WAVELET PACKET DIVISION MULTIPLEXING

By

JIANGFENG WU

M.Sc., B.Sc.,

Beijing University of Aeronautics & Astronautics

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Ph. D.

McMaster University

WAVELET PACKET DIVISION MULTIPLEXING

PH. D. (1998) MCMASTER UNIVERSITY

(Electrical and Computer Engineering) Hamilton, Ontario

TITLE: **Wavelet Packet Division Multiplexing**

AUTHOR: Jiangfeng Wu

M.Sc., B.Sc.,

Beijing University of Aeronautics & Astronautics

SUPERVISOR(S): Dr. K. M. Wong

Professor,

Department of Electrical and Computer Engineering

NUMBER OF PAGES: xiv, 151

ABSTRACT

Wavelet Packet Division Multiplexing (WPDM) is a multiple signal transmission technique in which the message signals are waveform coded on wavelet packet basis functions for transmission. It has the advantage of a parallel transmission system, but the overlapping nature of such waveforms in time and frequency provides a bandwidth efficiency improvement over the Frequency Division Multiplexing (FDM), whilst their orthogonality properties ensure that the overlapping message signals can be separated by a signal correlation receiver.

The interference caused by timing errors in the receiver is examined, and modeled as the information signal sequence filtered by interference filters. A design procedure which exploits the inherent degrees of freedom in the WPDM structure to mitigate the effects of timing error is introduced and a waveform which minimizes the energy of the timing error interference is designed. An expression for the probability of error due to the presence of Gaussian noise and timing error for the transmission of binary data is derived. The performance advantages of the designed waveform over standard wavelet packet basis functions are demonstrated by both analytical and simulation methods.

The performance of WPDM in impulsive noise and Gaussian noise is analyzed, simulated, and compared with the performance of a commonly used serial transmission system, Time Division Multiplexing (TDM). The results support the instinct that WPDM as a parallel transmission system provides greater immunity to impulsive noise than TDM.

The performance of WPDM in fading channels is analyzed, and the time domain equalization technique - Pilot Symbol Assisted Modulation (PSAM) - is applied to suppress the error floor which is common in digital communication with coherent reception in fading channels. The derivation and simulation show that in flat fading channels, the error floor in WPDM is suppressed as successfully as FDM with the PSAM, and the frequency of pilot symbols used in WPDM is less than that in FDM to acquire the sufficient channel estimation, thus the WPDM wastes less capacity and energy on pilot symbols than FDM. The analysis of WPDM in two-ray frequency selective fading channels with PSAM is done, and shows that error floor is also suppressed to some extent, but not as striking as for flat fading channels. The reason for this is revealed in the analysis.

The bandwidth efficiency improvement of WPDM, its simple implementation, its greater immunity to impulsive noise, and its good performance in fading channels indicate that WPDM holds considerable promise as a multiple signal transmission technique.

Acknowledgement

The author wishes to express sincere gratitude to her supervisor Dr. K.M. Wong and her co-supervisors Dr. Z. Luo and Dr. Q. Jin for their constant encouragement, continued assistance and expert guidance and supervision; to supervisor committee members Dr. J.P. Reilly and Dr. P. Yip for their encouragement, supervision and support, throughout the course of this work; to the Communication Research Laboratory (CRL) and staff, for providing good facilities.

Many thanks go to other friends for their valuable help. I would also like to acknowledge the financial support provided by Telecommunication Research Institute of Ontario and Department of Electrical and Computer Engineering (McMaster University).

Finally, I am indebted to my husband, Hai Qi, for his encouragement, assistance and support.

Contents

List of Figures

List of Tables

Acronyms

AWGN	Additive White Gaussian Noise
BPF	Bandpass Filter
BPSK	Binary Phase Shift Keying
CDM	Code Division Multiplexing
DFT	Discrete Fourier Transform
DSB	Double Side Band
FDM	Frequency Division Multiplexing
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
HP	High Pass
ISI	Intersymbol Interference
OFDM	Orthogonal Frequency Division Multiplexing
PAM	Pulse Amplitude Modulation
PSAM	Pilot Symbol Assisted Modulation
TDM	Time Division Multiplexing
WPDM	Wavelet Packet Division Multiplexing

