defined by its phase noise power density function, and power amplifier (PA) exhibiting a cubic nonlinearity. Similar BER degradation is produced in the receiver side also. " All the simulation results, unless otherwise noted, are for 8-bit quantization level, 4σ (RMS value of the signal power) clipping level, 115dB amplifier input, AWGN channel, and 100 dB backoff.

9.8.1 Data Rate

The data rate performance of the IEEE 802.11a system is shown in Figure 9.46. Multi-path channel is used to evaluate the modulation schemes in different coding gains where the channel changes in every symbol. Figure 9.47 shows the comparison of the performance in un- coded situation with the theoretical values. Theoretical values are

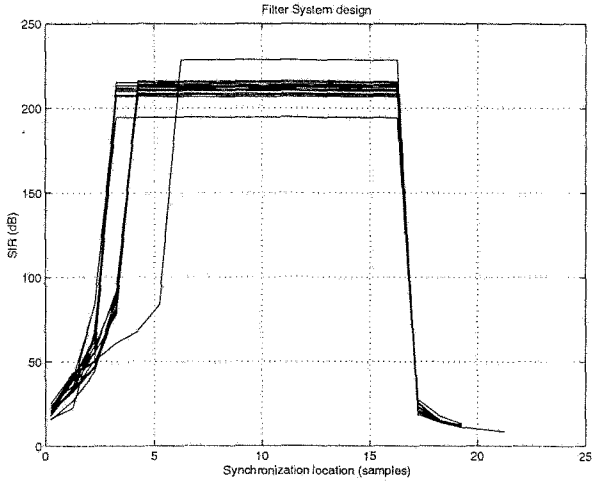


Figure 9.45: An example of FIR filter where the required SIR at the receiver is 30dB

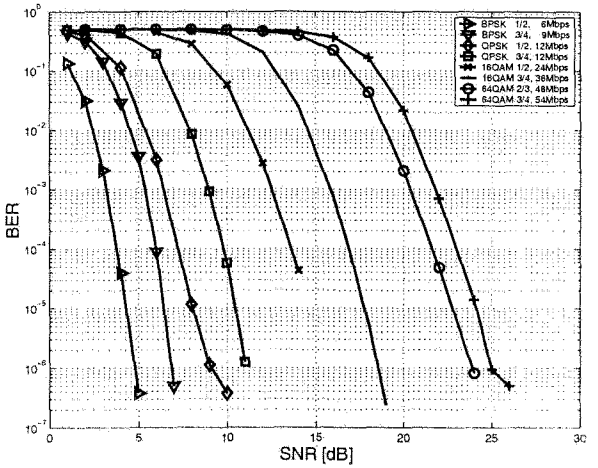


Figure 9.46: Bit error rate of IEEE 802.11a in Rayleigh Channel

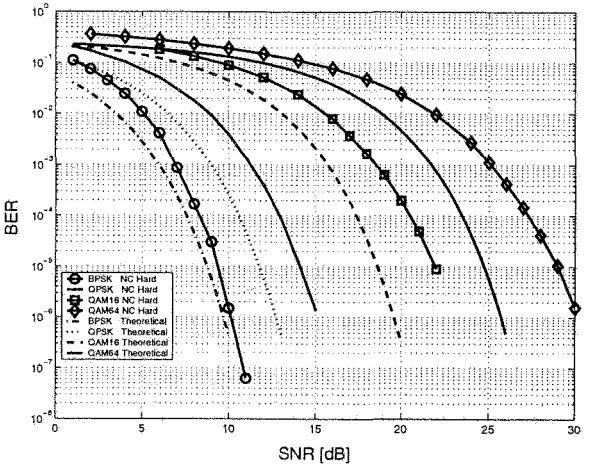


Figure 9.47: Bit error rate comparison of modulation schemes in AWGN

calculated as follows;

$$\begin{aligned} E_b/N_0 &= \frac{W}{R}\frac{N_g}{N_s}10^{\frac{SNR}{10}} \\ BER &= \frac{1}{2}erfc(\sqrt{E_b/N_0}) \end{aligned} \tag{9.25}$$

where signal bandwidth W is $20 * 10^6$, symbol length N_s is 64, and guard interval length N_g is 46.5. Data rate R changes according to modulation; values for BPSK, QPSK, 16QAM, and 64QAM are $12 * 10^6$, $24 * 10^6$, $48 * 10^6$, and $72 * 10^6$ respectively.

9.8.2 Phase Noise

Phase noise has two effects in OFDM: First, a random phase disturbance occurs in each symbol which is known as common phase error (CPE). All carriers suffer the same CPE. Second, inter-carrier interference (ICI) is introduced which takes the form of an additive noise-like signal. ICI is peculiar but CPE error needs to be compensated. CPE

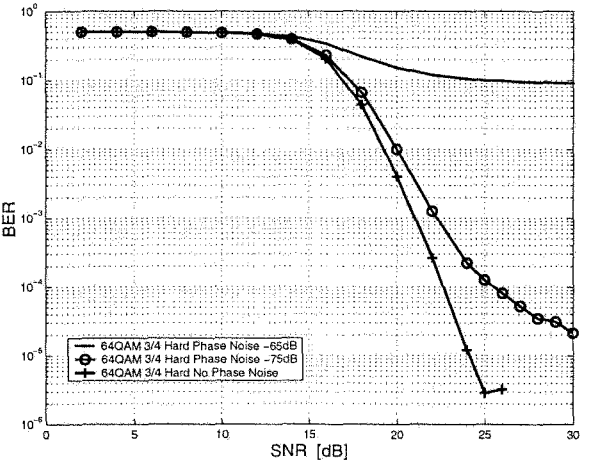


Figure 9.48: Phase noise (65dB of 10KHz)

generates a rotation of the constellation and the error is compensated in the opposite direction if the level of rotation is known.

Figure 9.48 shows the BER performance with phase noise with the following frequency shaping: -65, -75 dB till 10 KHz offset, slope -20 dB/dec, and -135 dBc noise floor. One can see the effect of phase noise from the figure.

9.8.3 Channel Estimation

Channel estimation in the IEEE802.11a is block type. As it is explained in detail in Chapter 6, block type estimation estimates the channel in the first symbol and use that channel information for the rest of the packet.

Figure 9.49 shows the BER performance with or without perfect channel knowledge (PCK). Hard or Soft decision types are investigated under 64QAM (3/4) coded or no coded (NC) modulation. One

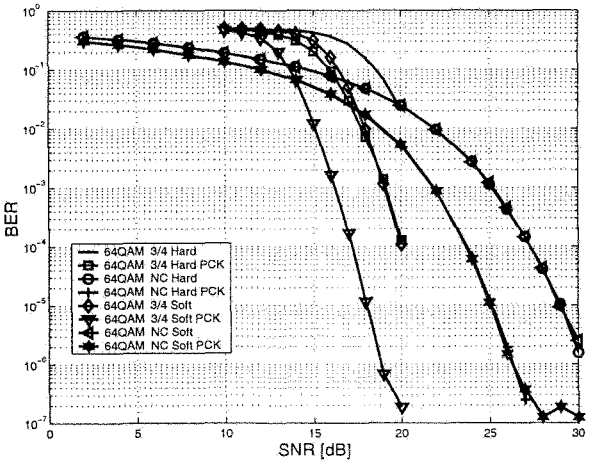


Figure 9.49: Channel estimation compared to Perfect Channel Knowledge (PCK)

can observe the effect of impairment that channel introduces from the figure.

9.8.4 Frequency Offset

One of the main drawbacks of OFDM is its sensibility to frequency offset which is caused by oscillator inaccuracies and Doppler shift induced by the channel. The effect of frequency offset is two fold, it reduces the signal power of the desired sub-channel outputs and also introduces ICI. The ICI power can be significant even for small frequency offsets in OFDM due to the high sidelobe power of the sub-channels.

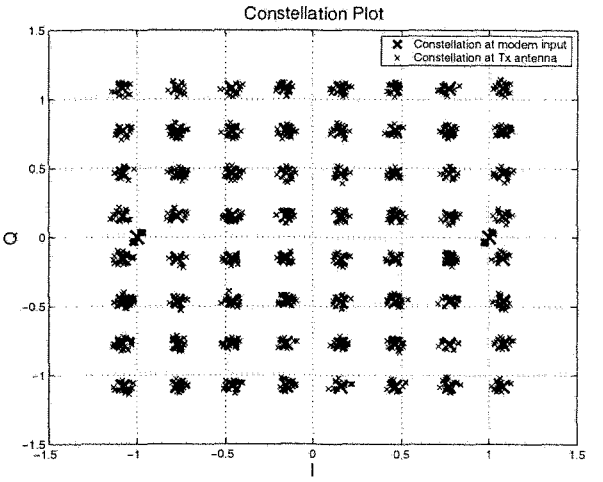


Figure 9.50: An example of constellation diagram with IQ imbalance

9.8.5 IQ Imbalance

OFDM is very sensitive to receiver IQ mismatch since if Zero-IF receivers are used, the IQ demodulation is done at RF which therefore can not be done digitally. A constellation diagram with IQ mismatch is shown in Figure 9.50. [131].

Figure 9.51 shows the BER performance when there is no coding and no perfect channel knowledge. Amplitude (A) and phase (P) mismatch is represented in percentages. Figure 9.52 shows the BER performance when there is coding and perfect channel knowledge. As it can be seen IQ mismatch affects more severely when there is no coding.

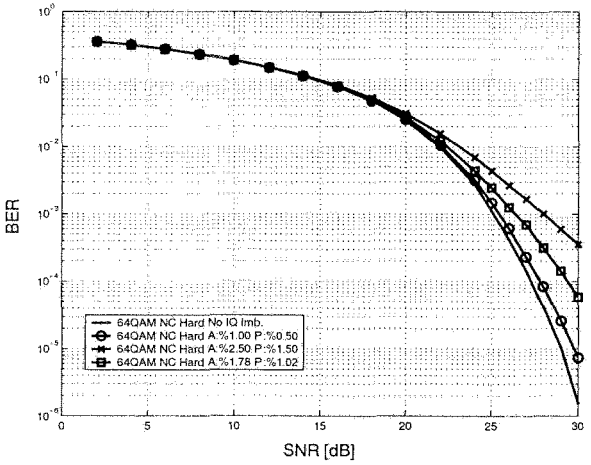


Figure 9.51: IQ imbalance when there is no coding and no perfect channel knowledge

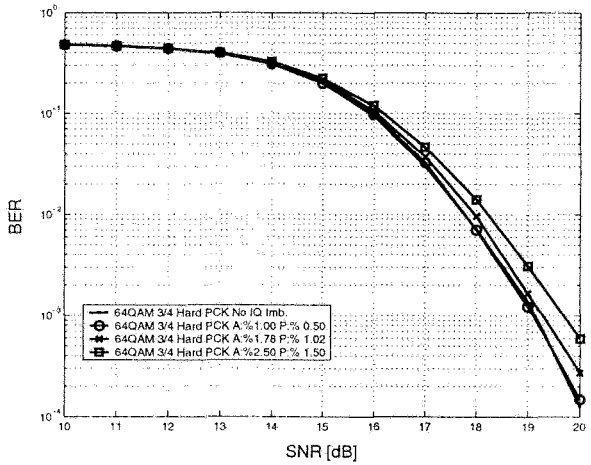


Figure 9.52: IQ imbalance when there is coding and perfect channel knowledge

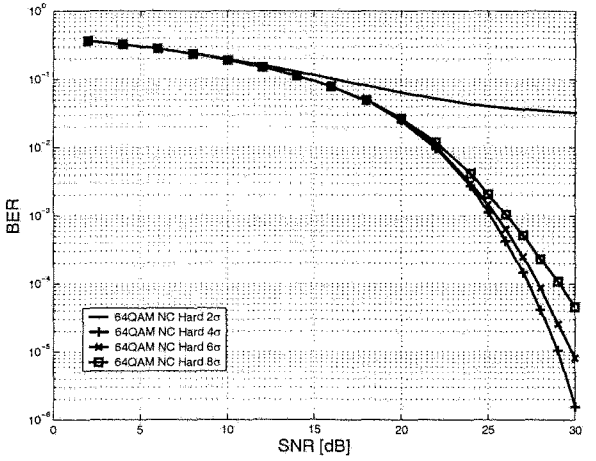


Figure 9.53: Impact of clipping threshold on performance

9.8.6 Quantization and Clipping Error

Quantization level (QL) has a major impact both on implementation cost and performance limitation. As QL decreases the power consumption and the complexity of the D/A and A/D converters decreases at the expense of the quantization noise [132]. QL can be limited by the clipping level thereby optimal QL and clipping level can be found after performing a joint optimization process [132].

Figure 9.53 shows the BER performance for different clipping levels with 8-bit QL. Clipping level is defined in terms of the RMS amplitude σ of the time domain signal. From the figure, one can see that 4σ shows superior performance than 2σ, 6σ and 8σ which fortifies the need to do joint optimization.

Figure 9.54 shows the BER performance under different QL levels and modulation schemes with 4σ clipping level. Quantization error can be observed in the lower QL level. The 8-bit QL has been found optimum with 4σ clipping level [132]. Figures 9.55 and 9.56 shows

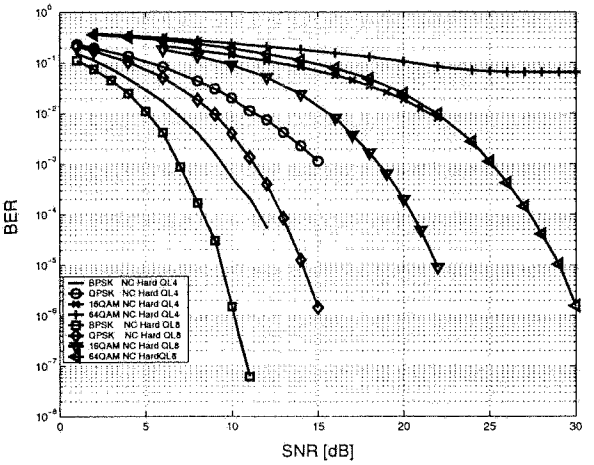


Figure 9.54: Quantization performance in AWGN

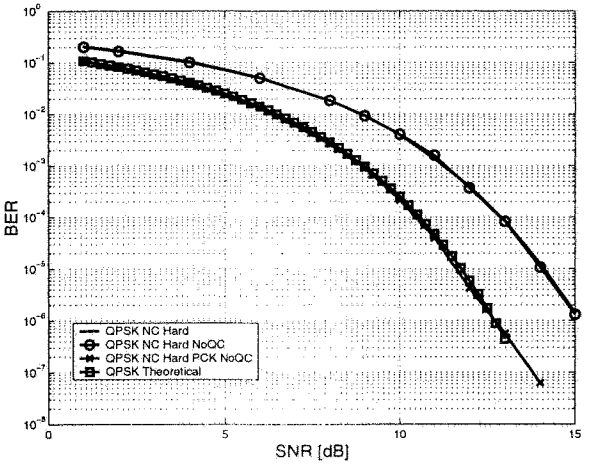


Figure 9.55: Quantization effect in QPSK (AWGN)

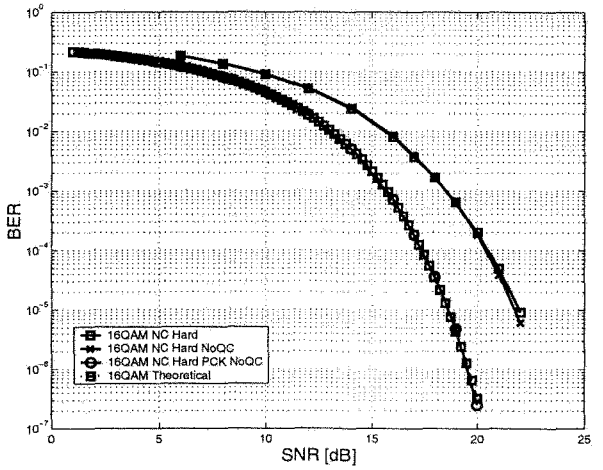


Figure 9.56: Quantization effect in 16QAM (AWGN)

the BER performance of quantization and clipping with no quantization and clipping (NoQC) for QPSK and 16QAM modulations.

9.8.7 Power Amplifier Nonlinearity

OFDM has a sensitivity to nonlinear distortion, which causes crosstalk between sub-carriers. The linearity of power amplifier (PA) is a significant metric for power consumption and BER. The nonlinear amplitude transfer function of a PA with a linear gain G, assuming that the PA is kept out of saturation, was modelled by a cubic nonlinearity:

$$y = G \times (x - \alpha.x^3) \; for x < x_{sat} \tag{9.26}$$

where x_{sat} is the limit of the saturation region and α equals $\frac{4}{(3.IIP3^2)}$ [132]. Figure 9.57 shows the BER performance for different IIP3 values. If IIP3 is bigger than 100dB than linear amplifier is applied.

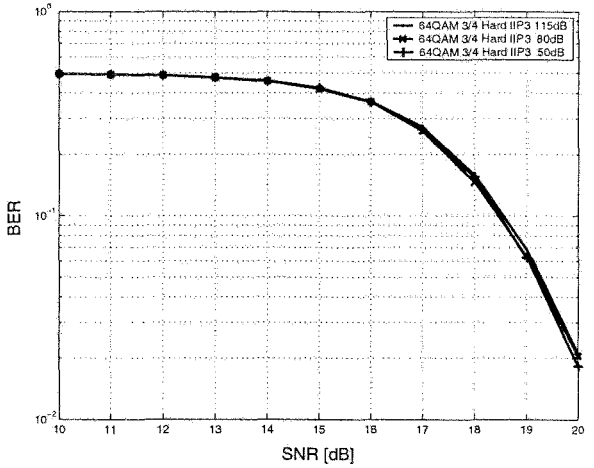


Figure 9.57: Nonlinear distortion in IEEE 802.11a

9.8.8 Hard or Soft Decision Decoding

Hard or Soft decision decoding comparison is shown in Figure 9.58. The performance is compared in coded and un-coded schemes with or without perfect channel knowledge (PCK). As it is seen from the figure, when there is no coding, Hard and Soft decision decoding has almost no effect. When there is coding and perfect channel knowledge, on the other hand, soft decision decoding has superior performance.

9.8.9 Co-channel Interference

Co-channel interference degrades the performance of an OFDM system. Figure 9.59 shows the BER performance of the IEEE 802.11a system when there are 10 adjacent channel. Relative Power of adjacent channels to the signal (AdjCh) is used to compare the effect of co-channels.

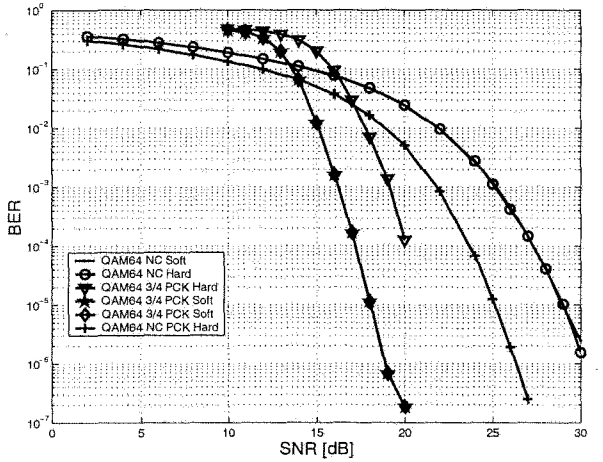


Figure 9.58: Hard or Soft decision coding

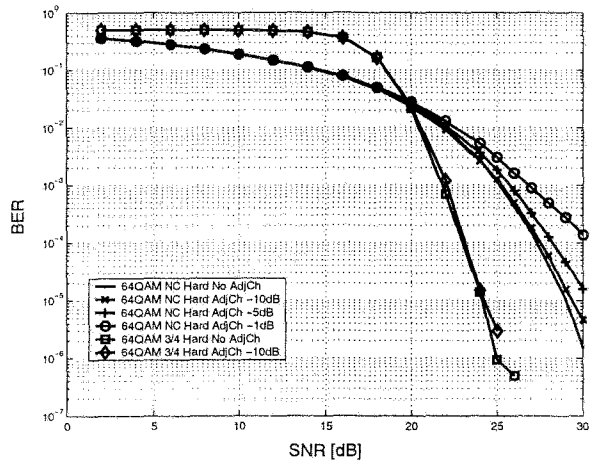


Figure 9.59: Co-channel interference

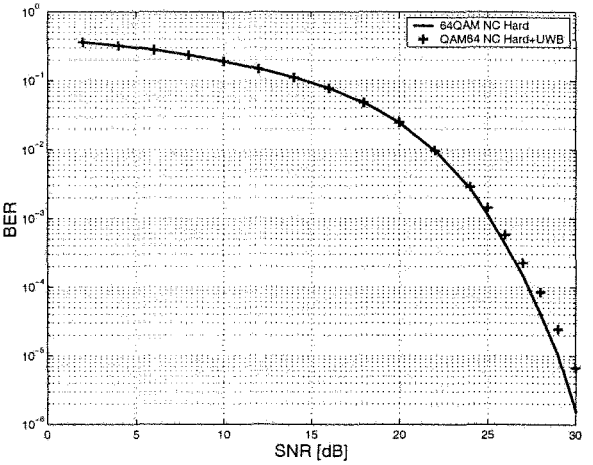


Figure 9.60: UWB interference

9.8.10 Narrowband Interference

The benefit of OFDM is its resistance to narrow-band interference. Since the OFDM waveform is composed of many narrow-band tones, a narrow-band interferer will degrade the performance of the limited portion of the spectrum. The performance degradation can be corrected by forward error correction coding.

9.8.11 UWB Interference

Ultra Wide Band (UWB) operates in an unregulated band where some portion of which also overlaps to the spectrum of IEEE 802.11a. Several interference mitigation methods are proposed.

Using notch filter is a way to filter out the spectrum of UWB pulses to suppress the content in certain bands. Impulsive noise is mitigated more effectively with OFDM. Since the OFDM symbols

are long in duration, energy from a low-level impulse noise can be integrated to a level benign to system performance[133].

UWB interference has an almost negligible effect on the OFDM receiver when it is transmitted in time-hopping patterns. Time hopping patterns distribute the power among the spectrum and power spectral density of the signal becomes almost equivalent to a white Gaussian noise when the number of UWB users is large due to the central limit theorem (See Figure 9.60). Concurrently, the degradation is higher when there is a small number of UWB signals [129]. Time hopping sequence can only be applied to impulse radio systems.

The multi-band approach is another solution. The multi-band approach divides the spectrum into sub-bands (See Chapter 12 for detail). As a result, sub-bands can be turned off to avoid interference or it can be combined with the above-mentioned methods in each sub-band [130].

9.8.12 Performance of 64QAM

Figures 9.61, 9.62 and 9.63 show BER performance of various components of an IEEE 802.11a system: (3/4) coded or non coded (NC) schemes compared with or without perfect channel knowledge (PCK). (Soft) or (Hard) decision decoding are used to investigate the effect of quantization and clipping with no quantization and clipping (NoQL). IQ imbalance (IQimb.) is characterized by amplitude mismatch of 1.78% and phase mismatch of 1.02 degrees. (Lorenz.) indicates the Lorenzian phase noise which is used in the local oscillator by a total integrated noise of -32dBc. (BO54) refers to back-off between the average power and the input power 1dB-compression point of 5.4dB and (pnc) is for phase noise compensation.

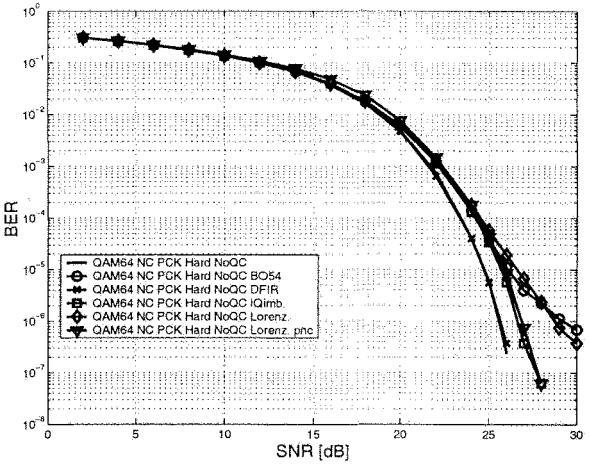


Figure 9.61: Impairments in 64QAM when there is no coding, and perfect channel knowledge

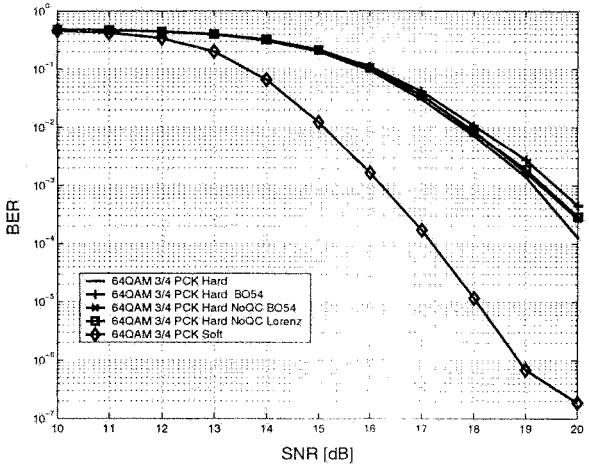


Figure 9.62: Impairments in 64QAM when there is coding and perfect channel knowledge

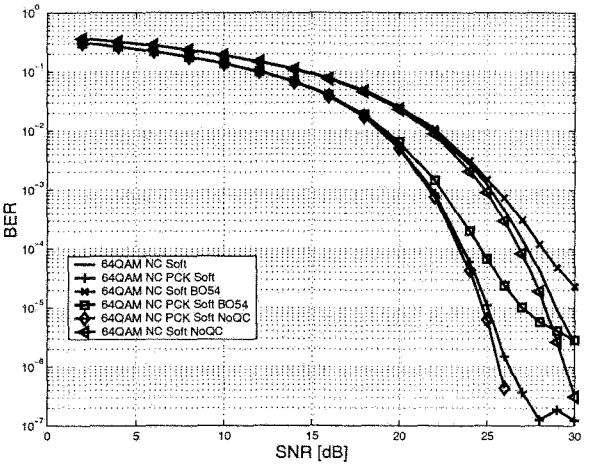


Figure 9.63: Impairments in 64QAM when there is no coding and soft decision decoding

9.9 Rate Adaptation

Rate Adaptation, the mechanism to select one out of available transmission rates at a given time, is important to improve the performance of WLAN. IEEE 802.11b operating in 2.4GHz supports four PHY rates; 1,2,5.5 and 11 Mbps and IEEE 802.11a operating in 5GHz supports eight PHY rates; 6,9,12,18,24,36,48, and 54 Mbps. There is a trade of between robust communication in lower rates and high speed communication in higher rates. The optimum performance is achieved making the rate selection adaptive with regard to channel conditions.

The standard leaves the rate adaptation out of scope. It is rewarding to give some examples. AutoRate Fallback (ARF) protocol alternates between 1 to 11 Mbps based on the timing function and missed ACK frames. After not receiving two consecutive ACKs, the data rate is reduced from default 11 Mbps to 5.5 Mbps and timer is reset. If there is a failure again, data rates falls back from 5.5 Mbps. When the

timer expires or the number of received ACKs reaches 10, then the transmission rate is raised to the higher data rate and the timer is cancelled. However, if an ACK is not received for the very next packet, the transmission rate is lowered again and the timer is reset.

Receiver-Based Auto-Rate (RBAR) is a proprietary mechanism which is based on RTS. Receiver selects a rate based on the signal strength of received RTS and notifies the transmitter by CTS frame. Then, the transmitter responds with data packet in the chosen rate.

The control, broadcast, and multi-cast frames are sent with rates defined in basic rate set and it should be compliant with all transceivers.

9.10 Zero IF Technology

The popularity of mobile connectivity leads to the integration of wireless LAN into notebooks, PDAs, and other portable computing devices. The Zero Intermediate Frequency (ZIF) transceiver is a solution to apply direct conversion. ZIF utilizes one mixer stage to convert the desired signal directly to and from the baseband without using any IF stages and external filters. The ZIF technology brings smallest size, highest sensitivity, lowest power and lowest cost which are very crucial for future mobile devices.

The DSSS and OFDM are strong candidates for ZIF since the signal to be processed does not have a significant DC component. DC component is very important since the circuits also introduce a DC component and can cause the RF stages to saturate. Thus, it would be hard to distinguish if the signal also has one.

9.11 IEEE 802.11e MAC Protocol

IEEE 802.11e is a supplementary to the MAC layer to support LAN applications with Quality of Service (QoS) requirements over IEEE 802.11 wireless LANs. 802.11e is designed as complementary to the 802.11 physical standards such as a,b and g. The classes of service is offered with a QoS management scheme where the applications are considered as data, voice and video.

The QoS facility is available to QoS enhanced stations (QSTAs) and QoS enhanced access point (QAP) in QoS BSS (QBSS) or in QoS IBSS (QIBSS). The standard improves the legacy DCF and PCF and introduces Enhanced Distributed Coordination Access (EDCA) and HCF Controlled Channel Access (HCCA) where HCF stands for Hybrid Coordination function. EDCA implements a contention based access mechanism. EDCA priorities the traffic and varies the amount of time a station would wait before transmission depending on the traffic type. HCF introduces hybrid coordinator (HC) which coordinates a reservation of transmission opportunities. A wireless station (WSTA) that is not attached to QAP, requests the HC for the transmission opportunity. HC schedules a transmission opportunity for the WSTA and deliver the frames based on a polling mechanism. A non QoS STA is allowed to associate in a QBSS and interact in legacy 802.11.

Distribution service introduces a traffic differentiation service in which the packet transmission is controlled based on the traffic class it belongs. This QoS functionality shall make IEEE 802.11e WLAN be a part of a large QoS network where end-to-end delivery may be performed or a last hop of the network where QoS is provided within its boundary.

Priority	Access Category (AC)	Designation (Informative)
1	0	Best Effort
2	0	Best Effort
0	0	Best Effort
3	1	Video Probe
4	2	Video
5	2	Video
6	3	Voice
7	3	Voice

Table 9.7: Priority access category mappings for IEEE 802.11e

9.11.1 IEEE 802.11e MAC Services

MAC service provides connectionless exchange of MSDUs. In legacy 802.11, the packet exchange is based on best effort service but in 802.11e, the QoS facility involves in a per MSDU basis and identifies the traffic. Depending on the traffic priority level, the MAC deliver MSDUs with a corresponding user priority. Traffic identifier gives zero priority to an MSDU which belongs to a non-QoS STA and QoS STA.

There are eight user priorities (UP) defined in the standard as seen in Table 9.7. Either each MSDU has a UP value or traffic specification (TSPEC). Traffic identifier (TID) field values, 0 through 7, are designated for UP. TID field values, 8 through 15, are interpreted as traffic stream identifiers and designated to TSPEC. Outgoing MSDUs are handled with priority parameter values 8 through 15 in accordance with the user priorities 0 through 7. MSDUs are permitted to reorder in the MAC layer which allow to implement priority. The general packet format depicted in Figure 9.64 includes a 2 octets QoS control field. The maximum MSDU size is 2304 octets which determines the maximum frame body, minus any encryption overhead.

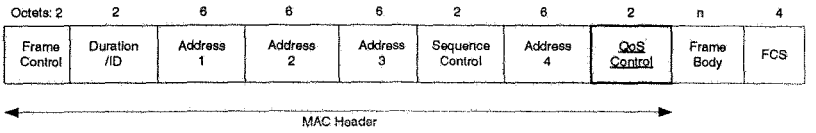


Figure 9.64: MAC frame format © IEEE

9.11.2 IEEE 802.11e MAC Architecture

The MAC architecture is shown in Figure 9.65. The MAC protocol introduces EDCA and HCF that can work with DCF and PCF.

9.11.3 Hybrid Coordination Function (HCF)

HC is a coordination function in QBSS. HCF introduces two access mechanisms: contention-based and polling based access mechanisms. Contention based channel access is called enhanced distributed channel access (EDCA) and contention free transfer is called controlled channel access (HCCA) mechanism. Transmission opportunity (TXOP) is acquired by using one or both of the access mechanisms by QSTA. Depending on the access mechanism used the TXOP is called either EDCA TXOP or polled TXOP.

HCF Contention-based Channel Access (EDCA)

EDCA provides a contention based differentiated access. Contention parameters varies according to user priorities. There are eight user priorities (UP) which define the access categories (AC). The idle time now is not constant (DIFS) but equals to AIFS[AC]. In the same way CWmin and CWmax are not fixed per PHY but variable and equal to CWmin[AC] and CWmax[AC] depending on the AC. In the contention, the higher AC valued station receives the TXOP and shall

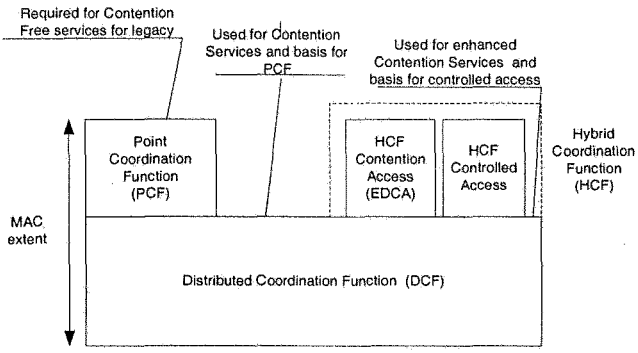


Figure 9.65: MAC architecture © IEEE

transmits multiple frames as long as it does not exceed the EDCA TXOP limit [AC]. A typical situation is depicted in Figure 9.66.

Figure 9.67 is reference model of mapping the frames into access categories. Each access category contains a queue, channel access function, and medium occupancy timer. When a channel access function gets the transmission opportunity (TXOP) it sets the medium occupancy timer to a value defined in MIB. As long as the channel access function continues to use of TXOP, medium occupancy timer is not reloaded but continues to count down to zero. The EDCA access mechanism is similar to DCF. But now a station has more than one channel access function each contend for the channel. Before a transmission, channel access function look for the channel idle for a time greater than AIFSD[AC][9]+Slot time, or a time greater than EIFS-DIFS+AIFS[AC] if the previous transmission was in error. If there is transmission in the medium the channel access function backoffs. The backoff timer is decremented after detecting the medium is idle for a slot time or AIFSD[AC] from the last indicated medium busy or a period of EIFS-DIFS+AIFSD[AC] if the previously received frame was erroneous. Continuation of TXOP is given to the frame of the same

[9]AIFSD[AC]=AIFS[AC]+Slot time+ SIFS

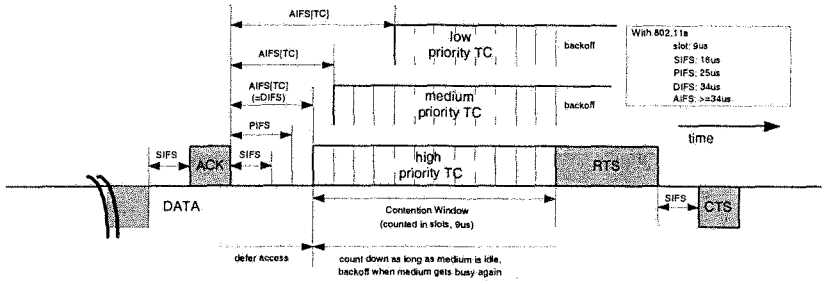


Figure 9.66: EDCA mechanism

access category as the transmitted frame.

Backoff procedure has a finite state machine. Each channel access function shall maintain a state variable CW[AC], and initialized to CWmin[AC]. The backoff procedure is invoked when the frame with that AC is requested to be transmitted and the medium is busy or when a frame with that AC is transmitted and TXOP is terminated or when a frame with that AC fails or when there is higher AC, another channel access function in the same STA, which is granted access at the same time. If there is a failure or if a higher AC grant the access, the CW[AC] is updated as long as it is less than the CWmax[AC] and as long as the retry limits are not exceeded. The update procedure is as follows:

$$CW[AC] = (CW[AC] + 1) * 2 - 1$$

when the backoff procedure is invoked the backoff time is chosen randomly between (1, CW[AC]).

The backoff timer is decremented after a slot time if the medium is idle or after AIFS[AC] if the medium is busy or after EIFS - DIFS + AIFS[AC] + SlotTime if the last frame was erroneous or after a ACK timeout interval if there is no ACK requiring transmission ini-

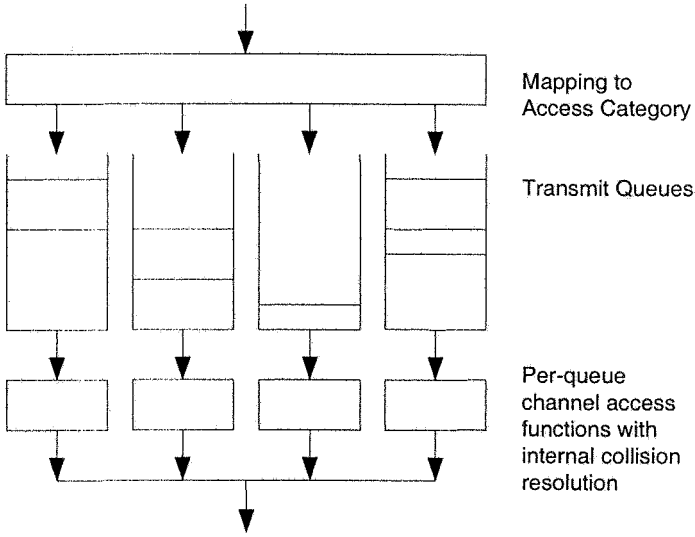


Figure 9.67: Reference implementation model © IEEE

tiated from the channel access functions of this QSTA. There might be internal collision between channel access functions which leads to increase in one of the retry counts.

The retransmission procedure is different from the legacy retransmission procedure. In the legacy transmission procedure if the transmission fails, STA or AP attempts to retransmit as long as the retry limits allow. In 802.11e if the STA or AP uses the non QoS data types, the same rule applies but if the STA or AP uses the QoS data types, it has right to initiate a transmission attempt of a frame of an access category which does not belong to the access category of the frame which is failed previously.

9.11.4 HCF Controlled Channel Access

The HCF uses a point controller called Hybrid controller (HC) and resides in QoS enhanced access point (QAP). HC is similar to PC but with additional functionality. HC provides TXOP to WSTAs in the called controlled access phase (CAP) to transfer QoS data. Contention free period (CFP) and contention period (CP) alternates in HC operation within a superframe. A non QSTA shall operate in CP using DCF access method. HC has the functionality to grant TXOP with duration specified to a non-AP QSTA. This allows a non-AP QSTA to transmit multiple frame exchange.

HC accesses to the WM by starting a CFP or a TXOP in CP. CAP procedure in conjunction with CFP and CP is depicted in Figure 9.68.

If there is a reception of an expected response during the first slot time following SIFS, the HC may initiate by transmitting at a PIFS after the end of the last transmission or non-AP QSTA shall initiate recovery by transmitting at a PIFS after the last transmission [120].

During the TXOP, NAV protects the transmission by suspending the all stations except the transmitting one. QSTA shall use the QoS

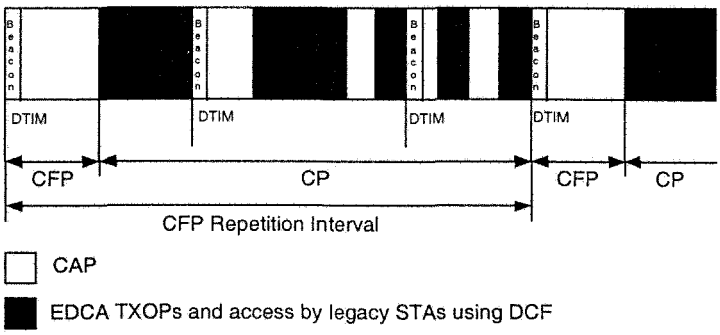


Figure 9.68: CAP/CFP/CP periods © IEEE

control fields in the packet to indicate the traffic belonging of the packet. Non-AP QSTAs may also request TXOP by setting the appropriate fields in the packet.

9.11.5 Admission Control

Admission Control Unit (ACU) is left open in the standard. ACU shall be implemented to control the admission criteria to a given request in polling. A reference implementation shall calculate the Scheduled Service Interval (SI) and TXOP duration for a given SI. The calculation of the Scheduled Service Interval can be done as follows: the scheduler determines the minimum of all Maximum Service Intervals for all admitted streams and takes closest number which is a multiple of beacon intervals. If ρ is mean data rate, L is nominal MSDU size and R is the physical transmission rate then

$$TXOP_i = max\left(\frac{N_i \times L_i}{R_i} + O, \frac{M}{R_i} + O\right) \qquad (9.27)$$

where $N_i = \lceil \frac{SI \times \rho_i}{L_i} \rceil$, O is overhead in time units, M is maximum allowable size of MSDU, i.e., 2304 bytes.

When a new stream is to be added, the criteria is as follows

$$\frac{TXOP_{k+1}}{SI} + \sum_{i=1}^{k} \frac{TXOP_i}{SI} \leq \frac{T - T_{CP}}{T} \tag{9.28}$$

where k is the number of existing streams and T and T_{CP} indicates the beacon interval and time used for EDCA traffic respectively. If the inequality in Equation 9.28 holds, the new arriving stream is admitted.

9.11.6 Block Acknowledgement

Block acknowledgement is introduced to save bandwidth by acknowledging multiple frames with one ACK. Block acknowledgement is granted after an exchange of request/response frames. BlockACK control frame is sent after receiving blocks of QoS data type frames.

9.11.7 Multi-rate Support

A dynamic rate switching capability is allowed in the standard. 802.11e defines set of rules for the possible algorithms. All control, multi-cast or broadcast frames except the ACK frames shall be transmitted at a rate that belongs to basic rate set. Data, ACK, management with unicast receiver address shall be sent in any rate that is being agreed in both sides.

9.11.8 Direct Link Protocol

IEEE 802.11e introduces direct link protocol (DLP) which allows to transmit frames directly to a STA from another STA. DLP is established by three way handshake. STA (A) which is willing to communicate STA (B) sends request to AP which forwards that request to STA (B). If STA (B) accepts the direct stream, STA (B) replies back

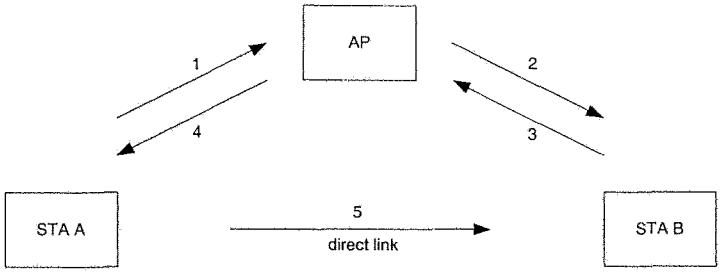


Figure 9.69: Direct Link handshake © IEEE

to AP and AP forwards the response to STA (A) (See Figure 9.69). AP may also pair them without any request. STA (B) shall not go into power save as long as there is a packet transmission or for a duration of DLP idle time out period.

DLP is useful if the recipient is in power save mode and needs to be awakened. DLP enables more secure sessions between the sender and the receiver, thereby the rate set and other information can be exchanged.

9.12 HIPERLAN/2

A parallel converging line of activities in WLAN is conducted by European Telecommunications Standards Institute (ETSI) under Project Broadband Radio Access Networks (BRAN). HIPERLAN is one of the BRAN standard and designed to provide high-speed access to networks.

HIPERLAN/1 is designed for ad-hoc networks. HIPERLAN/1 operates in 5.1-5.3 GHz bandwidth and does not include spread spectrum technology. Data rates goes up to 23.5 Mbps in 5 fixed channels. The MAC protocol is based on a variant of CSMA/CA. It has centralized MAC and supports QoS functions. It has ad hoc routing capabil-

ity to manage intermediate stations act as relays. HIPERLAN/1 does not guarantee QoS and considers only best effort service. This is what motivated ETSI to develop HIPERLAN/2.

HIPERLAN/2 is an OFDM WLAN standard. HIPERLAN/2 is a counterpart of HIPERLAN/1 and focuses for managed infrastructure and wireless distribution system. The primary difference between IEEE 802.11a arises in MAC layer. HIPERLAN/2 MAC is a time division multiple access/time division duplexing (TDMA/TDD) protocol. As a result, HIPERLAN schedules the access in a deterministic manner but IEEE 802.11 schedules in random.

HIPERLAN has attractive features as compared to IEEE 802.11. Although the maximum data rates are same in both standards, the goodput, the true usable maximum throughput, of HIPERLAN/2, however, is 42 Mbps while the goodput of IEEE 802.11a is 18 Mbps for a average size of 512 bytes packets.

HIPERLAN/2 has already designed to support QoS in different information types such as data, video, and voice. Connection oriented feature of HIPERLAN/2 makes it straightforward to implement QoS. Each connection can be assigned different specific QoS parameters such as bandwidth, jitter, BER, etc.

Automatic frequency allocation has a built-in support in AP. There is no need for manual frequency planning. The mechanism works as follows: AP monitors the medium and selects the most appropriate channel that minimizes interference.

The heterogenous networking is noble to seek and HIPERLAN/2 technology has the ability to integrate with 3G, ATM, and other legacy backbone networks and has been designed to provide high-speed seamless access to them.

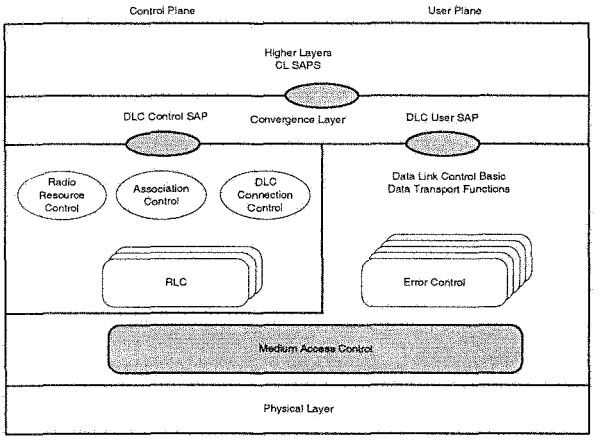


Figure 9.70: HIPERLAN/2 MAC protocol stack © ETSI-BRAN

9.12.1 Protocol Architecture

Figure 9.70 shows the protocol model. The HIPERLAN/2 has three layers; Physical Layer (PHY), Data Link Control Layer (DLC), and the Convergence Layer (CL). PHY layer is OFDM layer and mainly resembles IEEE 802.11a. HIPERLAN uses different scrambler and training, has extra puncturing and 400ηs guard interval.

9.12.2 Data Link Control

DLC is a set of sub-layers: Medium Access Control (MAC), Error Control (EC), Radio Link Control (RLC), DLC Connection Control (DCC), Radio Resource Control (RRC), Association Control Function (ACF).

MAC protocol

MAC protocol defines the access mechanism to the medium. HIPERLAN/2 uses TDMA/TDD scheme. There are two different access mechanisms; centralized and direct mode. The centralized mode attaches HIPERLAN/2 to an infrastructure. AP is the one that offers resources and all the frames are sent through the AP. Direct mode is also referred as ad-hoc mode. A station is selected central controller (CC) to do the resource sharing. Although the resource request are sent to CC, stations can directly send data to each other without passing the frames through CC.

The time slotted structure of the medium allows for simultaneous communication in both downlink and uplink within the same time frame. The AP and CC is informed about the state of the buffers in the MT by Resource Request (RR) messages. The allocation of resources are conveyed by Resource Grant (RG) messages.

The standard defines informative logical and transport channels to improve readability and to define the structures of MAC frames. The mapping between logical and transport channels are showed in Figure 9.71.

Logical Channels

Logical channels identify the various data paths. Each logical channel is defined for a different kind of data transfer service. Logical channels are considered to operate between logical connection end points and logical entities.

1-Broadcast Control CHannel BCCH is used in downlink to convey broadcast control channel information.

2-Frame Control CHannel FCCH is sent downlink by resource grants (RGs) to describe the structure of the MAC frame.

3-Random Access Feedback CHannel RFCH is a feedback channel which informs the stations about the previous transmission attempt at Random Channel (RCH).

4-RLC Broadcast CHannel RBCH is used to convey information concerning the whole radio cell when necessary.

5-Dedicated Control CHannel DCCH conveys RLC sub-layer signals between an MT and the AP. DCCH forms a logical connection and is established implicitly during association with a terminal.

6-User Data CHannel UDCH conveys user data between the AP and an MT. User Broadcast Channel (UBCH) and User Multi-cast Channel (UMCH) are also for broadcast and multi-cast packets respectively.