Notations

$\sum_n$	Summation over n
$\prod_n$	Product over n
$\oplus$	Orthogonal sum
$\int_y$	Integration with respect to y
$\|\|A\|\|_2$	$\mathcal{L}_2$ norm of A
A^*	Conjugate of A
$\langle A, B\rangle$	The inner product of A and B
$E\{A\}$	Expected value of A
$\mathrm{R}e\{A\}$	Real part of A
$\mathbf{A}$	Vector or matrix
$\mathbf{A}^\dagger$	Hermitian transpose of $\mathbf{A}$

Chapter 1

Introduction

1.1 Wavelet Packet Division Multiplexing

Multiplexing is a commonly used technique in multi-user communication systems [1–3], whereby a number of independent message signals can be combined into a composite signal suitable for transmission over a common channel. In practice, there are two commonly used methods for multiplexing signals: (1) frequency division multiplexing (FDM) and (2) time division multiplexing (TDM). In FDM, the band-limited baseband signals are shifted upwards in frequency by modulating different sinusoidal carriers to occupy non-overlapping bands as shown in Fig. 1.1. On the other hand, in TDM, the baseband signals are sampled and are usually represented in digital codes. The transmission of these samples engages the channel periodically for different non-overlapping time slots as illustrated in Fig. 1.2.

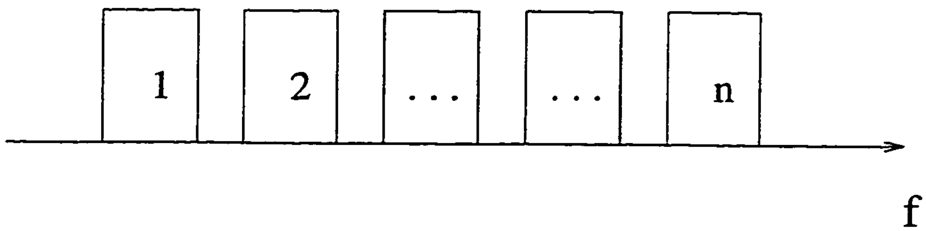


Figure 1.1: FDM: nonoverlapping frequency bands.

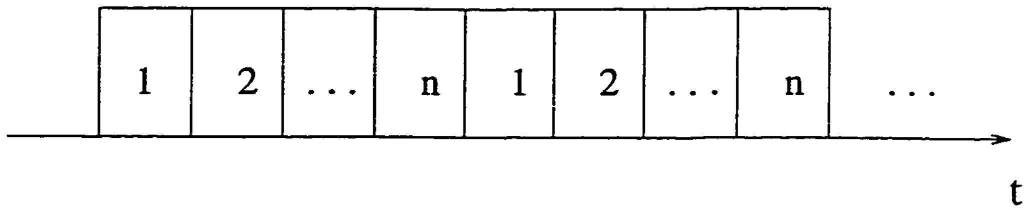


Figure 1.2: TDM: nonoverlapping time slots.

The non-overlapping frequency band in FDM or the non-overlapping time slot in TDM is usually called a *subchannel.*

To facilitate separation of the subchannels at the receiver, specific orthogonality constraints are imposed in multiplexing. Let $\psi_i(t)$ be the modulating signal of ith subchannel, which is assumed to be in the space of finite energy signals, $\mathcal{L}_2(\mathcal{R})$. The orthogonality constraint is

$$\langle \psi_i(t), \psi_j(t-kT) \rangle = \delta[i-j]\delta[k], \tag{1.1}$$

where $\langle \psi_i(t), \psi_j(t) \rangle$ denotes the $\mathcal{L}_2$ inner product of the two functions $\psi_i(t)$ and $\psi_j(t)$ defined as

$$\langle \psi_i(t), \psi_j(t) \rangle = \int_{-\infty}^{\infty} \psi_i(t)\psi_j(t)\, dt,$$

and $\delta[k]$ is the Kronecker delta defined by

$$\delta[k] = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0. \end{cases}$$

In FDM and TDM, the orthogonality of different subchannels is obtained by non-overlapping frequency bands and non-overlapping time slots respectively.

FDM was the earliest technique used to enable multiple simultaneous voice conversations to be transmitted over a common channel, and originally was used exclusively in communication systems. Gradually, FDM has been replaced by TDM. This evolution to TDM was based upon the advantages associated with digital signaling in comparison to analog signaling, such as reliability, voice and data integration capability, ease of encryption, etc.