7-Link Control CHannel LCCH conveys information between the error control (EC) functions in the AP and an MT for a specific UDCH.

8-Association Control CHannel ASCH is only used in the uplink and conveys new association request and handover request on behalf of the RLC.

Transport Channels

The transport channels are that describe the basic message format. Transport channels also take care of the random access method and collision resolution although the message content and their interpretation belong to logical channels.

1-Broadcast CHannel BCH carriers BCCH that carries control information that is sent in every MAC frame to mainly enable some RRC functions.

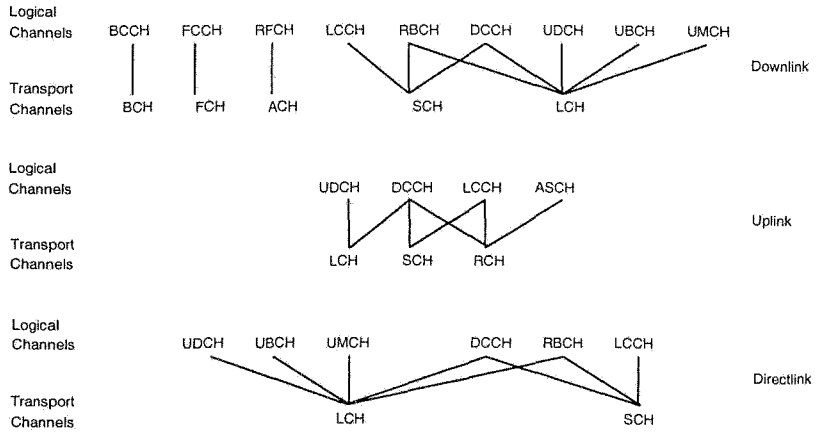


Figure 9.71: Mapping between logical channels and transport channels © ETSI-BRAN

2-Frame CHannel FCH carriers FCCH that contains an exact description on allocation of resources.

3-Access Feedback CHannel ACH shall be used for sending RFCH which conveys information on previous random access attempts.

4-Long Transport CHannel LCH transports user protocol data unit. (U-PDU).

5-Short Transport CHannel SCH transports short control information (C-PDU).

6-Random CHannel RCH gives an MT the opportunity to send control information to the AP or CC.

Transmission

After the MT initiates connection with AP, MT receives a unique MAC identifier. In order to transmit data MT initially request slot by

sending a resource request (RR). Information about the pending User Protocol Data Units (U-PDU) is sent with an RR. The contention slots in the RCH is used to send the RR. This is a control mechanism for AP to manage the delay. AP gives transmission right by sending resource grant (RG). AP shall also poll MT to get information from time to time.

DLC connection identifier represents the connection for unicast, multicast, or broadcast sessions. A connection is distinguished by MAC identifier and DLC connection identifier where this combination is also called DLC user connection (DUC). For unicast traffic, each MT is allocated a MAC identifier and one DLC for each connection. For multi-cast traffic, there is a separate MAC-ID allocated for the group. Each U-PDU is transmitted once without ARQ. For broadcast traffic, transmission without ARQ can be repeated multiple times to increase probability of a successful transmission.

Error Control

EC is to detect erroneous frames and frames not in-sequence. EC applies ARQ and gives each transmitted U-PDU a sequence number and seeks for ACK in signalled LCCH. Erroneous U-PDU is retransmitted. For real time data, there is a mechanism to discard outdated U-PDU. This lacks data but retains connection alive.

Radio Link Control

The RLC protocol provides a transport service for the ACF, RRC, DCC. These protocols define the DLC control plane responsible for signalling between AP and MT.

Association Control Function

MT locates AP by listening Broadcast Channel (BCH). Selects AP among the APs heard which has the best signal to interference (C/I). MT initiates association mechanism and exchanges information about supported PHY rates, supported convergence layers, and supported authentication and encryption procedures. After association, the MT can request for one DLC user connections for each connection. Disassociation can occur either by notification of MT or by being unreachable for a certain amount of time.

DLC User Connection Control

The DCC controls the resources to grant DLC connection identifier. No traffic is allowed to send if there is no DLC user connection between the MT and AP.

Radio Resource Control

RRC is responsible for handover, dynamic frequency selection, and power management. These are management features that leverage the signals. Observed signals are used to optimize the network.

1-Dynamic Frequency Selection DFS function is a distributed mechanism to select the operating frequency. The AP selects the frequency in a manner to avoid the used ones.

2-Link Adaptation HIPERLAN/2 defines signal to interference ratio (C/I). Link adaptation scheme copes with varying radio quality by C/I and adapts its data rate to make the link robust.

3-Multi Beam Antennas HIPERLAN/2 supports multi-beam antennas to service a station in high C/I.

4-Handover When MT finds itself away from its AP and near to another AP, MT initiate handover mechanism.

5-Power Control MT and AP control their power to decrease the interference. This also enables power save mode by putting MT into sleep mode. MT negotiates with AP about the sleeping times in number of frames unit and goes to sleep. After the sleep time, either MT wakes up or AP have MT continue to sleep.

6-MT alive AP sends "alive" message to active MTs which do not transmit any traffic for a time. If there is no response from MT, then the MT is disassociated.

9.12.3 Convergence Layer

The convergence layer (CL) is responsible for mapping the service request of higher layers to the services offered by DLC and converting the higher layer packets into fixed size DLC format. Padding, segmenting, and reassembling make the DLC compatible with any network, e.g. 3G, Ethernet, ATM.

9.12.4 HIPERLAN/2 vs 802.11a

HIPERLAN/2 (H2) has a physical layer which is similar to IEEE 802.11a. There are various implementation differences.

- Modulation schemes

  HIPERLAN/2 and IEEE 802.11a supports different data rates. In 16-QAM, HIPERLAN/2 uses coding rate of 9/16 to achieve 27 Mbps and IEEE 802.11a uses coding rate of 1/2 to achieve 24 Mbps.

Modulation scheme	Coding rate	Physical data rate	Data bits per symbol	H2	802.11a
BPSK	1/2	6Mbps	24	√	√
BPSK	3/4	9Mbps	36	√	√
QPSK	1/2	12Mbps	48	√	√
QPSK	3/4	18Mbps	72	√	√
16-QAM	3/4	36Mbps	144	√	√
16-QAM	9/16	27Mbps	108	√	
16-QAM	1/2	24Mbps	96		√
64-QAM	2/3	48Mbps	192		√
64-QAM	3/4	54Mbps	216	√	√

- Preambles

  IEEE 802.11a uses only one preamble but HIPERLAN/2 uses several preambles for different purposes: Broadcast burst preamble, Downlink burst preamble, short preamble for uplink burst, long preamble for uplink burst. These preambles are created by chaining the number of short and long symbols.

- Data scrambler

  They both use the scrambler with a length 127 pseudo random sequence but the initialization process is different.

- Puncturing

  HIPERLAN/2 uses two puncturing units: P1 and P2. P1 is applied to the first 156 bits of the packet, and the rest of the packet is treated with P2 to obtain the desired coding rate. IEEE 802.11a only uses P2. P2 is designated to provide 1/2, 3/4, 9/16 coding rates in HIPERLAN/2 and 1/2, 2/3, 3/4 coding rates in 802.11a.

- Frame format

  IEEE 802.11a has PLCP header together with PLCP preamble and payload in the frame format to contain length and rate infor-

mation. On the other side, HIPERLAN/2 does not have a header since this information is determined on MAC level.

9.13 MMAC-HiSWAN

Multimedia Mobile Access Communication (MMAC) System is a system that is designated to transmit high speed, high quality multimedia information in Japan. The system introduces two WLAN standards: HiSWANa (High Speed Wireless Access Network), operating in 5GHz, and HiSWANb, operating in 25GHz. HiSWANa adopts OFDM physical layer and it is similar to HIPERLAN/2. It provides a data rate between 6-36 Mbps. In MMAC HiSWANa, access points are equipped with a carrier sense functions to do the frequency planning and access points have a interim synchronization mechanism to avoid interference [124].

9.14 Overview of IEEE 802.11 Standards

In addition to the 802.11a, 802.11b and 802.11e, 802.11 standards are being developed that extend the physical layer options, improve security, add quality of service (QoS) features or provide better interoperability [127].

9.14.1 IEEE 802.11c - Bridge Operation Procedures

802.11c identifies the procedure in the case of bridge operations. This standardizes the access point development.

9.14.2 IEEE 802.11d - Global Harmonization

802.11d is supplementary to the Media Access Control layer to promote worldwide use of 802.11 WLANs. It will allow access points to communicate information on the permissible radio channels with acceptable power levels for user devices. The 802.11 standards cannot legally operate in some countries; the purpose of 11d is to add features and restrictions to allow WLANs to operate within the rules of these countries.

9.14.3 IEEE 802.11f - Inter Access Point Protocol

802.11f is a "recommended practice" document that aims to achieve radio access point inter-operability within a multi-vendor WLAN network. The standard defines the registration of access points within a network and the interchange of information between access points when a user is handed over from one access point to another.

9.14.4 IEEE 802.11g - Higher Rate Extensions in the 2.4GHz Band

802.11g is an extension to 802.11b and introduces higher data rates than 802.11b using OFDM technology within the 2.4 GHz band. The maximum link rate is 54-Mbps per channel compared with 11 Mbps for 802.11b. The 802.11g standard uses OFDM modulation but, for backward compatibility with 802.11b, it also supports complementary code keying (CCK) modulation and, as an option for faster link rates, allows packet binary convolutional coding (PBCC) modulation.

The 802.11g operates in 2.4 GHz and uses approximately 30 MHz which provides three non-overlapping channels and will likely to cause the same problems that 802.11b networks face. The RF interference problem is also present for 802.11g networks. The control signals

should be sent by CCK in order for 802.11b to receive them. This may cause problem since the CCK preamble is longer than that of OFDM thereby degrades the network performance.

9.14.5 IEEE 802.11h - Spectrum Managed 802.11a

802.11h is called spectrum and transmit power management extensions in the 5GHz band in Europe. This standard is supplementary to the MAC layer to comply with European regulations for 5 GHz WLANs. European radio regulations for the 5 GHz band require products to have dynamic frequency selection (DFS) and transmission power control (TPC).

Dynamic Frequency Selection (DFS)

DFS implement a mechanism to avoid co-channel operation with radar systems and to ensure uniform channel spreading. DFS shall be used to do automatic frequency planning which is to select the radio channel at the access point to minimize interference with other systems.

The DFS services provide for:

- Association of STA based on STA's supported channels.
- Reserving the current channel to be tested for the presence of radar.
- Testing channels for radar before using the channel.
- Terminating the channel use after detecting radar.
- Requesting and reporting of measurements in the channels.
- Advertising a new channel to mitigate the BSS after the radar is detected.

Transmission Power Control (TPC)

TPC limits the transmitted power to the minimum needed to reach the furthest user. TPC shall also be used to do interference. The TPC services provides

- Association of STA based on the STA's power capability.
- Specification of regulatory and local maximum transmit power levels for the current channel.
- Selection of a transmit power for each transmission in a channel within constraints imposed by regulatory requirements.
- Adaptation of transmit power based on range, path loss and link margin.

Completion of 11h will provide better acceptability within Europe for IEEE compliant 5 GHz WLAN products. Essentially 802.11h is 802.11a with TPC and DFS.

9.14.6 IEEE 802.11i - MAC Enhancements for Enhanced Security

IEEE 802.11i is called medium access control security enhancements. Supplementary to the MAC layer to improve security. It will apply to 802.11 physical standards a,b and g. It provides an alternative to WEP with new encryption methods and authentication procedures. IEEE 802.1x forms a key part of 802.11i. Weakness in WEP encryption is damaging to the 802.11 standard perception in the market. In addition, the WEP algorithm weakness has been exposed. The 11i specification is part of a set of security features that addresses and overcomes these issues. Solutions will start with firmware upgrades using the Temporal Key Integrity Protocol (TKIP), followed by new silicon with AES (an iterated block chipper) and TKIP backwards compatibility.

9.14.7 IEEE 802.1x - Port Based Network Access Control

802.1x is a framework for regulating access control of client stations to a network via the use of extensible authentication methods. It forms a key part of the important 802.11i proposals for enhanced security. It applies to 802.11 physical standards a,b and g.

9.14.8 IEEE 802.1p - QoS on the MAC Level

802.11p is a standard for traffic class and dynamic multi-cast filtering. It provides a method to differentiate traffic streams in priority classes in support of quality of service offering. It forms a key part of the 802.11e proposals for QoS at the MAC level. This applies to 802.11 physical standards a,b and g.

Chapter 10

Digital Broadcasting

10.1 Broadcasting of Digital Audio Signals

It has been recognized, particularly in Europe, that multi-carrier modulation is especially appropriate for broadcasting of digital signals. European standards bodies have adopted OFDM as the modulation choice for several variations of digital audio and digital television, for both terrestrial and satellite transmission [12, 135, 136]. We will concentrate here on one particularly interesting form of terrestrial digital audio broadcasting.

Substantial quality improvement can be obtained over conventional analog FM by using digital techniques, just as digital compact disk (CD) recording has widely supplanted analog disks and tapes. With modern compression techniques, high quality stereo audio can be produced at approximately 200 kbps, an order of magnitude reduction from CD technology, which uses a straight PCM encoding without compression.

The European Digital Audio Broadcast (DAB) standard provides for three modes of operation. Mode 1 is for terrestrial Single Fre-

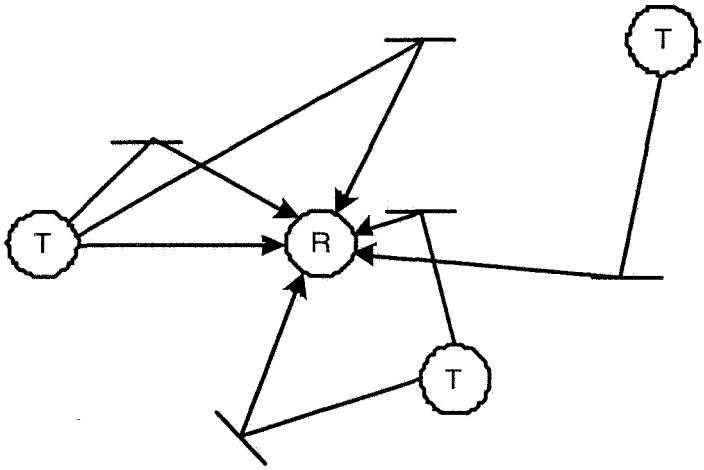


Figure 10.1: Received signal in an SFN

quency Networks (SFN). Mode 2 applies to conventional terrestrial local broadcasting. Mode 3 is for satellite broadcasting. The ability to provide an SFN is a unique capability of multi-carrier modulation, and the remainder of this section will discuss this application.

An SFN involves the provision of several geographically dispersed transmitters, all transmitting the same signal, synchronized in time. This permits coverage of a large area, and also fills in dead spots due to shadowing at a particular receiver location. Any receiver therefore receives a sum of signals from more than one transmitter, coincident in frequency, format, and time, except for differences in propagation delay. This is illustrated in Figure 10.1, showing multi-path propagation from several transmitters.

Assuming that each path provides a Rayleigh distributed signal, the powers will add. If the spread of arrival times, which includes both multi-path spread and spread of propagation times from different transmitters, is less than the guard interval between OFDM sym-

bols, then orthogonality among sub-carriers is preserved. If this is not the case, then degradation caused by interference among sub-carriers will result. This interference may be approximated as Gaussian noise, and is proportional to the degree that the inter-symbol interference duration exceeds the guard interval. As an example of a guard interval requirement, consider a receiver receiving direct signals from two transmitters 40km apart. The time separation of the signals could be up to 133μsec. Considering the addition of multi-path from each transmitter, and reception from additional transmitters, the guard interval should be substantially greater than this value. Note that as in a cellular radio system, signals from far distant transmitters lead to system degradation.

A constraint in the other direction, due to Doppler shift and frequency variation in the receiver's local oscillator, sets a minimum spacing between sub-carriers, and thus a maximum symbol duration. For a vehicle travelling at 80 mph (31m/s), the Doppler shift at 240 MHz can be up to 25 Hz. The sub-carrier spacing must be large compared to this value in order to minimize this source of inter-carrier interference. It is particularly important in an OFDM system that a good frequency lock be maintained in the receiver.

The Mode 1 standard meets these constraints by providing for 1536 sub-carriers with 1KHz spacing. A guard interval, which employs a cyclic prefix, has a duration of 246μsec.

Since the useful symbol duration is 1msec, the loss in spectral efficiency is approximately 20%. Each sub-carrier is modulated by $\pi/4 - DPSK$, that is a 4-point constellation in which two bits of information are carried by differential phase modulation of

$$\pm\frac{\pi}{4} or \pm\frac{3\pi}{4} \tag{10.1}$$

A 2048-point IFFT is performed in the transmitter, treating the unused sub-carriers as zero, so the bandwidth remains 1.536 MHz. The signal is complex modulated, using I and Q components, onto an

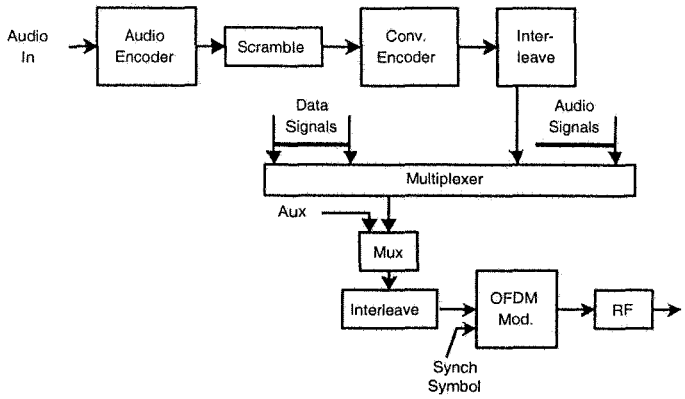


Figure 10.2: Block diagram of a DAB transmitter

RF carrier in the range 175 - 240 MHz.

The transmitted OFDM signal carries several audio programs, each of which may be monophonic or stereo, and of varying quality. Each audio channel is encoded using sub-band encoding, at one of several allowed bit rates between 32 and 384 kbps. That bit stream is then convolutionally encoded at rate of approximately $\frac{1}{2}$ and interleaved. The bit streams are then time multiplexed and input to the OFDM modulator. The gross bit rate, as described below, is 2.3 Mbps. Because of the variability of the original audio encoding, the number of channels carried is also variable. A typical system carries six programs each using 192 kbps.

10.2 Signal Format

The construction of an overall transmitted DAB signal is shown in Figure 10.2.

The audio encoder accepts digital samples of the audio input, which in general is 2-channel stereo. The samples occur at a 48KHz. rate, 16 bits per sample, for a total rate of 768 kbps. per channel. A sub-band encoding technique is used to reduce the rate to a lower value, depending on the desired subjective quality.

The output of the audio encoder is scrambled by modulo-2 addition with a pseudo-random sequence formed from a 9-bit feedback shift register. This is different from the more common scrambling technique of division by a randomizing polynomial. It does not suffer from error multiplication effects, but does require synchronization of the descrambling sequence at the receiver. The purpose of this scrambling is to insure that the final transmitted frequency spectrum is suitable dispersed. In addition, the standard allows for additional encryption scrambling if appropriate.

The scrambled data is then convolutionally encoded. This is essential for satisfactory operation over a radio channel. The first step is a mother code of rate $\frac{1}{4}$ with constraint length 7. That code is then subjected to puncturing to yield a rate of the form $8/n$, where n can be any integer value between 9 and 32. Each audio channel can have a different code rate, depending on the required error rate, to produce a given audio quality over the expected worst case channel. Typically a rate of approximately $\frac{1}{2}$ is used. Interleaving by a factor of 16 is then applied.

Several audio channels are multiplexed and combined with other signals, which may include data channels in addition to audio. The signals are arranged in frames, as shown below. A frame consists of a time sequence of 2 synchronization symbols, 3 digital overhead symbols, followed by 72 symbols of multiplexed audio information. A symbol is a block of 3072 bits which will be mapped into an OFDM symbol. The length of a frame is therefore $77 * 1.246 \approx 96$ $msec$.

The overhead and information symbols are interleaved before assignment to sub-carriers. This provides further randomization of the signal. The synch symbols are added and the OFDM modulation per-

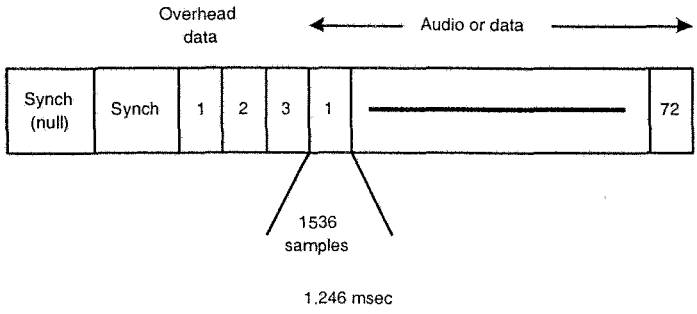


Figure 10.3: DAB frame structure

formed.

The first synch symbol is null (no signal), so that course frame alignment can be obtained by simple envelope detection. The second synch symbol carries a fixed pattern. It is used both for fine acquisition and to provide phase reference for differential detection of the following symbols. The overhead bits constitute a fast information channel, which carries parameters necessary for decoding subsequent information.

The raw bit rate can be calculated as

$$\frac{72\ symbols*1536\ subcarriers/symbol*2\ bits/subcarrier}{0.096\ sec} = 2.3Mbps \tag{10.2}$$

As describer previously, each sub-carrier employs 4-point DPSK modulation. The differential modulation applies to the sequence of symbols of each sub-carrier, rather than among sub-carrier symbols as is sometimes proposed. Differential modulation eases the requirements of synchronization and simplifies receiver implementation, at the expense of approximately $2dB$ in noise performance. For satisfactory performance, the OFDM symbol rate must be large compared

to the maximum Doppler shift. An equivalent statement is that the OFDM symbol duration must be small compared to the time coherence of the channel.

10.3 Other Digital Broadcasting Systems

10.3.1 DAB in the U.S.A

The system requirements for DAB in the U.S.A. are considerably different from those in Europe. Most importantly, it is desired to operate in the same 88 - 108 MHz band that is used for conventional analog FM broadcasting, using the same 200 KHz channels. This leads to a potentially severe problem of co-channel and adjacent channel interference between the digital and analog signals.

The problem of interference from the digital signals into analog receivers in reduced by transmitting the digital signals at a much lower power level than the analog signals. The FM capture effect then assists in rejection of the interference. Careful geographic frequency assignment and co-ordination is essential.

As for reception of the digital signal, interference from analog stations is a form of additive noise. Robust signal format and coding are essential. One suggested enhancement involves detecting an analog interferer and attempting to subtract it out [137, 138].

Standardization of the signal format is in progress, but some form of OFDM will most likely be adopted.

A standard similar to the DAB standard has been adopted in Europe for broadcasting of digital television signals. Using MPEG-2 coding, approximately 3 - 4 Mbps are required for high quality SECAM or PAL programming, while up to 20 Mbps are needed for future high definition television.

A bandwidth of 8 MHz is used, as in PAL or SECAM, to carry either several standard channels or one high definition channel. The same SFN scheme as in DAB will be deployed. In this case 8000 sub-carriers spaced approximately 1KHz apart are used. Larger constellation sizes, up to 64 points, provide the higher bit rates.

The large constellation sizes require more robust techniques than are used in DAB [139, 140]. Modulation is coherent as opposed to differential. A concatenated inner convolution code and outer block code are employed. Several pilot tones are transmitted for more precise equalization and synchronization.

10.4 Digital Video Broadcasting

Digital video broadcasting services cover more than digital television. The wireless video concept includes applications such as multimedia mobile and wireless data services. In addition, DVB systems can be integrated with other wireless services such as cellular systems to provide a highly asymmetric data access system which is desirable for new wireless Internet access. Source coding is one of the most critical components of DVB systems. Different data rates are required for the variety of services offered by DVB. Quality television requires about 216 Mbps and stereo audio requires about 1536 kbps with no source coding applied. MPEG 2 utilizes a compression algorithm to remove spatial, temporal and statistical redundancy of images, and sub-band coding techniques in conjunction with psycho-acoustic models to reduce the bit rate. The total bit rate, video and audio is reduced to approximately 3.5 Mbps. The choice of Coded Orthogonal Frequency Division Multiplexing (COFDM) for DVB-T applications is based on several design factors:

- *Interference rejection:* Co-existence of digital broadcasting systems with current analog services requires a system design which

restricts interference into analog systems and is robust against narrow band interference. The former can be met by reducing power level and utilizing coding techniques of digital communications to achieve required bit error rates. The latter can be met by multi-carrier modulation efficiently due to its robustness against narrow band interference such as VSB-AM.

- *Multi-path effect:* Multi-path propagation is a major impairment of a terrestrial broadcasting channel. The SFN architecture of DVB produces strong echo (0 dB) generated by adjacent transmitters. Presence of strong echo-like interference requires a robust modulation scheme. COFDM is again suitable to combat against strong multi-path as long as the guard interval length is properly designed. Larger guard intervals, as explained before, allows larger distance between transmitters at the expense of reduced bandwidth utilization. Proper positioning of pilot signals in time and frequency slots provides an acceptable performance in higher Doppler conditions such as mobile multimedia terminals.

After down conversion at the receiver's analog front-end and sampling, a digital AGC is used to adjust the VGA gain in the front-end circuit. A coarse frequency offset correction is necessary to reduce the inter-carrier interference. Frequency offset estimation uses the fixed pilots for initial estimation and scattered pilots for further fine synchronization. Time synchronization is initially performed by using a correlated structure and further refinement is done by methods explained in Chapter 5 for fine synchronization using the output of the OFDM demodulator. Channel estimation is required for detection and soft decoding reliability measure. Pilot symbols are used for channel estimation, and by using time and frequency interpolation we can obtain a channel estimate for all sub-carriers adaptively. Performance of the inner decoder is significantly enhanced by using a soft decoder and a reliability measure derived from channel estimation. A typical receiver is shown in Figure 10.5.

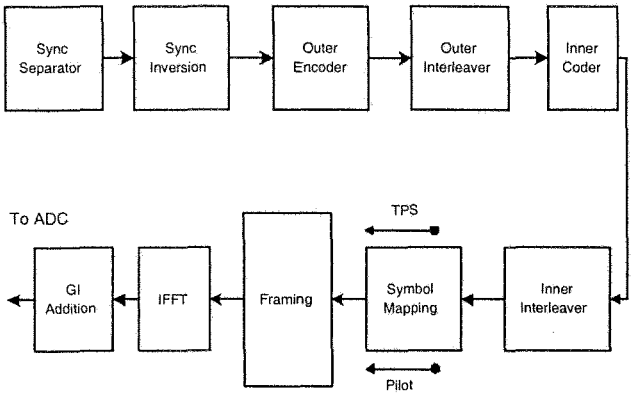


Figure 10.4: DVB transmitter structure

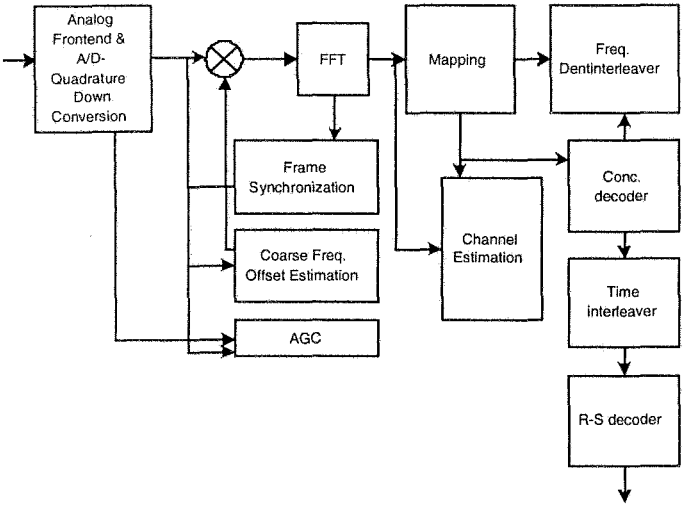


Figure 10.5: DVB receiver architecture

Chapter 11

OFDM based Multiple Access Techniques

11.1 Introduction

Spectrum is a shared resource among the users and needs an allocation technique to provide unique access per user. The technique called multiple access aims to minimize the impact of interference and use spectrum efficiently. The resource allocation schemes are basically distributing space of a 3D volume with code, time and frequency axes among users. There are several spectrum allocation schemes proposed.

Random access protocols are suitable when the traffic is bursty. A station initiates transmission whenever it has data to transmit. As a result, the access scheme does not pre-allocate bandwidth to a station and consequently, bandwidth is not wasted when the station is idle. Distributed behavior eventually results in collision, when two or more stations transmit at the same time. A contention based medium access protocol is provided in order to prevent and recover from corrupted

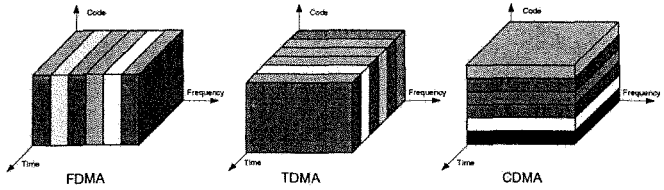


Figure 11.1: FDMA, TDMA and CDMA schemes

data packets. Aloha, Ethernet, IEEE 802.11 WLAN, HIPERLAN/1 are all based on random access schemes.

Unlike random access protocols where the access is not guaranteed, token ring based protocols guarantee access to each station by a token frame passing from one station to another. A station is allowed to transmit whenever it receives the token frame. A station transmits for a certain time and passes the token to its successor. In this access scheme, each station is granted access for sure and it is the stations responsibility to decide whether to use that time for transmission. IEEE 802.4 and Wireless Token Ring Protocol (WTRP) are based on token ring based multiple access schemes [123].

We focus mainly on OFDM related multiple access techniques. Although OFDM modulation shall be used with random access or token ring based, multiple access protocols, there are OFDM based multiple access protocols that are defined mainly in physical layer and derivation of frequency division multiple access (FDMA), time division multiple access (TDMA) and code division multiple access (CDMA) schemes.

11.2 OFDM-FDMA

Frequency division multiple access (FDMA) divides the available bandwidth (say 25 MHz) into smaller bandwidths (30 KHz) and assigns each sub-channel to a user as seen in Figure 11.1. If all sub-channels are occupied, the next call request can not be placed unless a user hangs up or is handed off to another cell, thus freeing up the channel. The channels are non-overlapping and interpolated by a guard band. FDMA has a simple algorithm and consequently requires simple hardware. Thus, it is efficient when the network size is small but has disadvantages when the size of the network exceeds the number of channels available. This forces to use the same channel for different users and increases the interference. FDMA is not conducive to varying station population. Unused channels and guard bands waste bandwidth. There is ISI unless the symbol period is bigger than the delay spread. FDMA does not have broadcast capability. Advanced Mobile Phone Service (AMPS) standing in 1G, Cordless telephone (CS2), Digital European cordless telephone (DECT) all deploy FDMA.

OFDM-FDMA systems split the entire band into sub-carriers. The user is allowed to do OFDM modulation only with a predetermined sub-carrier band. In each allocated subcarrier adaptive bit-loading is possible. Adaptive bit-loading assigns a different modulation schemes in order to assign a different number of bits to each sub-carrier. It is preferred only when the channel shows different gains in the different sub-carriers, as in ADSL.

Allocating sub-carriers makes different variations possible in the OFDM-FDMA system. Block FDMA and Interleaved FDMA are possibilities. Block FDMA assigns a group of adjacent sub-carriers to each user. Normally, adjacent sub-carrier channel gains show high correlation and bit-loading is unnecessary.

Interleaved FDMA, on the other hand, assigns sub-carriers distributed across the frequency band. Adaptive bit loading can be performed since at this time different subcarrier channels show different

gains because of less correlation.

11.3 OFDM-TDMA

Time division multiple access (TDMA) segments the time frame into slots. Integer number of slots are assigned to a user as seen in Figure 11.1. This makes a user utilize all spectrum for a bounded time interval. TDMA transmission is discontinuous and consequently bursty which also makes the handoff easier and transmission in low power. The frequency allocation is easier and allows lower transmit power than FDMA but TDMA has synchronization overhead due to the varying station population. Unlike FDMA, entire channel bandwidth is used. The throughput is good for uniform traffic.

Transmission is bursty and the burst contains a preamble, information bits and trail bits. Information bits are transmitted in slots. Each slot has trail bits, synchronization bits, information data and guard bits. TDMA suffers from ISI due to the high transmission rates.

After FCC's approval in 1987 for deployment of new technologies that do not interfere with the existing AMPS networks in order to satisfy the growth of cellular users, TDMA was introduced in 1991 based upon the Interim Standard 54 (IS-54), called D-AMPS in 2G cellular system. It was revised and the final version was standardized in 1994 as IS-136. Global System for Mobile (GSM) and Japanese Digital Cellular (JDC) also deploys TDMA.

In OFDM-TDMA system, a particular user uses all sub-carriers within the predetermined TDMA time slot. The bit loading can be performed depending on the channel condition. WLAN based systems work with OFDM-TDMA architecture. Each user uses OFDM modulation and gets transmission right through the MAC layer channel access mechanism.

11.4 Multi Carrier CDMA (OFDM-CDMA)

Code division multiple access (CDMA) is a spread spectrum technique using Direct sequence spread spectrum (DSSS). All users in CDMA use the entire time and frequency as illustrated in Figure 11.1. CDMA was used first during World War II to transmit secure messages between Churchill and Roosevelt. CDMA uses a narrow-band signal which is multiplied by a spreading signal. The spreading signal is a pseudo-noise code sequence that has a rate much higher than the data rate of the message. The destination can recover the message using the correct code. Each code is unique and all subscribers use the same channel concurrently. There is no need for spatial reuse since same channel can be used in the adjacent cell.

CDMA has soft capacity limit; an increase in the number of users increases the noise floor, because all other users' signals on a channel may be treated as noise. Thus, the performance degradation limits the number of users indirectly. Multi-path fading is considerably reduced since the signal is spread over frequency and time. The near-far problem occurs if there is an undesired user that has high detected power.

The very first standard of CDMA is called IS-95 standing in 2G and completed in 1993. It is referred cdmaOne in CDMA jargon. It achieves 10 times the capacity of AMPS and support data rates up to 14.4 kbps. The next CDMA standard is IS-95B standing in 2.5G. IS-95 uses eight codes each supports 9.6 or 14.4 kbps. Maximum theoretical data rates of 78.6 kbps, or 115.2 kbps can be achieved when eight codes are consumed by one user. CDMA2000 is the next generation CDMA deployed standard for 3G cellular networks.

In frequency selective fading channels, orthogonality of codes are not guaranteed due to inter-chip interference. This causes multiple access interference and decreases the capacity. In order to suppress this effect, several proposals have been made to combine the use of CDMA with the desirable properties of OFDM. The longer symbol interval in OFDM greatly eases the problem of acquiring synchro-

nization of the spreading code in the receiver. The sub-carriers are spaced sufficiently close such that each sub-carrier constitutes a frequency non-selective fading channel. This eliminates the need for RAKE combining. Figure 11.2 shows one scheme. The same data symbol, multiplied by a different bit of the spreading code $\{c\}$, modulates each sub-carrier. This in effect provides spreading in the frequency domain. The transmitted signal element duration is the same as the original symbol duration, rather than the chip duration as in ordinary CDMA. Synchronization is therefore a much simpler process. Provided that adequate guard intervals, using cyclic prefixes, are used, multi-path correction is easier and better than RAKE combining in a single carrier system. This OFDM-CDMA technique provides frequency diversity. Best performance in this environment is achieved through the use of optimum combining in the choice of weighting factors $\{q\}$ that are applied in the receiver after demodulation and despreading. In the downstream direction (from a central station to the mobile) multiple access is achieved by superimposing signals for different users on the same sub-carriers using a set of orthogonal spreading codes such as Walsh functions. In the upstream direction a set of spreading codes, such as Gold codes, with good auto-correlation and cross-correlation properties, are employed. However, in this case a multi-user detection scheme in the receiver is essential because the asynchronous arrival times destroy orthogonality among the OFDM sub-carriers.

The above description assumes that the number of spreading chips per symbol, or the processing gain, is equal to the number of sub-carriers. This is not necessary. If more sub-carriers are desirable, either to achieve flat spectra or to achieve greater spectral efficiency with a given guard interval, then more than one input symbol may be spread to form each OFDM symbol. Another form of combined OFDM-CDMA involves feeding different spread input symbols to the sub-carriers, as illustrated in Figure 11.3. The spreading codes for a given user may all be the same as shown, or could be different. No integral relationship between the processing gain and the number of

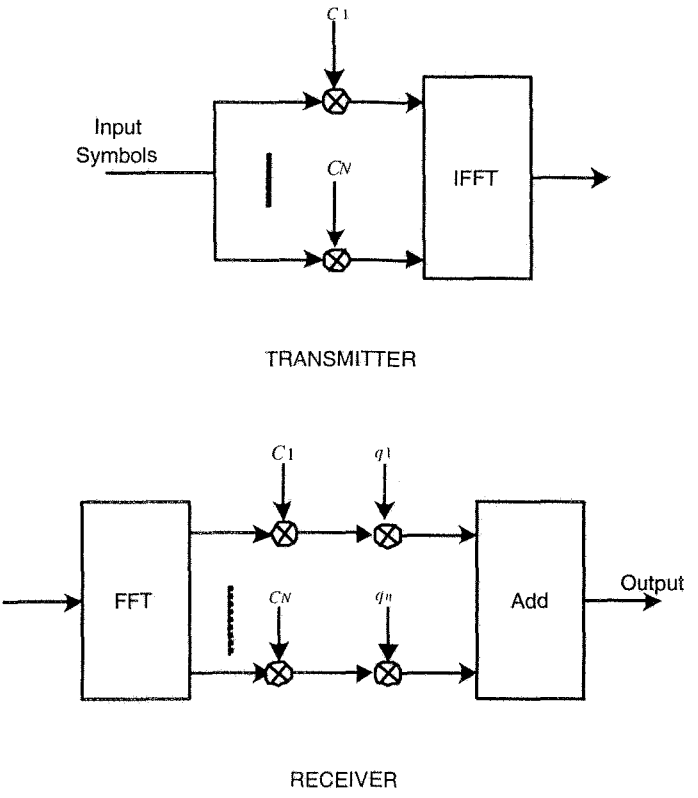


Figure 11.2: OFDM-CDMA system with frequency spreading

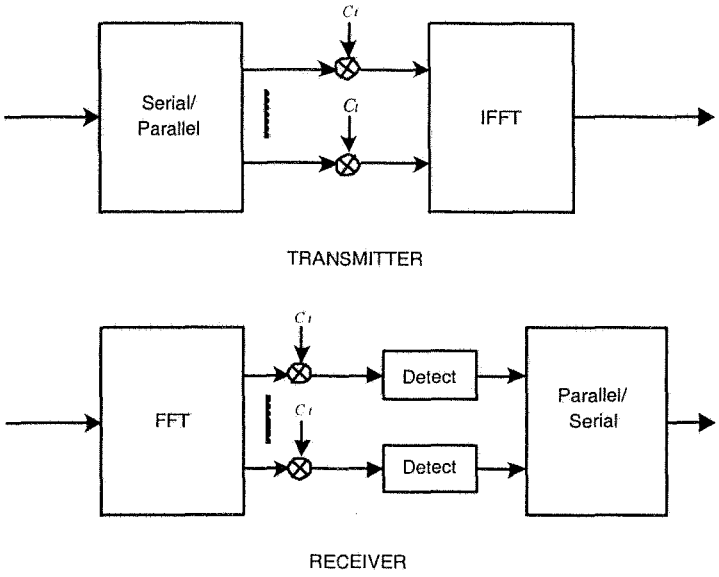


Figure 11.3: OFDM-CDMA system with time spreading

sub-carriers is required. This process uses spreading in the time domain as in standard CDMA, but at a rate much lower by a factor of $1/N$. For the same reasons as in the first scheme, code acquisition is eased and RAKE combining is not needed.

Other variations have been proposed, including very high rate modulation so that sub-channel spectra severely overlap. RAKE type techniques are required to receive this signal. A different interpretation of OFDM-CDMA using a duality principle in digital communications is presented in [169]. Using the duality principle, we can design the receiver as dual of a single carrier CDMA RAKE receiver. This duality is shown in Figures 11.4 and 11.5.

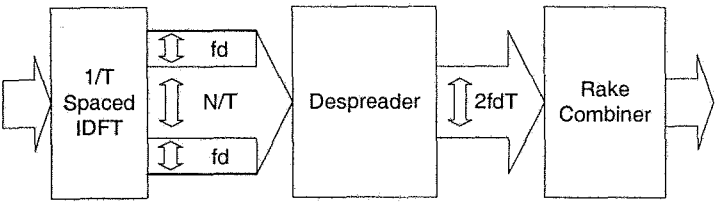


Figure 11.4: Discrete time MC-SS RAKE in frequency domain

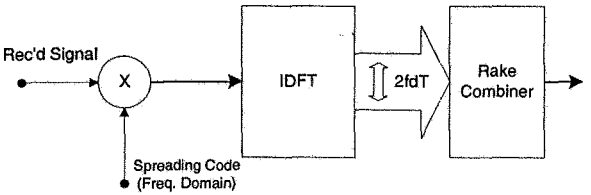


Figure 11.5: Simplified RAKE receiver

11.5 OFDMA

Broadband Wireless Access (BWA) is an appealing system for providing a flexible and easy deployment solution to high-speed communications. It appears as an alternative to wireline broadband access techniques such as copper line, coaxial cable, xDSL and cable modems.

The IEEE 802.16 group issued standards in the 10-66 GHz bands and the 802.16a group was formed to develop standards to operate in the 2-11 GHz bands wherein the channel impairments, multi-path fading and path loss become more significant with the increase in the number of subscribers (See Chapter 13). Improved and flexible multiple access methods are needed to cope with these impairments. Orthogonal Frequency Division Multiple Access (OFDMA) is a promising multiple access scheme that has attracted interest. OFDMA is based on OFDM and inherits its immunity to inter-symbol interference and frequency selective fading.

In multiuser environment, a good resource allocation scheme leverages multiuser diversity and channel fading. It was shown that the optimal solution is to schedule the user with the best channel at each time. Although in this case, the entire bandwidth is used by the scheduled user, this idea can also be applied to OFDMA system, where the channel is shared by the users, each owning a mutually disjoint set of sub-carriers, by scheduling the sub-carrier to a user with the best channel among others. Of course, the procedure is not simple since the best subcarrier of the user may be also the best sub-carrier of another user who may not have any other good sub-carriers. The overall strategy is to use the peaks of the channel resulting from channel fading. Unlike in the traditional view where the channel fading is considered an impairment, here it acts as a channel randomizer and increases multiuser diversity. Scheduling the sub-carriers in a optimum manner is called resource allocation in the literature.

Resource allocation problem can be defined differently depending on the constraints of the network. A real-time resource allocation

problem is defined when Quality of Service (QoS) requirements are fixed by the application. QoS requirement is defined as achieving the data transmission rate and bit error rate (BER) of each user in each transmission. Capacity maximization resource allocation problem, on the other hand, is based on water-pouring schemes wherein the aim is to approach Shannon capacity under the power constraint.

11.5.1 OFDMA Architecture

A typical OFDMA architecture is introduced in [160].

Unlike OFDM systems, K users involved in the OFDMA system share the N sub-carriers. The difference arises in the forming and deforming of an FFT block. The rest is the same as in an OFDM system as seen in Figure 11.6. Each user allocates a non-overlapping set of sub-carriers S_k where the number of sub-carriers per user is $J(k)$. We denote by $X_k(l)$ the lth subcarrier of the FFT block belonging to k^{th} user. $X_k(l)$ is obtained by coding the assigned bits c with the corresponding modulation scheme. In the down-link the $X_k(l)$ are multiplexed to form the OFDM symbol[1] of length $(N + L)$ with the appended guard prefix L in order to eliminate ISI. At the uplink, the OFDM symbol is formed in the base station with a synchronization error.

$$x(l) = \sum_{k=0}^{K-1} \sum_{n=0}^{J(k)-1} X_k(n) e^{j\frac{2\pi}{N}(I_k(n))l}, \tag{11.1}$$

with $n = -L, ..., N-1$, where $I_k(n)$ denotes the subcarrier assigned to the kth user. Resource allocation problem comes into the picture when associating the set of sub-carriers to the users with different bits loaded into them. The received signal from the jth user is

$$y_j(l) = x(l) \bigotimes h_j(l) + w(l), \tag{11.2}$$

[1] A OFDMA symbol is defined as one OFDM FFT block.

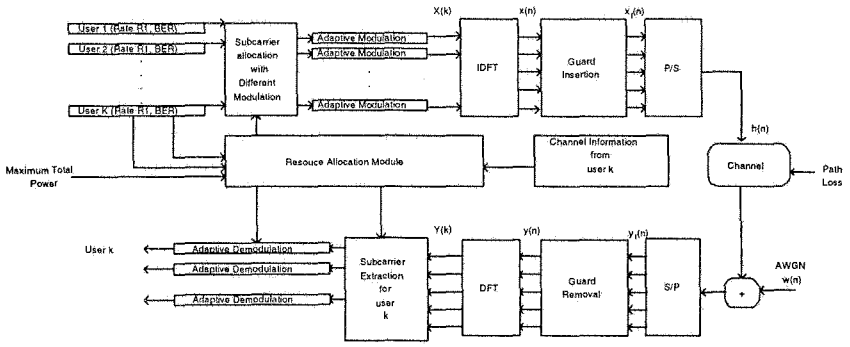


Figure 11.6: Orthogonal Frequency Division Multiple Access System

where $h_j(t)$ is the baseband impulse response of the channel between base station (BS) and j^{th} user. Equation (11.2) is the received signal $y(t)$ sampled at rate $1/T$. The first L samples are discarded and the N-point FFT is computed. The data of the jth user is

$$Y_j(n) = \begin{cases} X_j(n)H_j(i_j(n)) + W(n), & if\, i_j(n) \in S_j \\ 0, & otherwise, \end{cases} \tag{11.3}$$

where $H_j(n) := \sum_i h_j(i) exp(j2\frac{\pi}{N}ni)$ is the frequency response of the channel of kth user.

In a perfectly synchronized system, the allocation module of the transmitter assigns sub-carriers to each user according to some QoS criteria. QoS metrics in the system are rate and bit error rate (BER). Each user's bit stream is transmitted using the assigned sub-carriers and adaptively modulated for the number of bits assigned to the sub-carrier. The power level of the modulation is adjusted to overcome the fading of the channel. The transmission power for AWGN channel can be predicted. In addition the channel gain of sub-carrier n to the corresponding user k should be known. The channel gain of the subcarrier is defined as

$$\alpha_{k,n} = H_k(n) * PL_k, \tag{11.4}$$

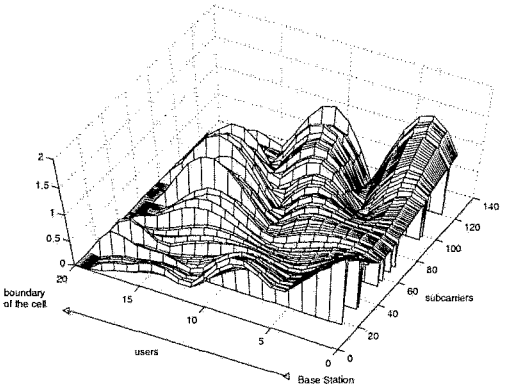


Figure 11.7: An example of channel gain

where PL is the path loss, defined by

$$PL_k = PL(d_o) + 10\alpha log_{10}(d_k/d_o) + X_\sigma,$$

where d_o is the reference distance, d_k is the distance between transmitter and receiver, α is the path loss component and X_σ is a Gaussian random variable for shadowing with a standard deviation σ. An example of channel gain can be seen in Figure 11.7.

The problem above is called resource allocation in the OFDMA literature; the channel information is assumed to be known at the transmitter and receiver. The channel is assumed to be reciprocal; the BS is able to estimate the channel of all BS-to-mobile links based on the received uplink transmission as long as the channel variation is slow. As a result, the resource allocation should be done within the coherence time. With the channel information, the objective of resource allocation problem can be defined as maximizing the throughput subject to a given total power constraint regarding the user's QoS requirements. As we clarify further, BER_k of the transmission should not be higher than the required BER_k and data rate of every user should be equal to the requirement R_k.