Nevertheless, FDM can also be used to transmit digital signals. In such cases, FDM is favored over TDM for the following reasons. Firstly, in contrast with TDM in which each data symbol occupies the entire available bandwidth, an individual data element in FDM occupies only a small part of the bandwidth. Therefore, FDM is less sensitive to wide band impulse noise interference, and less sensitive to channel distortion [4]. Secondly, FDM is essentially spreading out the total signal interval, thereby reducing the sensitivity to delay spread. Thirdly, in TDM, certain types of noise, such as narrow-band interference or strong attenuation in some frequency band, may cause the entire system to be disabled. This can be avoided in FDM by dropping out only the subchannels affected.

The advantages of FDM over TDM are, in fact, the advantages of a parallel transmission system over a serial transmission system. FDM is one of parallel transmission systems (in which the signals of different subchannels are transmitted in parallel), and TDM is a specific serial transmission system (in which the signals of different subchannels are transmitted in sequence).

However, nonoverlapping spectra of the different channels in FDM results in inefficient use of the bandwidth. Sufficient guard band is needed between adjacent subchannels to isolate them at the receiver using conventional filters. When much more efficient use of bandwidth is required, one of the popular multiplexing is *orthogonal FDM* (OFDM), which was first proposed in 1960's [5, 6], and is now attracting widespread interest from both researcher and product developer. OFDM abandoned the use of steep bandpass filters that completely separated the spectra of individual subcarriers in FDM. Instead, in OFDM the spectra of subcarriers are permitted to overlap as in Fig. 1.3, with the orthogonality in waveforms.

That is, for each subchannel, the waveform $\psi_i(t)$ to modulate one symbol has to be orthogonal to its translated version in order to eliminate intersymbol interference (ISI)

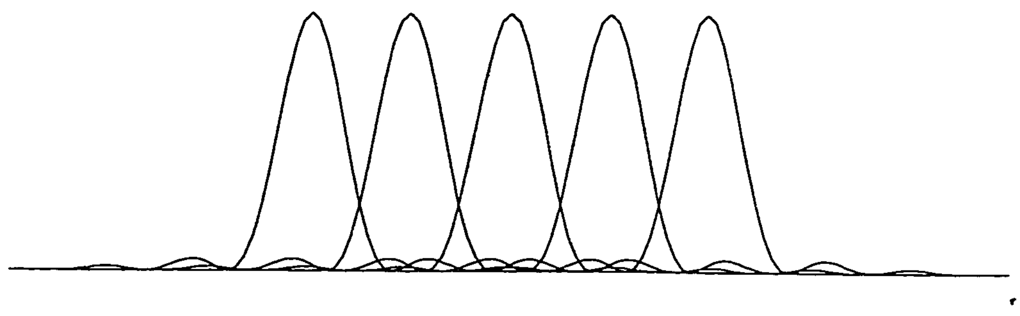


Figure 1.3: OFDM: overlapping frequency bands.

and different subchannels remain mutually orthogonal to eliminate interchannel interference. Design of these functions $\psi_i(t)$ overlapping in both time and frequency domains was discussed in [6].

With each subchannel of OFDM modulating a subcarrier, for a large number of subchannels, the array of sinusoidal generator and coherent demodulator would become unreasonably expensive and complex. However, it has been shown that OFDM can be effectively realized by discrete Fourier transform [4, 7], which has a fast algorithm and can be easily implemented by using a digital signal processor. It is the digital realization that makes OFDM attractive, and proposed as part of the standards in High-speed Digital Subscriber Line (HDSL) and Asymmetric Digital Subscriber Line (ADSL) [8]. Now OFDM is also considered to be employed in mobile radio applications and wireless multimedia communications [9]. It is well known that mobile communication channels are not reliable transmission media since they are time varying and experience various fading. Depending on the environment, some sub-frequency bands may lie in the frequency null (i.e., the attenuation in these bands is extremely high) or suffer severe interference from other transmitters [10]. OFDM has the flexibility of allotting the given signal power to different subchannels according to the channel condition to maximize the channel capacity. The bandwidth efficiency advantage, together with the simple implementation, and the flexibility in assignment of signal powers, make OFDM attractive in wireless communications.