Let's define $\gamma_{k,n}$ as the indicator of allocating the nth sub-carrier to the kth user. The transmission power allocated to the nth subcarrier of kth user is expressed as

$$P_{k,n} = \frac{f_k(c_{k,n}, BER_k)}{\alpha_{k,n}^2}, \tag{11.5}$$

where $f_k(c_k, n)$ is the required received power with unity channel gain for reliable reception of $c_{k,n}$ bits per symbol.

Required transmission power for c bits/subcarrier at a given BER with unity channel gain for M-QAM modulation is

$$f(c, BER) = \frac{No}{3}[Q^{-1}(\frac{BER}{4})]^2(2^c - 1) \tag{11.6}$$

where $Q^{-1}(x)$ is the inverse function of

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt.$$

We can formulate the resource allocation problem with an imposed power constraint as

$$\begin{aligned} &\max_{c_{k,n}, \gamma_{k,n}} \quad R_k = \sum_{n=1}^{N} c_{k,n}\gamma_{k,n} \; for \; all \; k \\ &subject \; to \quad P_T = \sum_{k=1}^{K} \sum_{n=1}^{N} \frac{f_k(c_{k,n}, BER_k)}{\alpha_{k,n}^2}\gamma_{k,n} \leq P_{Max} \end{aligned} \tag{11.7}$$

where the limit on the total transmission power is expressed as P_{Max} for all $n \in \{1, ..., N\}$, $k \in \{1, ..., K\}$ and $c_{k,n} \in \{1, ..., M\}$.

If there is no power constraint, Equation (11.7) is changed in order to minimize P_T subject to allocating R_k bits for all k (i.e problem is to find the values of the $\gamma_{k,n}$ and the corresponding $c_{k,n}$ while minimizing P_T). As it can be seen the cost function in our system is the power consumption matrix in Equation (11.5). Rather than using $\alpha_{k,n}^2$, we adopted using $P_{k,n}$ since in this case modulation type and BER get involved in the decision process.

11.5.2 Resource Allocation Regarding QoS

In a multiuser environment with multiple modulation techniques, the solution to the problem is complicated since the optimal solution needs to pick the sub-carriers in balance. We can classify the problem according to each set of bits assigned to a sub-carrier.

For a user k, $f_k(c_{k,n}) \in \{f_k(1, BER_k), ..., f_k(M, BER_k)\}$. We can construct M times $[K \times N]$ power matrices$\{P^c\}$ for each c. For a constant c, $\{f(c)\}$ can be computed and the transmission power requirement can be found with Equation (11.5). The dimension of the indicator function is incremented and represented by $\gamma_{k,n,c}$ and defined as follows:

$$\gamma_{k,n,c} = \begin{cases} 1, & if\ c_{k,n} = c \\ 0, & otherwise \end{cases} \tag{11.8}$$

The above problem can be solved with Integer Programming (IP). IP approach is considered to be the optimal solution for the resource allocation problem.

There are $K \times N \times M$ indicator variables and M power matrices where the entries of each matrix for a given c can be found from

$$P^c_{k,n} = \frac{f_k(c, BER_k)}{\alpha^2_{k,n}}. \tag{11.9}$$

Using Equation (11.9) as an input, the cost function now can be written as

$$P_T = \sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{c=1}^{M} P^c_{k,n} \gamma_{k,n,c} \tag{11.10}$$

and the description of the IP problem is

$$\min_{\gamma_{k,n,c}} P_T,\ for\ \gamma_{k,n,c} \in \{0, 1\} \tag{11.11}$$

subject to

$$R_k = \sum_{n=1}^{N} \sum_{c=1}^{M} c.\gamma_{k,n,c},\ for\ all\ k,$$

and

$$0 \leq \sum_{k=1}^{K} \sum_{c=1}^{M} \gamma_{k,n,c} \leq 1, \ for\, all\, n.$$

Although the optimal solution gives the exact results, from an implementation point of view, it is not preferred since in a time varying channel, in order to allocate the sub-carriers within the coherence time, the allocation algorithm should be fast and the complexity of IP increases exponentially with the number of constraints of the network. A real-time requirement leads to searching suboptimal solutions that are fast and close to the optimal solution. Several suboptimal allocation schemes are proposed for different settings in the literature. The main idea is to approximate the optimal scheme iteratively [150, 146, 160].

11.5.3 Resource Allocation Regarding Capacity

In the previous sections, we considered the problem occurs when the QoS requirements per symbol is fixed. Another way to approach the resource allocation is in terms of capacity. Suppose there is no fixed requirements per symbol and the aim is to maximize capacity.

It has been shown in [171] that for point-to-point links, a fair allocation strategy maximizes total capacity and the throughput of each user in the long run, when the users' channel statistics are the same. This idea underlying the proposed fair scheduling algorithm is exploiting the multiuser diversity gain.

With a slight modification, we can extend the fair scheduling algorithm for point-to-point communication to an algorithm for point-to-multipoint communication. Suppose the user time varying data rate requirement $R_k(t)$ is sent by the user to the base station as feed back of the channel condition. We treat symbol time as the time slot, so t is discrete, representing the number of symbols. We can keep track of average throughput $t_{k,n}$ of each user for a subcarrier in a past window

of length t_c. The scheduling algorithm will schedule a subcarrier $\bar{n}$ to a user $\bar{k}$ according to the following criterion

$$\{\bar{k}, \bar{n}\} = \arg\max_{k,n} \frac{r_{k,n}}{t_{k,n}} \tag{11.12}$$

where $t_{k,n}$ can be updated using an exponentially weighted low-pass filter described in [171]. Here, we are confronted with determining the $r_{k,n}$ values. We can set $r_{k,n}$ to R_k/N, where N is the number of carriers. With this setting, the peaks of the channel for a given subcarrier can be tracked. The algorithm assigns a user to a subcarrier when the channel quality in that subcarrier is high relative to its average condition in that subcarrier over the time scale t_c. When we consider all subcarriers the fairness criterion match with the point-to-point case as follows

$$\bar{k} = \max_{k} \frac{R_k}{T_k}, \tag{11.13}$$

where $T_k = \sum_{n=1}^{N} t_{k,n}$. The theoretical analysis of fairness property of Equation (11.13) for point-to-point communication is derived in [171]. We can apply those derivations for point-to-multipoint communication.

11.6 Flash-OFDM

Data communication over legacy cellular voice systems is inefficient since data is fundamentally different from voice. Flash-OFDM [174] is an OFDMA system where the subcarriers are distributed among the mobile stations. Spectrum is divided into N tones and each base station transmits OFDM signal on all tones but different mobiles in the same cell use different tones. Flash-OFDM uses fast hopping across all tones in a pseudorandom predetermined pattern, making it a spread spectrum technology. With fast hopping, a user that is assigned one tone does not transmit on the same tone every symbol, but uses a hopping pattern to jump to a different tone every symbol duration. As

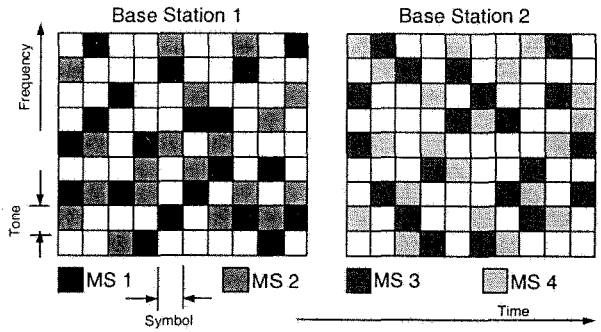


Figure 11.8: Flash-OFDM tone allocation

Carrier frequency	0.5 to 3.5 GHz
Bandwidth	1.25 MHz uplink/downlink
Number of tones	113
Cyclic prefix overhead	11%
Max delay spread	10us
Max Doppler spread	200Hz
Peak downlink rate	2.7 Mbps
Peak uplink rate	750 Kbps

Table 11.1: Flash-OFDM parameters

seen in Figure 11.8, tone allocation is non-overlapping and they are hopped to average the inter-cell interference. Therefore, users are orthogonal as in TDMA/FDMA system and users see average out-of cell interference as in CDMA. Typical parameters are seen in Table 11.1.

Flash-OFDM implements a rate adaptation scheme. It provides higher data rates when it is close to the base station. Thereby, in a given OFDM symbol tones could be differently modulated.

Session control only allows 31 mobiles in ON state at a time out of 125 active mobiles. This is a hard limit since non-overlapping

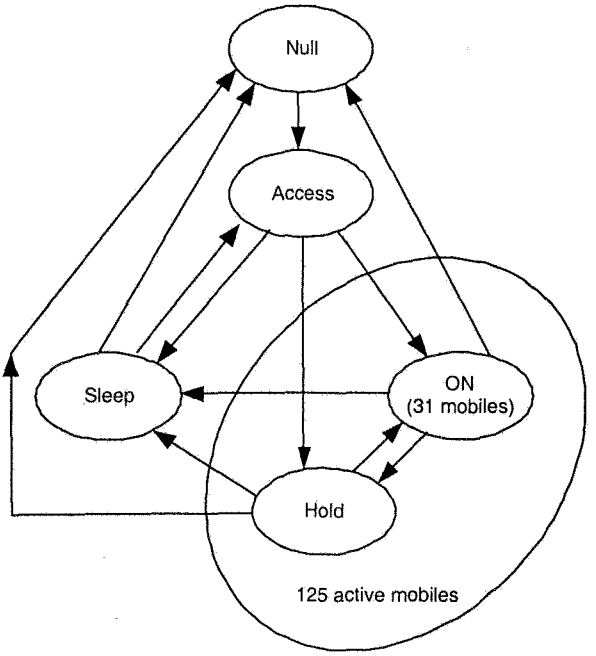


Figure 11.9: Flash-OFDM session control

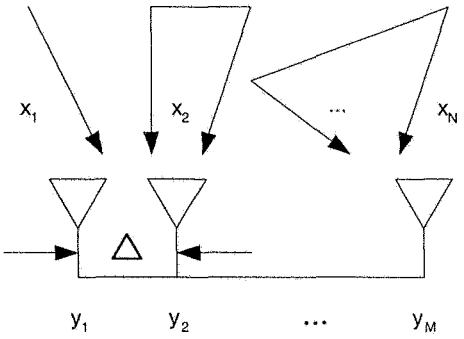


Figure 11.10: A SDMA system

tones are allocated. Flash-OFDM takes advantage of the granularity of OFDM in its control layer design enabling the MAC layer to perform efficient packet switching over the air. Figure 11.9 illustrates the session control mechanism.

11.7 OFDM-SDMA

Space division multiple access (SDMA) is a technique that employs a receiver antenna array. SDMA improves the spectrum efficiency significantly by allowing several users reuse the same channel. SDMA leverages spatial diversity which is one of the key resources in wireless communication. The spatial resource is highly difficult to deal with in MAC layer but in physical layer with a coordinated MAC by using a technique called spatial beam-forming. The SDMA scheme is suitable for fixed networks where the spatial diversity is stable. P receiver antenna can support P simultaneous users [163, 161].

A combined orthogonal frequency division multiplexing/spatial diversity multiple access (OFDM-SDMA) approach is proposed to

tackle two major obstacles introduced by the indoor wireless networks: The multi-path distortion can easily be mitigated with OFDM while the bandwidth efficiently can be increased by SDMA.

Let's assume that there are N mobile stations each with single antenna and one access point with M receiving antennas. The channel is indicated as

$$h_{i,j}(t) = \sum_{l=1}^{L} a_{i,j}(l)\delta(t - \tau_{i,j}^{(l)}) \tag{11.14}$$

where $a_{i,j}(l)$ and $\tau_{i,j}^{(l)}$ are the complex gain and delay of the ith path of a L path channel. For an OFDM system entire channel is divided into K sub-carriers. The gain of the kth sub-carrier between the ith user and j^{th} receiving antenna is given by:

$$H_{i,j}(k) = \sum_{l=1}^{L} a_{i,j}(l)e^{-j2\pi(k-l)f_s\tau_{i,j}^{(l)}} = (w_{i,j}^{(k)})^T h_{i,j} \tag{11.15}$$

where f_s is the bandwidth of the subcarrier. The received signal in M antennas for the kth subcarrier is given by

$$y[k] = H[k]x[k] + n[k] \tag{11.16}$$

where

$$\begin{aligned} y[k] &= [y_1[k]...y_M[k]]^T \\ x[k] &= [x_1[k]...x_N[k]]^T \\ n[k] &= [n_1[k]...n_M[k]]^T \\ H[k] &= [H_{i,j}(k)]_{i=1,...,N,j=1,...,M} \end{aligned} \tag{11.17}$$

and noise $n[k]$ variance is $\sigma^2 = N_o/2$. $H_{i,j}(k)$ is Gaussian and $|H_{i,j}|$ is Rayleigh distributed.

On the receiver side, a linear Minimum Mean Squared Error (MMSE) filter is used in order to suppress the multi-user interference. The MMSE estimation of the transmitted signal $x[k]$ is given by

$$x_{(MMSE)} = (R_{yy}^{-1}R_{yx})^H y[k] \tag{11.18}$$

where

$$\begin{aligned} R_{yy} &= E[yy^H] = H[k]H^H[k] + 2\sigma^2 I_{MxM} \\ R_{yx} &= E[yx^H[k]] = H[k] \end{aligned} \tag{11.19}$$

The direct matrix inversion introduced in MMSE makes the receiver very complicated. More accurate result can be achieved by more complicated optimum maximum likelihood (ML) receiver is given by

$$x_{(ML)} = min_{x_{(MMSE)}[k]}(|y[k] - H[k]x_{(MMSE)}[k]|^2) \tag{11.20}$$

OFDM-SDMA is an alternative to support multiple users with high spectrum efficiency when compared to OFDMA. OFDM-SDMA uses lower level of constellation than OFDMA since in order to support same number of users number of bits loaded into a sub-carrier should be more in OFDMA than OFDM-SDMA since the number of sub-carriers used in OFDMA is less than OFDM-SDMA. Since there is only one transmission, multi user interference is not present in OFDMA. The demodulation in OFDMA involves taking only a FFT which is much simpler than OFDM-SDMA where the signals from the antennas are combined with Maximum Ratio Combining. It has been shown in [170] that OFDM-SDMA can outperform OFDMA when the optimal ML detector is applied but not with MMSE receiver.

ML and MMSE detectors are complex and normally not preferred. In order to keep the network's cost acceptable simplified SDMA processors are preferred which perform somewhat worse than OFDMA. OFDMA-SDMA arises an attractive choice due to its lower complexity receiver.

Chapter 12

Ultra WideBand Technologies

Ultra WideBand (UWB) is a radio technology where the bandwidth is at least %25 greater than its center frequency and also greater than 500 MHz. Although UWB is a new term, the technology itself is not new and had been known since the early 1800's and was used to transmit morse codes by Spark Gap radio. In 1960's, the research on impulse measurements led to usage of pulse based transmission for radar and communications. UWB entered a new era in 2002 right after FCC's approval to use the 3.1-10.6 GHz band for unlicensed UWB devices (See in Figure 12.1).

What makes a previously existing system (UWB) receive a signifi-

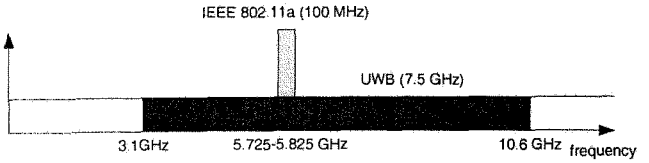


Figure 12.1: UWB spectrum

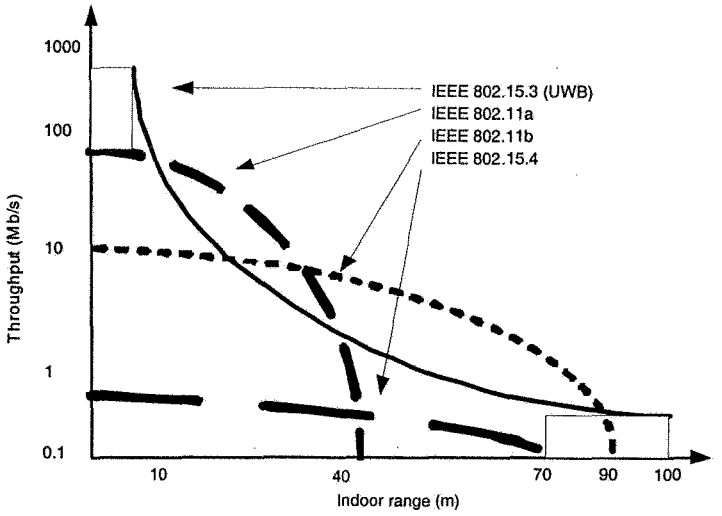


Figure 12.2: UWB performance

cant attention is that UWB promises to deliver data rates that can scale from 110 to 480 Mbps at a distance of 10 to 2 meters respectively as seen in Figure 12.2. Achieving high data rates is possible since the spectrum is very wide, consequently according to Shannon's formula, capacity of the network increases linearly with the used bandwidth as follow;

$$C = Wlog_2(1 + SNR)$$

where C and W are capacity and bandwidth respectively and SNR is signal to noise ratio.

12.1 Impulse Radio

The original UWB approach was a impulse radio where very short duration pulses ($\sim$ nanoseconds) are sent through the air. This cor-

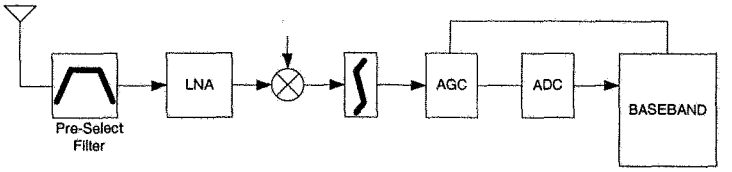


Figure 12.3: Impulse radio receiver

responds to large spectrum in frequency domain since bandwidth is inversely proportional to duration in time ($F \sim 1/t$).

Having a large spectrum results in a low energy density which reduces the interference. Large spectrum also makes it hard to determine the center frequency which causes the technique to be also named as "carrier-free", "baseband", or "impulse" technology in the past. Large spectrum spreads the signal energy over the bandwidth and this results in a low power density. Other than achieving high data rates, multipath degradation is also readily reduced since the direct and reflected paths do not overlap due to the short duration pulses. As a result reflected path does not cause any cancellation to the direct path.

The key advantage, however, is its low cost implementation due to the low system complexity. As it can be seen from the block diagram of impulse radio receiver in Figure 12.3 the received signal is directly converted into a baseband signal and there is no need for bulky filters and mixers, but only a high-speed RAKE receiver.

Impulse radio is used in radar applications since short duration pulses have the ability to penetrate through the obstacles that tend to reflect signals at limited bandwidths, and short duration pulses enable accurate positioning and fine radar range resolution.

12.2 Multiband Approach

With the FCC specification for a UWB system to use a bandwidth of 500 MHz at minimum, UWB is started to be considered as a multi-band signal, where each band is approximately 500 MHz. Now, it is possible to separate spectrum into 3-16 sub-bands unlike using the whole spectrum at once. Although single-band has been considered as the original approach like in Impulse radio, multi- band is becoming the strong candidate as the standard to be adopted. The multi-band approach has several advantages since now the information can be processed in smaller bandwidth [173];

- Lower Complexity.
- Less Power consumption.
- Lower cost.
- Flexible spectral efficiency.
- Lower rate ADCs.

12.3 Multiband OFDM

Single carrier (pulse based) or multi-carrier (OFDM) techniques can be used to send information in a sub-band. Multi-path fading is not present in impulse systems as much as the conventional radio systems since Rayleigh fading is effective in continuous signaling. But in multi-band systems there is a heavy multi-path environment. An OFDM receiver is advantageous over a RAKE receiver in combating multi-path since adding a long enough cyclic prefix to the end of the packet makes the FFT an ideal detector. A RAKE receiver on the other hand requires many fingers to capture the majority of multi-path energy and to assign fingers to correct delays is complicated [173].

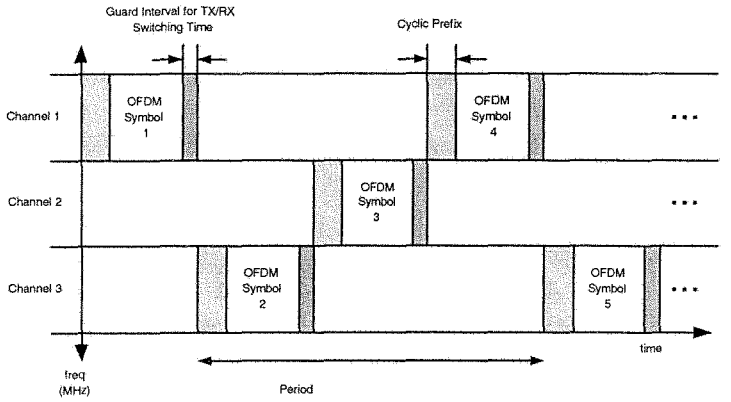


Figure 12.4: An example of time-frequency interleaving for the multi band OFDM in three band. © IEEE

Multi-band OFDM, a recent and strong proposal for UWB, implements time frequency interleaving (TFI) OFDM modulation. OFDM symbols are interleaved across both time and frequency. An illustration of three band TFI-OFDM is seen in Figure 12.4. As it can be seen from the figure a cyclic prefix (CP) and a guard band is appended to the symbol. Guard band (GB) is inserted in order to provide enough time for transmitter and receiver to switch from one channel to another. This is very crucial at reducing the complexity of transmitter since the FFT length is reduced: Let's say in the three band system the number of sub-carriers needed to cover a band is N. If the system is tuned to use full band then the FFT length will be $3 \times N$ but still data will be carried within the sub-carriers of one band which is N. By using GB, a single transmitter and receiver chain can treat the multiple transmissions as a full band transmission since they can switch from one band to another and transmit and collect like in full band by only using a FFT length equal to the number of sub-carriers in one band (N). Figure 12.5 shows a TFI-OFDM transmitter architecture. It is similar to the conventional OFDM transmitter except that

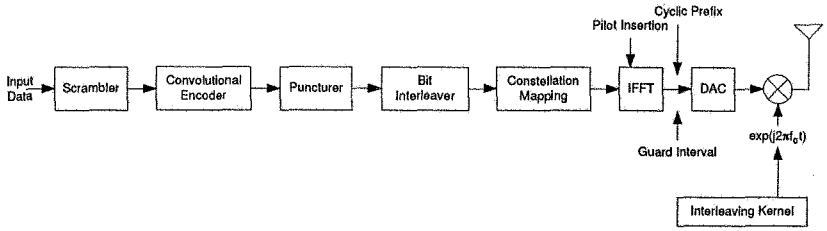


Figure 12.5: Example transmitter architecture for the proposed UWB PHY © IEEE

there is now a mechanism to change the carrier frequency based on time-frequency interleaving pattern.

Inheriting the features of OFDM, Multiband OFDM has several advantages [173];

- Robustness to multi path: Adding cyclic prefix makes OFDM modulation robust to multi path channel environments and it only needs a single-tap frequency equalizer.
- One transmitter and receiver chain: Adding the guard interval requires the use of only one transmitter and receiver chain.
- Compliance with worldwide regulations and coexistence with future systems: Since the TFI-OFDM clusters the spectrum into bands and bands into sub-carriers, any channel or individual sub-carrier can be dynamically turned on and off in order to comply with any change desired.
- Excellent robustness to narrow-band interference: OFDM is robust against single tone and narrow-band interferers.
- Antenna is easier to design: OFDM has an inherent robustness against gain, phase, and group delay variation which are crucial metrics for antenna design.

UWB has a distinctive property of providing high data rates at only limited range. This will provide enhancements for wireless personal area networking (WPAN): wireless content transfer, high speed data rates in low power for mobile users, cable-less communication.

Low power feature makes it a good choice for sensor networks where the tiny, low cost, low powered sensor nodes form a network for ubiquitous access. They gather the local information and send them to a center for processing. The sensor network is considered to be deployed wherever there is a need for gathering local information for making life easier such as finding intruders, empty parking spots, forest fires, and maintaining buildings, etc.

Chapter 13

IEEE 802.16 and WiMAX

13.1 Introduction

The IEEE 802.16 Air Interface Standard is a wireless technology proposed to revolutionize the broadband wireless access industry. It is designed to provide wireless last-mile broadband access in the Metropolitan Area Network, delivering replaceable performance as in traditional cable, DSL or T1.

The initial version of the 802.16 standard employs a point-to-multipoint (PMP) architecture in 10-66 GHz and requires line-of-sight towers. The 802.16a extension, ratified in March 2003, does not require line-of-sight transmission and lowers the frequencies to the range of 2-11 GHz which could be the last mile link between the national network and the high speed WLANs. It boasts a 31-50 mile range and 70 Mbps data transfer rates in the form of PMP and mesh technologies.

There are additional sub-standards; 802.16b for quality of service, 802.16c for interoperability with protocols and test-suite structures, 802.16d for developing access, 802.16e for mobility support. IEEE

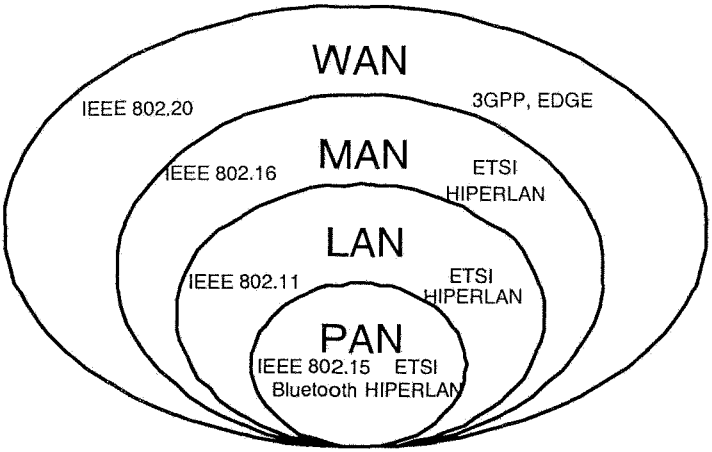


Figure 13.1: Global wireless standards

802.20 will be another standard of WiMAX to address wide-area networks as seen in Figure 13.1.

The overall vision for 802.16 is to allow carriers to set up base stations connected to a public where each base station creates a mesh network to support fixed subscribers, probably mounted on rooftops. The base stations would have MAC protocol to employ interoperable network and allocate uplink and downlink bandwidth to subscribers. 802.16 network could anchor 802.11 hotspots as well as servicing end user directly.

13.2 WiMAX

WiMAX (Worldwide Interoperability for Microwave Access) is the popular name of the 802.16 that is currently being developed. WiMAX network consists of base station (BS) and subscriber station (SS). BS controls the cell in terms of activity, admission, quality of service.

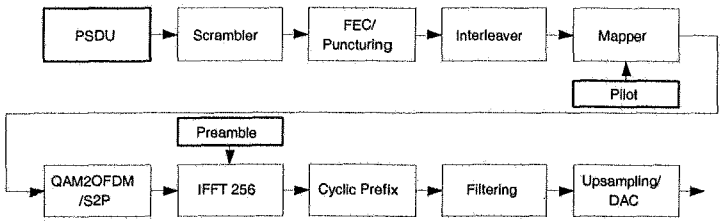


Figure 13.2: WirelessMAN OFDM transmitter © IEEE

They often use fixed antenna at the SS but adaptive-antennas in subscriber appears to be another good choice.

Range is practically around 5 miles or less but when OFDM is used, it could go up to 30 miles via lowest data rate. 802.16 can be implemented by point-to-point (P2P) or mesh topology where the range is increased compared to PMP mode.

13.3 WirelessMAN OFDM

WirelessMAN OFDM deploys 256-FFT in which there are 55 guard band sub-carriers and 8 pilot signal sub-carriers. The number of used sub-carriers are fixed to 192. Table 13.1 shows the symbol parameters.

The guard time is selected according to the terrain type, the worst case delay spread is around 5.24 μsec and can be supported with 1/4 guard time for a channel bandwidth of 10 MHz. In a typical mobile environment with omni-directional antenna in subscriber stations, a typical 10 μs delay spread can be supported with 1/4 guard time for 5 MHz channel bandwidth [176]. Data rates achieved can range from 1 Mbps to 74 Mbps as shown in Table 13.2.

IEEE 802.16 offers some optional features: Sub-channelization in uplink is a useful feature when the SS is power limited. Dividing the

Parameters	Value
NFFT	256
Nused	192
Fs/BW	1.75 (licensed) 8/7 (license exempt) 7/6 (others)
Tg/Tb (Guard Time)	1/4,1/8,1/16,1/32
Guard Band	28 Low, 27 High
Npilot	8

Table 13.1: OFDM symbol parameters for WirelessMAN OFDM

Modulation	QPSK	QPSK	16QAM	16QAM	64QAM	64QAM
Code Rate	1/2	3/4	1/2	3/4	2/3	3/4
1.75 MHz	1.04	2.18	2.91	4.36	5.94	6.55
3.5 MHz	2.08	4.37	5.82	8.73	11.88	13.09
7.0 MHz	4.15	8.73	11.64	17.45	23.75	26.18
10.0 MHz	8.31	12.47	16.63	24.94	33.25	37.40
20.0 MHz	16.62	24.94	33.25	49.87	66.49	74.81

Table 13.2: Data rates achieved in 802.16 with a 1/32 guard time (Mbps)

sub-carriers into 16 clusters can achieve 12-dB link budget enhancement and each cluster can be assigned to a subscriber in uplink. More frequent repetition of preamble is another provision to handle time variation in the channel [141].

13.4 802.16 MAC

The MAC is responsible for air transmission. The FDD system creates separate uplink and downlink subframes. The TDD system on the other hand, in a given channel assigns a uplink or downlink channel.

The mapping of locations is done by a preamble (DL_MAP and UL_MAP). In the downlink, SS listens to all the MPDUs in turn and identifies the frame location destined for himself by inspecting at DL_MAP. Uplink management is handled by a contention-based multiple access scheme. Depending on the service type, the slot can be guaranteed or achieved by polling from BS or contention. Contention access takes place in slots and each slot is divided into minislots where SS contend for that minislot [176]. The subscription requires exchanging timing to account for transit delays and path loss and power adjustment information. For this purpose SS transmits a ranging request (RNG-REQ) packet to enable BS for identify the timing and receive ranging response (RNG-RSP) packet from the BS that gives timing and power adjustment information. There are bandwidth request contention slots used by SS and when it is granted, SS uses non-contention slots allocated by the BS. BS controls the length of the contention slot.

802.16 establishes a connection oriented service. Connections are categorized into two types: management and transport. Management connections are established initially to allow transport connections which carry typically user data. A user seek for a service first scans a downlink signal from a base station and synchronizes to it. SS uses uplink channel descriptor to identify the parameters. SS then does ini-

tial ranging to alert the BS about its presence. Using the management messages, SS and BS settle the basic connection parameters. SS then gains access to the system and establish IP connectivity. During the connection, periodic ranging is necessary to maintain a link, this synchronizes the timing and power of the SSs transmission to the BS. Parameters are updated depending on the channel condition or distance to the BS. Same messages are used as in the initial ranging but now the connection id is given not 0 [141, 177].

13.5 Conclusion

The cost and complexity associated with traditional wired cable and telephone infrastructure could not provide satisfying coverage and caused gaps in the broadband coverage. WiMAX is expected to extend the potential of Wi-Fi to far longer distances. It gives the opportunity to offer the speed, multimedia support and ubiquity and provides economically viable market opportunities for operators, wireless Internet service providers, and equipment manufacturers.

Chapter 14

Future Trends

Multi-carrier communications are currently in a state of very rapid development, both in the introduction of new techniques and in application to a wide variety of channels and services. Advances in algorithmic development and in the capabilities of signal processing devices will surely lead to substantial improvement in performance of multi-carrier systems. This in turn will lead to further application in systems where the unique advantages of multi-carrier modulation can be applied, while overcoming its drawbacks, at reasonable cost.

14.1 Comparison with Single Carrier Modulation

In Reference [178] a comparison was made of the relative advantages of OFDM and single carrier modulation in the ADSL application. The Table 14.1 indicates which system is favored for each of several issues.

The comparisons also apply for environments other than ADSL.

Issue	Single Carrier	Multi Carrier	Equivalent
Performance in Gaussian noise		√	
Sensitivity to impulse noise (uncoded)	√		
Sensitivity to narrowband noise (uncoded)	√		
Sensitivity to clipping	√		
Sensitivity to timing jitter and phase noise	√		
Latency (delay)	√		
Computations per unit time		√	
Cost and power consumption in analog sections	√		
Adaptability of bit rate		√	

Table 14.1: Relative advantages of single carrier and multi-carrier modulation. An "√" denotes the system with better performance or lower cost.

As shown in Chapter 3, bandwidth efficiency and performance in the presence of colored Gaussian noise ideally are equal, if the single carrier system uses optimum decision feedback equalization and the multi-carrier system employs optimum bit loading. The relative computation per unit time is less for multi-carrier in the ADSL environment where the long equalizer needed to correct the highly dispersive channel response in a single carrier implementation requires more computation than the short equalizer and FFT in multi-carrier. The higher cost and power drain of analog components for multi-carrier, particularly the transmitter power amplifier, results from the necessity to keep clipping at a minimum. The sensitivity to narrowband noise and jitter are due to the overlapping and slowly decaying sub-carrier spectra in multi-carrier modulation.

Future enhancements to OFDM give a promise of reversing some of the cases in which OFDM compares unfavorably, in particular those associated with clipping and the overlapping sub-carrier spectra in current OFDM implementations.

14.2 Mitigation of Clipping Effects

In Chapter 4, it was shown that the error probability resulting from clipping may be quite tolerable in most cases. However, the need to limit the cost, and even more importantly the power drain, of components such as the transmitter power amplifier provides an incentive to keep the peak-to-average power ratio as small as possible. Several approaches have been proposed to accomplish this.

Probably the simplest solution is to allow clipping at a relatively low level, and to use standard error control coding to correct the resultant errors. Of course this also reduces errors caused by other degradations, in particular additive noise. In [179] it is shown that this simple approach may compare favorably with other techniques when the clipping level is not reduced to a value that is too low. Another receiving

scheme involves attempting to reconstruct the clipped signal [180].

A large number of proposals involve altering the transmitted signal so as to keep the peak value low, and in most cases notifying the receiver of the alteration. Because the probability of a clip is relatively low in any case, the side information required to perform this notification can be quite low compared with the actual data [97]. One relatively simple approach is to buffer each OFDM symbol before transmission and to detect if the clipping level would be exceeded. If so, the level is reduced and the reported to the receiver, possibly on a sub-carrier dedicated to this purpose.

Other techniques [28, 181, 29] use block codes to produce more than one mapping of input data to OFDM symbols, and transmit the one with lowest peak. The mappings typically involve different rotations of the sub-carrier constellations. Side information is sent to the receiver as before to report which mapping was used. In practice, only one or two bits of side information per OFDM symbol are needed. If unused sub-carriers are available, then they could be modulated in a way that reduces the peak value of each OFDM symbol. In this case, no side information needs to be transmitted because those sub-carriers are ignored in any case. In Reference [181], it is shown that for an error probability of 10-7, the peak-to-average ratio can be reduced from 13.3 dB to 11.1 dB by the use of two mappings, or down to 9.5 dB with four mappings. Reduction of the peak value of an OFDM signal is now an active field of research, involving trade-offs among performance, spectral efficiency, and complexity of implementation. Incorporation of some of these techniques in certain applications of OFDM may be expected in the near future.

14.3 Overlapped Transforms

When the time window used in modulation is a square one equal to the OFDM symbol duration, the spectrum of each sub-carrier is of the

sinc form. Each sub-carrier spectrum therefore overlaps many of the others. This leads to high inter-channel interference when orthogonality is lost for any reason, such as inter-symbol interference, frequency offset, or phase noise. In addition, the susceptibility to narrow-band interference can be severe in that many sub-channels could be affected by a single interferer. The problem can partially be relieved by the use of windowing as described in Chapter 2. However, there is only the duration of the cyclic prefix to provide room for the window shaping, so that substantial overlap of sub-carrier spectra will remain.

Much better spectral shaping can be achieved if processing is no longer constrained to the duration of a block corresponding to a single OFDM symbol. If several blocks are allowed to overlap in time, then a very gradual windowing function can be introduced. Reference [182] describes the performance that can be achieved by a technique called Discrete Wavelet Multi-Tone (DWMT) modulation. In particular, for an overlap of 8 blocks, the first sidelobe of each sub-carrier spectrum is reduced from -13.5 dB in DMT to -45 dB. Performance in noise is substantially unchanged, and sensitivity to various inaccuracies is greatly reduced. Furthermore, no cyclic prefix is needed, leading to maximum spectral efficiency. However, DWMT requires substantially greater computational complexity, and latency is increased by the degree of overlap. The DWMT transmitter is shown in Figure 14.1.

As in DMT, the filter for sub-carrier n, $F_n(z)$ is centered at frequency n/T with nulls at the center frequencies of other sub-carriers. Unlike DMT, its impulse response is of length gT, where $g > 1$. The set of filters performs a wavelet transform that is not described in the reference. The equivalent of the filter bank can be implemented by a fast algorithm similar to the FFT. The impulse responses are orthogonal both over frequency index and over time shifts as in

$$\sum_k f_k^n f_{k-iN}^m = K\delta_i\delta_{n-m}. \tag{14.1}$$

The filters are also chosen so that their spectra decays very rapidly.

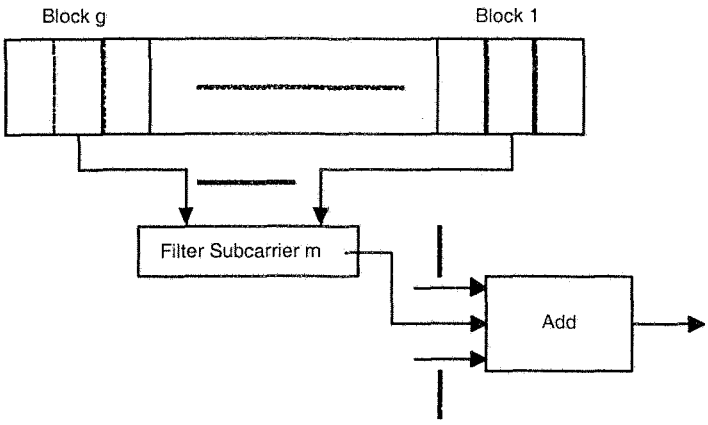


Figure 14.1: DWMT transmitter

Overlap of sub-carriers is therefore greatly reduced. Increasing g enhances this spectral confinement.

With input symbols D_m the line signal is of the form

$$d(z) = \sum_{n=0}^{N-1} D_n(z) F^m(z) \tag{14.2}$$

If there is any dispersion at all over the channel, then unlike in DMT with cyclic prefix, orthogonality among the sub-carriers is lost. This is corrected by means of a matrix equalizer whose size can be kept relatively small due to the very low spectral overlap of sub-carriers spaced a moderate number apart. The matrix equalizer must not only correct for inter-symbol interference in each sub-carrier, but also inter-channel interference among the sub-carriers. Because the frequency overlap among the sub-carriers decays very quickly, only immediately adjacent sub-carriers typically need be accounted for.

The DWMT receiver is illustrated below. The incoming signal

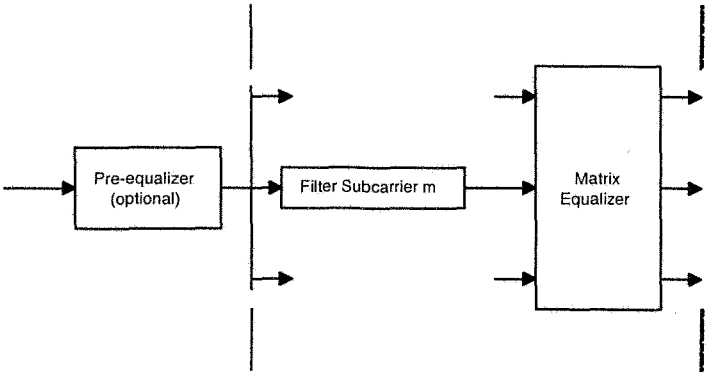


Figure 14.2: DWMT receiver

may first be partially equalized by an optional pre-equalizer. This can relieve the computational burden of the matrix post-equalizer. A wavelet transform, matched to the transmitter transform, is performed by the equivalent of the filter bank shown. Again a fast algorithm can be used. This is followed by the matrix equalizer which includes adjacent sub-carrier taps as well as the usual equalization for each sub-carrier.

The DWMT technique clearly requires more computation than DMT. Reference [182] reports that the complexity in a modulator and demodulator is increased by a factor of approximately $1.25\frac{1+g+\log_2 N}{\log_2 N}$, which is not especially large.

The DWMT technique is a promising one. Not only does it reduce susceptibility to narrow-band interference and loss of orthogonality, but the increase in spectral efficiency due to elimination of guard times between OFDM symbols may well be worth the additional cost in many applications. However, the increase in latency must be considered in some applications.

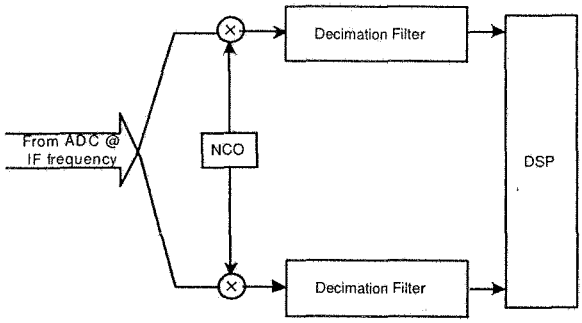


Figure 14.3: IF sampling and digital radio down conversion

14.4 Advances in Implementation

"Moore's Law" concerning the exponential increase in the number of elements that can be included on a chip is expected to hold for the next few years. One consequence will be the availability of general purpose signal processing chips that can perform advanced algorithms for fairly high rate OFDM systems. Another is the availability of circuitry that is dedicated to performing common functions efficiently, either as part of a signal processing chip or as an auxiliary chip. Functions now available include FFTs, Viterbi decoders, Reed-Solomon coders and decoders, etc. Implementation of the butterfly operation of FFT in a modern DSP is significantly more efficient than before. Execution of a 1024-point bit reversed complex Fast Fourier Transform in a new DSP takes less than 50 microseconds.

Analog/Digital converters have been critical components in terms of cost in the past. Because of the high peak-to-average ratio of OFDM signals, this has been a serious issue for such systems. However, recent advances in signal conversion devices have reduced this concern. Availability of high speed A/D devices with high spurious-free dynamic range (SFDR) and signal-to-noise ratio (SNR) at reasonable cost has resulted in further research and development in *Digital Radio*

architecture. By using a combination of high speed A/D and digital down conversion, complexity of an analog front-end is reduced and the system architecture is more flexible. Programmability of digital radio is an attractive feature for multi-mode and multi-band wireless application.

Characteristics of power amplifiers are important for OFDM system physical layer performance such as bit error rate level, as discussed in Chapter 4, and for power consumption requirements. Power amplifiers are the largest source of battery drain in wireless systems. Recent advances in GaAs MMIC have resulted in highly efficient wide-band power amplifiers suitable for OFDM applications.

Bibliography

[1] Walrand, J., Varaiya P., *High-Performance Communication Networks*. Morgan Kaufmann Publishers, 2000.

[2] Rappaport, T. S, *Wireless Communications*. Prentice Hall, 1996.

[3] Gitlin, R.D., Hayes J.F., Weinstein S.B. *Data Communications Principles*. New York: Plenum, 1992.

[4] Doelz, M.L., Heald E.T., Martin D.L. *Binary Data Transmission Techniques for Linear Systems*. Proc. I.R.E.; May 1957; 45: 656-661.

[5] Franco, G.A., Lachs G. *An Orthogonal Coding Technique for Communications*. I. R. E. Int. Conv. Rec.; 1961; 8: 126-133.

[6] Chang, R.W. *Synthesis of Band-Limited Orthogonal Signals for Multichannel Data Transmission*. Bell Sys. Tech. J.; Dec 1966; 45: 1775-1796.

[7] Shnidman, D.A. *A Generalized Nyquist Criterion and an Optimum Linear Receiver for a Pulse Modulation System*. Bell Sys. Tech. J.; Nov 1966; 45: 2163-2177.

[8] Saltzberg, B.R. *Performance of an Efficient Parallel Data Transmission System.* IEEE Trans. Commun.; Dec 1967; COM-15; 805-811.

[9] Weinstein, S.B., Ebert P.M. *Data Transmission By Frequency Division Multiplexing Using the Discrete Fourier Transform.* IEEE Trans. Commun., Oct 1971; COM-19; 5: 628-634.

[10] Hirosaki, B. *An Orthogonally Multiplexed QAM System Using the Discrete Fourier Transform.* IEEE Trans. Commun.; Jul 1981; COM-29; 7: 982-989.

[11] Rabiner,L.R., Gold B., *Theory and Application of Digital Signal Processing.* Englewood Cliffs, NJ: Prentice Hall, 1975.

[12] Alard, M., Lassalle R., *Principles of Modulation and Channel Coding for Digital Broadcasting for Mobile Receivers.* EBU Review; Aug 1987; 224: 47-68

[13] Bello, P.A., *Characterization of Randomly Time-Variant Linear Channels.* IEEE Trans Comm; Dec 1963; COM-11; 360-393.

[14] Kennedy, R.S., *Fading Dispersive Communication Channels.* New York: Wiley Interscience 1969.

[15] Klein, A., Mohr, W., *Measurement-Based Parameter Adaptation of Wideband Spatial Mobile Radio Channel Models.* IEEE 4th International Symp. on Spread Spectrum Techniques and App; ISSSTA'95; 91-97

[16] Cover, T. M., Thomas J. A., *Elements of Information Theory.* John Wiley & Sons, Inc.; 1991.

[17] Heiskala, J., Terry J., *OFDM Wireless LANs: A Theoretical and Practical Guide.* Sams Publishing; 2002.

[18] Haykin, S., *Communication Systems.* John Wiley & Sons, Inc.; 1994.

[19] Willink, T.J., Wittke P.H., *Optimization and Performance Evaluation of Multi-Carrier Transmission.* IEEE Trans. Info. Theory; Mar 1997; 43: 426-440.

[20] Gallagher, R.G., *Information Theory and Reliable Communication.* New York: Wiley; 1968.

[21] Kalet, I., *The Multitone Channel.* IEEE Trans. Commun.; Feb 1989; 37: 119-124.

[22] Feig, E., *Practical Aspects of DFT-Based Frequency Division Multiplexing for Data Transmission.* IEEE Trans. Commun.; Jul 1990; 38: 929-932.

[23] Sistanizadeh, K., Chow P.S., Cioffi J.M., *Multi-Tone Transmission for ADSL.* IEEE Int. Conf. Commun.; 1993; 756-760.

[24] Chow, P.S., Cioffi J.M., Bingham J.A.C., *A Practical Discrete Multitone Transceiver Loading Algorithm for Data Transmission Over Spectrally Shaped Channels.* IEEE Trans. Commun.; Feb-Apr 1995; 43: 773-775.

[25] Zervos, N.A., Kalet I., *Optimized DFE Versus Optimized OFDM for High-Speed Data Transmission Over the Local Cable Network.* IEEE Int. Conf. Commun.; 1993; 35.2.

[26] Cioffi, J.M., Dudevoir GP, Eyuboglu M.V., Forney G.D. *MMSE Decision Feedback Equalizers and Coding - Part I: Equalization Results.* IEEE Trans. Commun.; Oct 1995; 43: 2582-2594.

[27] Hughes-Hartog, D. U.S. Patent 4,731,816; U.S. Patent 4,833,706.

[28] Mestdagh, D.J.G., Spruyt P.M.P., *A Method to Reduce the Probability of Clipping in DMT-Based Transceivers.* IEEE Trans. Commun.; Oct 1996; 44: 1234-1238.

[29] Shepard, S., Orriss J., Barton S., *Asymptotic Limits in Peak Envelope Reduction by Redundancy Coding in OFDM Modulation.* IEEE Trans. Commun.; Jan 1998; 46: 5-10.

[30] Feig, E., Nadas A., *The Performance of Fourier Transform Division Multiplexing Schemes on Peak Limited Channels.* IEEE Globecom; 1988; 35.4.

[31] Rinne, J., Renfors M., *The Behavior of OFDM Signals in an Amplitude Limiting Channel.* IEEE Int. Conf. Commun.; 1994: 381-385.

[32] Mestdagh, D. J. G., Spruyt P.M.P., Biran B., *Analysis of Clipping Effect in DMT-Based ADSL Systems.* IEEE Int. Conf. Commun.; 1994; 293-300.

[33] Gross, R., Veeneman D., *SNR and Spectral Properties for a Clipped DMT ADSL Signal.* IEEE Int. Conf. Commun.; 1994; 843-847.

[34] VanVleck, J.H., Middletown D., *The Spectrum of Clipped Noise.* Proc. IEEE; Jan 1966; 54: 2-19.

[35] Mazo, J.E., *Asymptotic Distortion Spectrum of Clipped, DC-Biased, Gaussian Noise.* IEEE Trans. Commun.; Aug 1992; COM-40; 8; 1339-1344.

[36] Minkoff, J., *The Role of AM-to-PM Conversion in Memoryless Nonlinear Systems.* IEEE Trans. Commun.; Feb 1985; COM-33; 2: 139-144.

[37] Rice, S.O., *Distribution of the Duration of Fades in Radio Transmission.* Bell Syst. Tech. J.; May 1958; 37: 581-635.

[38] Kac, M., Slepian D., *Large Excursions of Gaussian Processes.* Ann. Math. Stat.; Dec 1959; 30: 1215-1228.

[39] Cramer, H., Leadbetter M.R., *Stationary and Related Stochastic Processes.* New York: Wiley, 1967.

[40] Leadbetter, M.R., Lindgren G., Rootzen H., *Extremes and Related Properties of Random Sequences and Processes.* New York: Springer-Verlag; 1983.

[41] Bussgang, J.J., *Crosscorrelation Functions of Amplitude Distorted Gaussian Signals.* Res. Lab. Of Electronics M.I.T., (Cambridge, Mass.) Tech. Rep; Mar 26, 1952; 216:3.

[42] Blachman, N.M., *Detectors, Bandpass Non-Linearities, and Their Optimization: Inversion of the Chebyshev Transform.* IEEE Trans. Inform. Theory, Jul 1971; IT-17: 398-404.

[43] Craig, J. W., *A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations.* Proc. MILCOM'91; 571-575.

[44] Bahai, A. R. S., Singh M., Goldsmith A. J., Saltzberg B. R., *A New Approach for Evaluating Clipping Distortion in Multi Carrier Systems.* IEEE J. in Sel. Areas in Communication, June 2002; 20: 1037-1046.

[45] Pollet, T., Moeneclay M., *Synchronizability of OFDM Signals.* IEEE Proc. of Global Telecom Conf. (GLOBECOM '95); Nov 1995; 3: 2054-2058.

[46] Lindsay, W.C., Chie C.M., *A Survey of Digital Phase-Locked Loops.* Proc. IEEE; Apr 1981; 9; 4: 410-431.

[47] Pollet, T., VanBladel M., Moeneclay M., *BER Sensitivity of OFDM Systems to Carrier Frequency Offset and Wiener Phase Noise.* IEEE Trans. Commun.; Feb/Mar/Apr 1995; 43; 2/3/4: 191-193.

[48] Robertson, P., Kaiser S., *Analysis of the Effects of Phase Noise in OFDM Systems.* IEEE Int. Conf. Commun.; 1995: 1652-1657.

[49] Armada, G., Calvo M., *Phase Noise and Sub-Carrier Spacing Effects on the Performance of an OFDM Communication System.* IEEE Commun. Letters; Jan 1998; 2; 1: 11-13.

[50] Steendam, H., Moeneclay M., Sari H., *The Effects of Carrier Phase Jitter on the Performance of Orthogonal Frequency Division Multiple Access Systems.* IEEE Trans. Commun.; Apr 1998; 46; 4: 456-459.

[51] Zogakis, T.N., Cioffi J.M., *The Effect of Timing Jitter on the Performance of a Discrete Multitone System.* IEEE Trans. Commun.; Jul 1996; 44; 7: 799-808.

[52] Coleri, S., Ergen M., Puri A., Bahai A., *Channel Estimation Techniques Based on Pilot Arrangement in OFDM Systems.* IEEE Trans on Broadcasting; Sept 2002; 48; 3; 223 -229.

[53] Beek, J.J.,Edfors O.,Sandell M.,Wilson S.K., Borjesson P.O., *On channel estimation in OFDM systems.* IEEE 45th Vehicular Technology Conf.; Jul. 1995; 815-819.

[54] Edfors, O., Sandell M., Beek J.J., Wilson S.K., Brjesson P.O., *OFDM channel estimation by singular value decomposition.* IEEE Trans. on Communications, July 1998; 46; 7; 931-939.

[55] Hsieh, M., Wei C., *Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels.* IEEE Trans. on Consumer Electronics, Feb. 1998; 44; 1.

[56] Steele, R., *Mobile Radio Communications.* London, England: Pentech Press Limited, 1992.

[57] Reimers,U., *Digital video broadcasting.* IEEE Comm. Magazine, June 1998; 36; 6; 104-110.

[58] Cimini,L. J., *Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing.* IEEE Trans. Commun., July 1985; 33; 7; 665-675.

[59] Zhao, Y., Huang A., *A novel channel estimation method for OFDM Mobile Communications Systems based on pilot signals and transform domain processing.* IEEE 47th Vehicular Technology Conf., May 1997; 2089-2093.

[60] Oppenheim, A. V., Schafer R. W., *Discrete-Time Signal Processing.* New Jersey: Prentice-Hall Inc., 1999.

[61] *Digital video broadcasting (DVB): Framing, channel coding and modulation for digital terrestrial television.* Draft ETSI EN300; 744; 1.3.1 (2000-08).

[62] Matloub, S., *Comparison of different channel estimation techniques for MIMO-OFDM systems.* www.stanford.edu.

[63] Li,Y., *Pilot-Symbol-Aided Channel Estimation for OFDM in Wireless Systems.* IEEE Trans on Vehicular Technology, July 2000; 49; 4.

[64] Bingham, J.A.C. *Multicarrier Modulation for Data Transmission: an Idea whose Time has Come.* IEEE Commun. Mag.: May 1990; 28: 5-14.

[65] Rohling, H., May T., *OFDM systems with differential modulation schemes and turbo decoding techniques.* 2000 International Zurich Seminar, February 2000; 251-255.

[66] Nee, R., Prasad R., *OFDM for Wireless Multimedia Communications.* Artech House, January 2000.

[67] Kominakis, C., Fragouli C., Sayed A. H., *Channel estimation and equalization in frequency selective fading.* Asilomar Conf. on Signals, Systems, and Computers, Oct 1999; 427-431.

[68] Bang, C. I., Lee M. H., *An analysis of pilot symbol assisted 16 QAM in the Rayleigh fading channel.* IEEE Trans. Consumer Electronics, Nov. 1995; 41; 1138-1141.

[69] Moon, J. K., Choi S. I., *Performance of channel estimation methods for OFDM systems in a multipath fading channels.* IEEE Trans on Consumer Electr, Feb. 2000; 46; 161-170.

[70] Agrawal, D., Tarokh V., Naguip A., Seshadri N., *Space-time coded OFDM for high data-rate wireless communication over wideband channels.* Proc. IEEE VTC, May 1998; 2232-2246.

[71] Bolcskei, H., Paulraj A., *Space-frequency coded broadband OFDM systems.* Proc. WCNC, May 2000; 1; 1-6.

[72] Yi, G., Letaief K., *Space-frequency-time coded OFDM for broadband wireless communications.* Proc. GLOBECOM, 2001; 1; 519-523.

[73] Li, Y., Seshadri N., Ariyavisitakul S., *Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels.* IEEE J. Select. Areas Commun., May 1999; 17; 461-471.

[74] Li, Y. *Simplified channel estimation for OFDM systems with multiple transmit antennas.* IEEE Trans. Wireless Commun., Jan. 2002.; 1; 67-75.

[75] Du, J., Li Y., *MIMO-OFDM channel estimation based on subspace tracking.* IEEE VTC, April 2003; 2; 1084-1088.

[76] Chow, J., Cioffi J., Bingham J. *Equalizer Training Algorithms for Multicarrier Modulation Systems.* IEEE Int. Conf. Commun; 1993: 761-765.

[77] Falconer, D., Magee F. *Adaptive Channel Memory Truncation for Maximum Likelihood Sequence Estimation.* Bell System Tech. Journal; Nov 1973; 52: 1541-1562.

[78] Lee, E.A., Messerschmitt D.G., *Digital Communication.* Boston: Kluwer Academic Publisher, 1994.

[79] Al-Dhahir, N., Cioffi J., *Optimum Finite Length Equalization for Multicarrier Transceiver.* IEEE Globecom; 1994: 1884-1888.

[80] Cioffi, J.M., Bingham J.A.C., *A Data Driven Multitone Echo Canceller.* IEEE Trans. Commun.; Oct 1994; 42; 10: 2853-2867.

[81] Yang, J., Roy S., Lewis N.H. *Data-Driven Echo Cancellation for a Multitone Modulation System.* IEEE Trans. Commun.; May 1996;44; 5: 2134-2144.

[82] Melsa, P.J.W., Younce R.C., Rohrs C.E., *Impulse Response Shortening for Discrete Multitone Transceivers.* IEEE Trans. Commun.; Dec 1996; 44; 12: 1662-1672.

[83] Ho, M., Cioffi J.M., Bingham J.A.C., *High-Speed Full-Duplex Echo Cancellation for Multitone Modulation.* IEEE Int. Conf. Commun.; 1993: 772-76.

[84] Ho, M., Cioffi J.M., Bingham J.A.C., *Discrete Multitone Echo Cancellation.* IEEE Trans. Commun.; Jul 1996; 44; 7: 817-825.

[85] Lee, D. L., Friedlander B., Morf M., *Recursive Ladder Algorithms for ARMA Modeling.* IEEE Trans. Automatic Control, Aug 1982; AC-27: 4.

[86] Kay, S.M., *Modern Spectral Estimation.* Upper Saddle River, NJ: Prentice Hall, 1988.

[87] Brodersen, R., *EE225 Lecture 16- OFDM Introduction.* http://bwrc.eecs.berkeley.edu/Classes/EE225C/lectures.htm.

[88] Wu, Z., Liu Y., Jing x., *Channel Estimation in OFDM Systems.* http://www.winlab.rutgers.edu/špasojev/courses/projects/

[89] Costello, D.J., Costello Lin., *Error Control Coding.* Englewood Cliffs, NJ: Prentice-Hall, 1983.

[90] Wicker, B., Bhargava V.K., *Reed-Solomon Codes and Their Applications.* New York: IEEE Press, 1994.

[91] Pommier D., Wu Y., *Interleaving or Spectrum Spreading in Digital Radio Intended for Vehicles.* EBU Review; Jun 1986; 217: 31-44.

[92] Ungerboeck, G., *Trellis Coded Modulation with Redundant Signal Sets.* IEEE Commun. Mag.; Feb 1987; 25; 2: 5-21.

[93] Biglieri, D., Divsalar D., McLane P.J., Simon M.K., *Introduction to Trellis-Coded Modulation with Applications.* New York: Macmillan, 1991.

[94] Hagenauer, J., Offer E., Papke L., *Iterative Decoding of Binary Block and Convolutional Codes.* IEEE Trans. Information Theory; Feb 1987; 25; 2: 5-21.

[95] Jung, P., *Comparison of Turbo-Code Decoders Applied to Short Frame Transmission Systems.* IEEE J. on Sel. Areas in Communications; Apr 1996; 14; 3.

[96] Hoeher, P., *New Iterative Turbo Decoding Algorithms.* Proceeding of The International Symposium on Turbo Codes; Brest, France; 1997.

[97] Starr, T., Cioffi J.M., Silverman P., *Understanding Digital Subscriber Line Technology.* Englewood Cliffs, NJ: Prentice Hall, 1999.

[98] Kyees, P.J., McConnell R.C., Sistanizadeh K., *ADSL: A New Twisted Pair Access to the Information Highway.* IEEE Commun. Mag.; Apr 1995; 33: 52-59.

[99] Ahamed, S.V., Bohn P.P., Gottfried N.L., *A Tutorial on Two Wire Digital Transmission in the Loop Plant.* IEEE Trans. Commun.; Nov 1981; 29: 1554-1564.

[100] Adams, P.F., Cook J.W., *A Review of Copper Pair Local Loop Transmission Systems.* Brit. Telecom. Tech. J.; 1989; 7: 17-29.

[101] Werner, J.J., *The HDSL Environment.* IEEE J. on Sel. Areas in Commun.; Aug 1991; SAC-9: 785-800.

[102] Gibbs, J., Addie R., *The Covariance of Near End Crosstalk and its Application to PCM System Engineering in Multipair Cable.* IEEE Trans. Commun.; Feb 1979; 27: 469-477.

[103] Lin, S.H., *Statistical Behavior of Multipair Crosstalk.* Bell Sys. Tech. J.; Jul-Aug 1980; 59: 955-974.

[104] Chow, P.S., Tu J.C., Cioffi J.M., *A Discrete Multitone Transceiver for HDSL Applications.* IEEE J. on Sel. Areas in Commun.; Aug 1991; SAC-9: 895-908.

[105] Barton, M., Chang L., Hsing T.R., *Performance Study of High Speed Asymmetric Digital Subscriber Line Technology.* IEEE Trans. Commun.; Feb 1996; 44: 156-157.

[106] Bingham, J. A. C., *ADSL, VDSL, and Multicarrier Modulation.* New York: Wiley, 2000.

[107] Gupta, P., Kumar P. R., *The Capacity of Wireless Networks.* IEEE Trans. on Information Theory, March 2000; IT-46; 2; 388-404.

[108] O'Hara, B., Petrick A., *The IEEE 802.11 Handbook: A Designer's Companion.* IEEE Product No. SP1118-TBR.

[109] Stallings, W., *Data and Computer Communications.* New York: Maxwell MacMillan, 1991.

[110] Helstrom, C. W., *Statistical Theory of Signal Detection.* London: Pergamon, 1966.

[111] Bahai, A.R.S., *Estimation in Randomly Time-Varying Systems with Applications to Digital Communications.* Ph.D. Dissertation, Univ. of California at Berkeley, 1993.

[112] Viterbi, A.J.,*CDMA, Principles of Spread Spectrum Communication.* Reading, MA: Addison Wesley, 1995.

[113] Takanashi, H., *Proposal of PHY Specification for 5 GHz Band.* Proposal to IEEE 802.11 Committee, Jan 1998.

[114] Bogenfeld, E, et al., *Influence of Nonlinear HPA on Trellis-Coded OFDM for Terrestrial Broadcasting of Digital.* IEEE Proc. of Global Telecom Conf. (GLOBECOM '93); 1433-1438.

[115] Bianchi, G., *Performance Analysis of the IEEE 802.11 Distributed Coordination Function.* IEEE J. on Sel. Areas in Comm., 2000; 18.

[116] Qiao, D., Choi S., Shin K. G., *Goodput Analysis and Link Adaptation for IEEE 802.11a Wireless LANs.* IEEE Trans. on Mobile Computing, 2002; 1; 4.

[117] IEEE 802.11, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.* Standard, IEEE, Aug. 1999.

[118] IEEE 802.11a, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer Extension in the 2.4GHz Band.* supplement to IEEE 802.11 Standard, Sept. 1999.

[119] *Part 11: Wireless LAN, Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer Extension in the 2.4GHz Band.* supplement to IEEE 802.11 Standard, Sept. 1999.

[120] IEEE 802.11e, *Medium Access Control (MAC) Enhancements for Quality of Service (QoS).* IEEE Std 802.11e/D4.4, June 2003.

[121] Xiao Y., *A Simple and Effective Priority Scheme for IEEE 802.11.* IEEE Communication Letters, Feb. 2003; 7; 2.

[122] Ergen, M., *IEEE 802.11 Tutorial.* http://www.eecs.berkeley.edu/~ergen/docs/ieee.pdf, UC Berkeley, June 2002.

[123] Ergen, M., Lee D., Sengupta R., Varaiya P., *Wireless Token Ring Protocol- Performance comparison with IEEE 802.11.* IEEE ISCC, July 2003.

[124] Yeh, H. J., Chen J. C., Lee, C. C., *WLAN standards.* IEEE Potentials, Oct/Nov 2003, 17-22.

[125] Feng, C.C., *IEEE 802.11a WLAN System.* NTU 2003.

[126] Bisdounis, L., *HiPERLAN/2 vs IEEE 802.11a: Physical Layer.* Intracom.

[127] Geier, J., *802.11 Alphabet Soup.* www.wi-fiplanet.com, August, 2002.

[128] Ergen, M., Varaiya P., *Admission Control and Throughput Analysis in IEEE 802.11.* ACM-Kluwer MONET Special Issue on WLAN Optimization.

[129] Tesi, R., Codreanu M., Oppermann I., *Interference Effects of UWB Transmission in OFDM Communication Systems.* Centre for Wireless Communications, University of Oulu, Finland.

[130] Bellorado, J., Ghassemzadeh S.S., Greenstein L. J., Sveinsson T., Tarokh V., *Coexistence of UltraWideband Systems with IEEE-802.11a Wireless LANs.* IEEE Globecom 2003.

[131] Tubbax, J., Come B., Perre L. V., Deneire L., Donnay S., Engels M., *Compensation of IQ imbalance in OFDM systems.* VTC, Fall 2002.

[132] Come, B., Ness R., Perre L. V., Eberle W., Wambacq P., Engels M., Bolsens I., *Impact of front-end non-idealities on Bit Error Rate performances of WLAN-OFDM transceivers.* Interuniversity MicroElectronics Center.

[133] *OFDM (Orthogonal Frequency Division Multiplexing).* Nova Engineering.

[134] *Interuniversity MicroElectronics Center.* http://www.imec.be

[135] Taura, K., Tsujishita M., Takeda M., Kato H., Ishida M., Ishida Y., *A Digital Audio Broadcasting (DAB) Receiver.* IEEE Trans. Consumer Elect.; Aug 1996; 42; 3: 322-327.

[136] Kuchen, F., Becker T.C., Wiesbeck W., *Analytic Bit Error Rate Determination for Digital Audio Broadcasting.* IEE Int. Broadcasting Conv.; Sep 1996: 261-266.

[137] Leclerc, M., Scalart P., Fautier P., Huynh H.T., *Performance Analysis of an In-Band COFDM/FM Digital Audio Broadcasting System.* IEEE Trans. Broadcasting; Jun 1997; 43; 2: 191-198.

[138] Kroeger, B.W., Peyla P.J., *Compatibility of FM Hybrid In-Band On-Channel (IBOC) System for Digital Audio Broadcasting.* IEEE Trans. Broadcasting; Dec 1997; 43; 4: 421-430.

[139] Mignone, V., Morello A., Visintin M., *An Advanced Algorithm for Improving DVB-T Coverage in SFN.* IEEE Int. Broadcasting Conv.; Sep 1997: 534-540.

[140] Sari, H., Karam G., Jeanclaude I., *Transmission Techniques for Digital Terrestrial TV Broadcasting.* IEEE Commun. Mag.; Feb 1995; 33; 2: 100-109.

[141] *IEEE Std. 802.16-2001 IEEE Standard for Local and Metropolitan area networks Part 16: Air Interface.* IEEE Std. 802.16-2001; 2002; 0_1- 322.

[142] Koffman, I., Roman V., *Broadband Wireless Access Solutions Based on OFDM Access in IEEE 802.16.* IEEE Communications Magazine, April 2002.

[143] Ergen, M., Puri A., *MEWLANA-Mobile IP Enriched Wireless Local Area Network Architecture.* IEEE 56th Vehicular Technology Conf., 2002; 4; 2449-2453.

[144] Wong, C. Y., Cheng R. S., Letaief K. B., Murch R. D., *Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation.* IEEE J. of Selec. Areas in Commun., 1999; 10; 17; 1747 -1758.

[145] Kivanc, D., Liu H., *Subcarrier Allocation and Power Control for OFDMA.* Thirty-Fourth Asilomar Conf. on Signals, Systems and Computers, 2000; 1; 147 -151.

[146] Pietrzyk, S., Janssen G. J. M., *Multiuser Subcarrier Allocation for QoS Provision in the OFDMA Systems.* IEEE 56th Vehicular Technology Conference, 2002; 2; 1077 -1081.

[147] Zhang, Y., Letaief K. B., *Multiuser Subcarrier and Bit Allocation along with Adaptive Cell Selection for OFDM Transmission.* IEEE Int. Conf. on Comm., 2002; 2; 861 -865.

[148] Kim, I., Lee H. L., Kim B., Lee Y. H., *On the Use of Linear Programming for Dynamic Subchannel and Bit Allocation in Multiuser OFDM.* IEEE Global Telecommunications Conf., 2001; 6; 3648 -3652.

[149] Rhee, W., Cioffi J. M., *Increase in Capacity of Multiuser OFDM System Using Dynamic Subchannel Allocation.* IEEE 51st Vehicular Technology Conf., 2000; 2; 1085 -1089.

[150] Wong, C. Y., Tsui C. Y., Cheng R. S., Letaief K. B., *A Real-time Sub-carrier Allocation Scheme for Multiple Access Downlink OFDM Transmission.* IEEE 50th Vehicular Technology Conference, 1999; 2; 1124 -1128.

[151] Koutsopoulos, I., Tassiulas L., *Channel state-Adaptive techniques for Throughput Enhancement in Wireless Broadband Net-*

works. IEEE 20th Annual Joint Conf. of the IEEE Computer and Communications Societies; 2; 757 -766.

[152] Nogueroles, R., Bossert M., Donder A., Zyablov V., *Performance of a Random OFDMA System for Mobile Communications*. Int. Zurich Seminar on Transmission and Networking, 1998; 37 -43.

[153] Ibars, C., Bar-Ness Y., *Comparing the Performance of Coded Multiuser OFDM and Coded MC-CDMA over Fading Channels*. IEEE Global Telecommunications Conf., 2001.; 2; 881 -885.

[154] *Digital video broadcasting (DVB): Framing, channel coding and modulation for digital terrestrial television*. Draft ETSI EN300 744 V1.3.1 (2000-08).

[155] Barbarossa, S., Pompili M., Giannakis G. B., *Time and Frequency Synchronization of Orthogonal Frequency Division Multiple Access Systems*. IEEE Int. Conf. on Communications; 6; 1674 -1678.

[156] Bohdanowicz,A., Janssen G. J. M., Pietrzyk S., *Wideband Indoor and Outdoor Multipath Channel Measurements at 17 GHz*. IEEE 50th Vehicular Technology Conf., 1999; 4; 1998 -2003.

[157] Corson, M. S., Laroia R., O'Neill A., Park V., Tsirtsis G., *A New Paradigm for IP-Based Cellular Networks*. 20th Annual Joint Conf. of the IEEE Computer and Communications Societies, Nov. -Dec. 2001; 2; 757 -766.

[158] Knopp, R., Humblet P., *Information capacity and power control in single cell multiuser communications*. IEEE Int. Conf. on Comm., 1995; 1; 331 -335.

[159] Khun, H. W., *The Hungarian Method for the Assignment Problem*. Naval Research Logistics Quarterly, 1955; 2; 83-97.

[160] Ergen, M., Coleri S., Varaiya P., *QoS Aware Resource Allocation Techniques for Fair Scheduling in OFDMA Based Broadband Wireless Access Systems.* IEEE Trans. on Broadcasting, December 2003; 49; 362-370.

[161] Guo, Z., Zhu W., *Performance Study of OFDMA vs. OFDM/SDMA.* IEEE 55th Vehicular Technology Conf., May 2002; 2.

[162] Bana, S. V., Varaiya P., *Space division multiple access (SDMA) for robust ad hoc vehicle communication networks.* IEEE Intelligent Transportation Systems, 2001; 962-967.

[163] Yin, H., Liu H., *Performance of Space-Divison Multiple-Access (SDMA) With Scheduling.* IEEE Trans. on Wireless Communications, October 2002; 1; 4; 611-618.

[164] Vook, F. W., Baum K. L., *Adaptive Antennas for OFDM.* IEEE 48th Vehicular Technology Conf., 1998; 1; 606-610.

[165] Kim, C. K., Lee K., Cho Y. S., *Adaptive Beamforming Algorithm for OFDM Systems with Antenna Arrays.* IEEE Trans. on Consumer Electronics, 2000; 46; 1052-1058.

[166] Thoen, S., Perre L. V. , Gyselinckx B., Engels M., *Adaptive Loading for OFDMA/SDMA-based Wireless Local Networks.* IEEE Global Telecommunications Conf., 2000; 2; 767-771.

[167] Royer, E. M., Toh C. K. , *A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks.* IEEE Personal Comm., April 1999; 46-55.

[168] Johnsson, M., *HiperLAN/2-The Broadband Radio Transmission Technology Operating in the 5 GHz Frequency Band.* HiperLAN/2 Global Forum, 1999.

[169] Bahai, A., Fettwies G., *Results on MC-CDMA Receiver Design.* IEEE Int. Conf. on Comm., June 1995; 915-918.

[170] Guo, Z., Zhu W., *Performance Study of OFDMA vs. OFDMA/SDMA*. Microsoft Research.

[171] Viswanath, P., Tse D. N. C., Laroia R., *Opportunistic Beamforming Using Dumb Antennas*. IEEE Trans. on Information Theory, June 2002; 48; 1277-1294.

[172] Aiello, R., Larsson T., Meacham D., Kim Y., Okado H., *Multi-Band Performance Tradeoffs*. IEEE P802.15-03/209r0, May, 2003.

[173] Batra, A., Balakrishman J., Dabak A.,*Time-Frequency Interleaved Orthogonal Frequency Division Multiplexing (TFI-OFDM)*. Physical Layer Submission to 802.15 Task Group 3a: IEEE P802.15-03/142r2, May, 2003.

[174] Rangan, S., *Flash-OFDM: A New Wireless Technology for Mobile Internet Access*. http://www.flarion.com/.

[175] *IEEE 802.16 and WiMAX-Broadband Wireless Access to Everyone*. Intel White Paper.

[176] Jonhston, D., Yaghoobi H., *Peering Into the WiMAX Spec: Part 1*. http://www.commsdesign.com.

[177] Jonhston, D., Yaghoobi H., *Peering Into the WiMAX Spec: Part 2*. http://www.commsdesign.com

[178] Saltzberg, B.R., *Comparison of Single Carrier and Multitone Digital Modulation for ADSL Applications*. IEEE Commun. Magazine; Nov 1998; 36; 11: 114-121.

[179] Wulich, D., Goldfeld L., *Reduction of Peak Factor in Orthogonal Multicarrier Modulation by Amplitude Limiting and Coding*. IEEE Trans. Commun.; Jan 1999; 47: 18-21.

[180] Kim, D., Stuber S.L., *Clipping Noise Mitigation for OFDM by Decision-Aided Reconstruction*. IEEE Commun. Let.; Jan 1999; 3; 1: 4-6.

[181] Baum, R. W., Fischer R. F. H., Huber J. B., *Reducing the Peak-to-Average Ratio of Multicarrier Modulation by Selective Mapping*. Electronic Letters; Oct 1996; 32: 2056-2057.

[182] Sandberg, S.D., Tzannes M.A., *Overlapped Discrete Multitone Modulation for High Speed Copper Wire Communication*. IEEE J. on Sel. Areas in Commun.; Dec 1995; SAC-13; 9: 1571-1585.

[183] Hara, S., Prasad R., *Overview of Multicarrier CDMA*. IEEE Commun. Magazine; Dec 1997; 35; 12: 126-133.

List of Figures

1.1 A basic PAM system 2

1.2 A basic QAM system 4

1.3 A QAM constellation 5

1.4 General form of QAM generation 6

1.5 The Collins Kineplex receiver 8

1.6 An early version of OFDM 9

1.7 OFDM modulation concept: Real and Imaginary components of an OFDM symbol is the superposition of several harmonics modulated by data symbols 10

1.8 Spectrum overlap in OFDM 11

1.9 Spectrum of OFDM signal 12

1.10 Very basic OFDM system 13

1.11 A typical wireless OFDM architecture 14

2.1 Wireless propagation 16

2.2 Received signal versus distance © [1] 18
2.3 Doppler effect 20
2.4 Fading illustration © [2] 24
2.5 A flat fading channel where f_s is sampling frequency 25
2.6 Components of a multi-carrier system 28
2.7 A communication channel in information theoretic view 29
2.8 Constellation diagram 31
2.9 An FFT implementation (decimation in time) 37
2.10 System with complex transmission 38
2.11 System with real transmission 40
2.12 Two different techniques for FFT butterfly 41
2.13 Partial FFT (DIT) 42
2.14 Prefix and postfix cyclic extension 43
2.15 Pilot positioning in time and frequency 45
2.16 Typical impulse response of a wireless channel . . . 46
2.17 Relationship between system functions 46
2.18 Relationship between correlation functions 49
2.19 OFDM time and frequency span 50
2.20 Time-varying channel 50

3.1 Optimum transmit power distribution by the water pouring theorem 62

4.1 Power amplifier 1 dB compression point 72

4.2 Relationship between instantaneous and envelope clipping 73
4.3 Excursions of a Gaussian random process above $|l|$. 74
4.4 Normalized instantaneous distortion spectrum 76
4.5 Memoryless nonlinear mapping 76
4.6 Conditional probability of symbol error 81
4.7 Symbol error probabilities due to clipping 84
4.8 Symbol error probabilities using the additive Gaussian noise approach 84
4.9 Simulated and analytical symbol error probability floor due to clipping 85
4.10 Symbol error probability floor in a Rayleigh fading channel 89
4.11 Baseband equivalent system model for clipping at OFDM receiver 89
4.12 BER in a Rayleigh fading channel, clip level = 3σ . . 91
4.13 BER in a Rayleigh fading channel, clip level = 6σ . . 91
4.14 BER for uncoded 16-QAM in a multipath fading channel 92
4.15 BER for uncoded 64-QAM in a multipath fading channel 92
4.16 Comparison of simulated and analytical symbol error probability floor for 64-tone OFDM signal in Rayleigh fading 93
4.17 Wireless LAN adjacent channels 97

5.1 ACI caused by frequency offset 104

5.2 Synchronization sequence in OFDM 105
5.3 Frame synchronization 106
5.4 Timing offset estimate 106
5.5 Frequency offset estimation 108
5.6 A digital implementation of coarse frequency offset estimation using maximum likelihood criteria 109
5.7 A typical oscillator phase noise power spectrum . . . 110
5.8 Classical phase-locked loop 111
5.9 A second order digital phase-locked loop 112
5.10 Typical spectrum of a phase-locked oscillator 113
5.11 Inter-carrier interference noise 116

6.1 Observation of the distortion by channel [87] 118
6.2 Pilot positioning in time and frequency 119
6.3 Baseband OFDM system 120
6.4 Pilot arrangement 123
6.5 Time domain interpolation 129
6.6 LMS scheme 132
6.7 16QAM modulation with Rayleigh fading (Doppler Freq. 70Hz) 133
6.8 16QAM modulation with Rayleigh fading (SNR 40dB) 134
6.9 MIMO-OFDM architecture 135
6.10 General decision feedback equalizer 139
6.11 Equalizer configuration in training mode 143

6.12 Equalization using order update 148
6.13 An OFDM system with frequency domain equalization 149
6.14 LMS adaptation of a frequency domain equalizer multiplier 152
6.15 Time and frequency domain equalization 153
6.16 An example system illustrating echo cancellation requirement 154
6.17 A basic echo canceller 155
6.18 Frequency domain echo cancellation 156
6.19 Combination echo cancellation for symmetric transmission 157
6.20 Echo cancellation where the transmit rate is lower than receive rate 158
6.21 Echo cancellation where the receive rate is lower than transmit rate 159
6.22 Geometric interpretation of linear predictors 162
6.23 Lattice filter representation of predictors 163

7.1 Two-dimensional coding for OFDM 168
7.2 Construction of a canonical block code 169
7.3 Performance with block coding over a Gaussian channel 170
7.4 Implementation of periodic interleaving 172
7.5 CRC implementation 173
7.6 Generation of a convolutional code 174
7.7 State diagram of a convolutional code 175

7.8 Viterbi algorithm applied to a simple convolutional code 176
7.9 Performance of a convolutional code over a Rayleigh fading channel 179
7.10 Concatenated coding with interleaving 180
7.11 An expanded constellation partitioned for Trellis coding 182
7.12 Implementation of a Trellis code 182
7.13 State diagram of the Trellis code 183
7.14 Metrics used for Viterbi decoding 184
7.15 A typical turbo encoder 186
7.16 A turbo decoder structure 188

8.1 Attenuation constant of 24-gauge (0.5 mm) wire-pair 193
8.2 A bridged tap and its echoes 193
8.3 Crosstalk mechanism 194
8.4 Capacitive crosstalk mechanism 195
8.5 Inductive crosstalk mechanism 196
8.6 Near-end crosstalk generation 198

9.1 Comparison of WLAN standards [124] 205
9.2 IEEE 802.11 networks 207
9.3 Relationship between IEEE 802.11 services © [108] 213
9.4 IEEE 802.11 architecture © IEEE 214
9.5 The hidden node problem 215
9.6 RTS and CTS solution 216
9.7 Exposed terminal 217

9.8 MAC frame 218
9.9 NAV Sub-module finite state machine 221
9.10 DCF finite state machine 222
9.11 Idle sub-module finite state machine 223
9.12 Backoff sub-module finite state machine 224
9.13 Frame sequence and retry sub-module finite state machine 225
9.14 Timing of the 802.11 DCF: Note that station 6 cannot hear station 2 but station 1 225
9.15 Time realization of DCF 227
9.16 Markov Chain model for the IEEE 802.11 DCF Model in normal operating condition 228
9.17 p and τ values versus n 231
9.18 Throughput versus number of active nodes 233
9.19 Throughput versus number of active nodes for fixed users and fixed total load 234
9.20 Macro states in PCF 235
9.21 Timing diagram for PCF 236
9.22 Contention free period determination 237
9.23 Access point CFP finite state machine 238
9.24 Station CFP finite state machine 239
9.25 Frequency hopping 246
9.26 Transmit and Receive FHSS © IEEE 247
9.27 Transmit and Receive DSSS © IEEE 251
9.28 Wireless LAN system architecture 256

9.29 Logical block diagram of OFDM architecture 256
9.30 Scrambler/Descrambler 257
9.31 Convolutional encoder (k=7) © IEEE 258
9.32 Transmitter spectrum mask 261
9.33 OFDM subcarrier allocation 261
9.34 Nonlinear HPA model 262
9.35 Format of an OFDM frame © IEEE 264
9.36 Logical representation of an OFDM frame © IEEE . 265
9.37 Format of an OFDM frame © IEEE 265
9.38 IEEE 802.11a PLCP/PMD transmitter state machine 267
9.39 Channel impulse response for typical wireless LAN medium 268
9.40 IEEE 802.11a PLCP/PMD receiver state machine . . 270
9.41 Frame synchronization and AGC 272
9.42 Channel estimation block 273
9.43 Simplified schematic of the IEEE 802.11a simulation model for the transmitter [132] 275
9.44 An example of power spectrum 276
9.45 An example of FIR filter where the required SIR at the receiver is 30dB 277
9.46 Bit error rate of IEEE 802.11a in Rayleigh Channel . 277
9.47 Bit error rate comparison of modulation schemes in AWGN 278
9.48 Phase noise (65dB of 10KHz) 279

9.49 Channel estimation compared to Perfect Channel Knowledge (PCK) 280

9.50 An example of constellation diagram with IQ imbalance 281

9.51 IQ imbalance when there is no coding and no perfect channel knowledge 282

9.52 IQ imbalance when there is coding and perfect channel knowledge 282

9.53 Impact of clipping threshold on performance 283

9.54 Quantization performance in AWGN 284

9.55 Quantization effect in QPSK (AWGN) 284

9.56 Quantization effect in 16QAM (AWGN) 285

9.57 Nonlinear distortion in IEEE 802.11a 286

9.58 Hard or Soft decision coding 287

9.59 Co-channel interference 287

9.60 UWB interference 288

9.61 Impairments in 64QAM when there is no coding, and perfect channel knowledge 290

9.62 Impairments in 64QAM when there is coding and perfect channel knowledge 290

9.63 Impairments in 64QAM when there is no coding and soft decision decoding 291

9.64 MAC frame format © IEEE 295

9.65 MAC architecture © IEEE 296

9.66 EDCA mechanism 297

9.67 Reference implementation model © IEEE 298

9.68 CAP/CFP/CP periods © IEEE 300
9.69 Direct Link handshake © IEEE 302
9.70 HIPERLAN/2 MAC protocol stack © ETSI-BRAN . 304
9.71 Mapping between logical channels and transport channels © ETSI-BRAN 307

10.1 Received signal in an SFN 318
10.2 Block diagram of a DAB transmitter 320
10.3 DAB frame structure 322
10.4 DVB transmitter structure 326
10.5 DVB receiver architecture 326

11.1 FDMA, TDMA and CDMA schemes 328
11.2 OFDM-CDMA system with frequency spreading . . 333
11.3 OFDM-CDMA system with time spreading 334
11.4 Discrete time MC-SS RAKE in frequency domain . . 335
11.5 Simplified RAKE receiver 335
11.6 Orthogonal Frequency Division Multiple Access System338
11.7 An example of channel gain 339
11.8 Flash-OFDM tone allocation 344
11.9 Flash-OFDM session control 345
11.10A SDMA system 346

12.1 UWB spectrum 349
12.2 UWB performance 350

12.3 Impulse radio receiver 351

12.4 An example of time-frequency interleaving for the multi band OFDM in three band. © IEEE 353

12.5 Example transmitter architecture for the proposed UWB PHY © IEEE 354

13.1 Global wireless standards 358

13.2 WirelessMAN OFDM transmitter © IEEE 359

14.1 DWMT transmitter 368

14.2 DWMT receiver 369

14.3 IF sampling and digital radio down conversion 370

List of Tables

6.1 Simulation parameters for channel estimation 131

9.1 Comparison of LANs 206
9.2 FCC rules for IEEE 802.11 FHSS 248
9.3 FCC rules for IEEE 802.11 DSSS 250
9.4 IEEE 802.11a OFDM PHY characteristics 255
9.5 Eight PHY Modes of the IEEE 802.11a 259
9.6 Key parameters of the IEEE 802.11a 263
9.7 Priority access category mappings for IEEE 802.11e 294

11.1 Flash-OFDM parameters 344

13.1 OFDM symbol parameters for WirelessMAN OFDM 360
13.2 Data rates achieved in 802.16 with a 1/32 guard time (Mbps) 360

14.1 Relative advantages of single carrier and multi-carrier modulation. An "$\surd$" denotes the system with better performance or lower cost. 364

Index

A

ACK, 215, 223, 226
Adaptive, 2, 138, 186, 201, 245, 254, 291, 329
ADSL(Asymmetric Digital Subscriber Line), 13, 79, 148, 189, 191, 200–202, 363, 365
Aliasing, 39, 44, 94
Alphabet, 2, 28, 171
AODV, 208
ARF (AutoRate Fallback), 291
ARMA model, 142, 145, 146, 151, 160, 163, 164
Association, 212
ATIM, 242
Authentication, 243
Auto-correlation, 55, 141, 145, 332

B

Bandwidth regrowth, 94, 262
Bit Allocation, 37, 201
Bluetooth, 219, 253
Bridged tap, 192
BSS, 209, 212
– IBSS, 206, 207, 240–242, 293
– Infrastructure BSS, 217, 241
Bussgang's theorem, 77

C

CDMA (Code Division Multiple Access), 328, 331, 334
– CDMA2000, 331
– cdmaOne, 331
Channel estimation, 33, 117–120, 122, 124–127, 130, 132–135, 273
Clipping, xvi, 69, 70, 72, 75, 77, 78, 80, 82, 86, 88, 90, 93, 94, 96, 262, 275, 276, 289, 365
Coding
– block, 168, 172, 179
– concatenated, 29, 168, 179, 324
– convolutional, 29, 168, 173, 174, 178–181, 183, 185, 258, 313

– Reed-Solomon, 171, 179, 180, 201, 202, 370
– turbo, 29, 168, 185, 187
Coherence bandwidth, 19, 49
Coherence distance, 21
Coherence time, 21, 49
Conjugate appending, 201
CRC (Cyclic Redundancy Check), 172, 173, 201, 202, 250
Crosstalk, 167, 194, 197, 199
CSMA/CA, 204, 242, 248
Cyclic extension, 44, 45, 103, 156

D

DAB (Digital Audio Broadcast), 33, 317, 320, 323, 324
DCF, 217, 220, 227, 299
DFE (Decision Feedback Equalization), 64, 65, 138, 143
DIF (Decimation in frequency), 42
Distribution system, 303
DIT (Decimation in time), 42
DMT (Discrete Multi-Tone), 12, 13, 143, 151, 190, 367
Doppler, 19, 21, 23, 119, 121, 133, 134, 280, 323
DSDV, 207
DSR, 208
DSSS, 245, 250, 251, 253
– HR/DSSS, 252

E

EAP, 243, 244
Echo, 138, 153–155, 325
Echo cancellation, 153, 156, 201
Entropy, 28
Equalization, xvi, 7, 48, 117, 137, 138, 142, 149, 151, 154, 273, 324, 365, 369
Error probability, 58, 60, 61, 63, 66, 78, 87, 90, 97, 167, 170, 181, 184, 365
ESSID, 208

F

FDMA (Frequency Division Multiple Access), 329, 330
FFT (Fast Fourier Transform), 12, 37, 41, 44, 104, 152, 255, 337, 338, 348, 352, 353, 365
– butterfly, 255, 274, 370
– partial, 41, 42, 44
FHSS, 245, 248, 249, 251
Flash-OFDM, 343
Fragmentation, 219
Frame synchronization, 272
Frequency offset, 44, 99, 102–104, 108, 269, 280
Frequency selective fading, 22, 120, 249, 331

G

Gap, 64
Gating, 7, 52
Gram-Schmidt, 161
Gray coding, 57, 86
GSM, 330

Guard interval, 44, 262, 304, 318, 319, 354

H
Hard or Soft Decision Decoding, 279, 286
HIPERLAN
– HIPERLAN/1, 204, 302
– HIPERLAN/2, 204, 303, 309, 311

I
ICI (Inter Carrier Interference), 102, 121, 134, 278
IEEE, 203, 328
– 802.11, 203, 204, 206, 208, 211, 212, 215, 241, 244, 328
– – DIFS, 220, 222, 229
– – EIFS, 220, 223
– – PIFS, 220, 235
– – SIFS, 219, 224, 226, 235, 299
– 802.11a, 253, 254, 258, 260, 272, 276, 286, 303, 310, 311
– 802.11b, 204, 245, 252, 253, 312, 313
– 802.11c, 312
– 802.11d, 313
– 802.11e, 204, 293, 294, 301, 316
– – AIFS, 295, 297
– – EDCA, 293, 295, 296
– – TXOP, 295, 296, 299
– 802.11f, 313
– 802.11g, 313
– 802.11h, 314
– 802.11i, 244, 315
– 802.16, 336, 357, 359
– 802.1x, 243, 244, 315, 316
IFFT, 154, 260, 274, 275
Interference, 7, 18, 44, 288, 289, 312, 314, 319, 323, 325, 367
– inter-carrier, 100
– inter-channel, 102, 104, 137, 367, 368
– inter-symbol, 2, 3, 7, 44, 99, 102, 121, 137, 139, 142, 155, 254, 319, 336, 367, 368
– Narrow-band, 275, 288, 354, 367, 369
– RF, 253
Interleaving, 13, 172, 178, 185, 187, 259, 321, 354
IPv4, 210
IPv6, 211
IR, 252
ISI (Inter Symbol Interference), 22, 88, 122, 137, 138, 262, 329, 337

J
Jitter, 99, 102, 115, 255

K
Kineplex, 7

M
Markov Model, 227
Matched filter, 47, 108, 142, 254
Mesh network, 357

MIMO, 134
MMAC, 204, 312
Mobile IP, 210, 211
Multiband OFDM, 354

N
NAV, 221, 226, 237
Noise, 70, 85, 99, 109, 112, 121, 139
– Phase, 108, 111, 114, 115, 275
Nyquist criterion, 2, 9, 27, 82

O
OFDM, xv, xvii, 7, 10, 11, 13, 14, 30, 33, 35, 44, 51, 59, 60, 69, 86, 94, 100, 115, 117, 328, 332, 352
– MIMO-OFDM, 119, 134, 135
– OFDM-CDMA, 332
– OFDM-FDMA, 329
– OFDM-SDMA, 346, 348
– OFDM-TDMA, 330
– OFDMA, 336, 337, 348
– TFI-OFDM, 353, 354
Open system authentication, 212

P
PAM, 1, 3, 4, 56
PCF, 217, 237
Peak-to-average ratio, 71, 94, 262, 365, 366, 370
Phase-locked loop, 110, 112
Pilot arrangement, block-type, 122, 124, 133
Pilot arrangement, comb-type, 126
Point-to-multipoint (PMP), 342, 357
Point-to-point (P2P), 359
Power amplifiers, 71, 371
Power management, 241, 242
Puncturing, 174, 259, 304, 311, 321

Q
QAM, 4, 9, 32, 56, 59, 69, 181
– 16QAM, 133, 285
QoS, 293, 299, 303
QSTA, 295, 299

R
RAKE, 332, 334, 351, 352
Rate Adaptation, 291
Rayleigh fading, 23, 35, 80, 96, 133, 318
Ricean fading, 25
Roll-off, 3, 101, 103
RTS/CTS, 215, 216, 233

S
Scattering function, 16, 49
Security, 206, 244, 315
SFN (Single Frequency Network), 13, 318, 325
Shadowing, 17, 269, 318
Shaping gain, 59
Shared key authentication, 212
SSID, 240
SSR, 208
Synchronization, xvii, 90, 99, 104, 107, 186, 239, 269, 321, 325, 331

T
TDMA, 328, 330
Throughput, 209, 231, 249, 330
Time-varying channels, 35, 48, 49, 51
TORA, 208

U
UWB, 288, 289, 349, 350
– Impulse radio, 289, 350, 351

V
Viterbi algorithm, 175, 176, 178, 183, 258, 370

W
Water pouring, 61
Weiner, 112
WEP, 243, 315
WiMAX, 358, 362
Windowing, 44, 103, 367
Wire-Pair channel, xvii, 60, 152, 190–192, 199, 200
WLAN (Wireless Local Area Network), 203, 210, 253, 293
WPAN, 355
WRP, 207
WTRP (Wireless Token Ring Protocol), 328

Z
ZIF (Zero IF technology), 292

Information Technology: Transmission, Processing and Storage
Series Editor: Jack Keil Wolf

Multi-Carrier Digital Communications, 2nd Edition, is a concise summary of the present state-of-the-art of the theory and practice of OFDM technology. It is unique in that it addresses both the theory and practical applications of multi-carrier communications over a wide variety of wired and wireless media. It offers a unified presentation of OFDM theory and high-speed and wireless applications; in particular, ADSL, wireless LAN, digital broadcasting, and ultra wideband technologies are explained.

Multi-Carrier Digital Communications is essential both to developers of communications systems and to researchers and graduate students involved in analysis of digital communications, and will remain a valuable summary of the technology, providing an understanding of new advances as well as the present core technology.

Ahmad R. S. Bahai is a Fellow and the Chief Technologist of National Semiconductor. He is also an adjunct/consulting professor at Stanford University and the University of California at Berkeley. He previously worked at AT&T Bell Laboratories as a technical manager in the Advanced Communications Lab. He served as an editor of IEEE Communication Letters for four years. His research interest includes mixed signal processing, VLSI for communication systems, and communication theory. He is the patent holder and author of many fundamental works in OFDM and wireless communications. He received his Ph.D. degree from the University of California at Berkeley in 1993.

Burton R. Saltzberg received his Sc.D. from New York University in 1964. Dr. Saltzberg is currently an independent consultant in digital communications. For most of his career at Bell Laboratories, from 1957 through 1996, he was engaged in communication-theory research, as well as in analysis and initiation of new data communications offerings over a wide variety of channels. He currently holds 29 patents, including one of the very first for OFDM. He was made a Fellow of the IEEE in 1976 for his contributions to data communications. Dr. Saltzberg received the IEEE Communications Society's Armstrong Award in 1991.

Mustafa Ergen received his B.S. degree from Middle East Technical University as a METU Valedictorian in 2000, his M.S. degree from the University of California Berkeley in 2002, and is currently a Ph.D. student in the University of California Berkeley, all in Electrical Engineering. He has been conducting research in wireless communication networks with an emphasis on WLAN and OFDM systems and is author of many works in the field.

Cover illustration by Mustafa Ergen



ISBN 0-387-22575-7

9 780387 225753

springer.com