Looking at the wavelet transform developed in late 1980's, we find that a set of basis functions from a wavelet packet matches the waveform requirement in OFDM. These basis functions are orthogonal with one another and are orthogonal across translations. Using these functions in a multiplexing system guarantees efficient use of the spectrum because of the overlap spectra of these basis functions. Moreover, these functions are essentially generated from one function, therefore the problem encountered in OFDM due to inaccurate subcarrier frequency and phase never occur in the multiplexing based on wavelet packet. Following the convention in [11–13], the multiplexing based on wavelet packet is called *wavelet packet division multiplexing* (WPDM).

Furthermore, using directly the properties of wavelet transform, WPDM elegantly corresponds to a digital realization and can be implemented in two steps. First, data streams of each subchannel are digitally multiplexed by the discrete wavelet packet reconstruction filter, which is a finite impulse response (FIR) filter bank. The second step is the waveform coding of the reconstructed sequence. The wavelet packet demultiplexer is just the inverse of the above procedures; i.e., demodulate the waveform coding, then decompose the demodulated sequence through the discrete wavelet packet decomposition filter. Again, the filter is in fact an FIR filter bank.

It is the interesting properties of wavelet transform and OFDM that motivated us to propose the WPDM. On the other hand, WPDM and OFDM have a fundamental difference that motivated us to investigate WPDM in more details.

1.2 Major contributions of thesis

By exploring the various properties of wavelet transform, applications in several areas have been found, such as image processing [14], acoustic signal processing [15], communications [13, 16–24], etc. As one of the applications in communications, multiplexing based on wavelet packet is proposed and studied in the thesis.

The basis functions from a wavelet packet may be orthogonal and each function may be orthogonal to its translated versions, although these functions overlap in both time and frequency domains. This property satisfies the waveform requirement of OFDM exactly. Therefore, WPDM is proposed as an alternative bandwidth efficient OFDM. Wherever the OFDM is used, the WPDM can be used with more bandwidth efficiency and less complexity, which will be addressed in more detail.

One difference between OFDM and WPDM is that for OFDM, each subchannel has the same spectral shape, while this does not hold for WPDM. Thus, although WPDM and OFDM should behave similarly, exactly identical performance cannot be expected. The performance of OFDM in various noise environments has been extensively researched [7, 25] since it was first proposed in 1960's. For WPDM, performance analysis and simulation are also needed. This constitutes another main topic of the dissertation beside the system model [11–13, 26–28]. The following environments are focused on: 1. timing error and jitters existing between transmitter and coherent receiver; 2. impulsive noise channels; and 3. fading channels.

First, the proposed multiple signal transmission and reception system WPDM exploits the self and mutual orthogonality of a set of basis functions so that the message signals can be separated by coherent detection. Signal transmission schemes based on coherent detection are vulnerable to errors caused by timing discrepancy, so we analyze the effects of timing errors and jitters [11–13]. The interference due to a timing discrepancy is made up of two components, *intersymbol interference* (ISI) and *interchannel interference* (ICI) or crosstalk, which we model as the message signal from the given channel and the adjacent channels passing through interference filters. The simple model leads us further to propose an optimum design of wavelet, which minimizes the timing error interference. We also obtain expressions for the probability of error under the effects of timing error and additive white Gaussian noise (AWGN). Corresponding simulations are employed to confirm the validity of our interference model.

Second, impulsive noise may be a primary source of system performance degradation in some cases, such as in the transmission of data over telephone networks. Intuitively, multiplexing in frequency domain performs better than multiplexing in time domain in such environment. Here, we derive an expression for the probability of error for WPDM in the presence of both impulsive and Gaussian noise [26, 27]. The derivation extends previous work on the performance of digital communication systems when impulsive noise and Gaussian noise are the causes of error [29–32] to WPDM. For the purpose of comparison, the performance of a TDM system in the same channel environment is also evaluated. The calculations and corresponding simulation results illustrate that WPDM can provide greater immunity to impulsive noise than TDM. This result is intuitively expected, since, one noise burst, which can destroy one bit in TDM, will be spread over all the channels of a WPDM system. In each channel, the spread effects of the impulsive noise may not be strong enough to cause an error.

As in OFDM, besides bandwidth efficiency, in WPDM the limited transmission power can be allotted to different subchannels according to the channel characteristics to maximize the overall transmission rate. This property makes WPDM attractive in mobile communications, in which communication channel varies from time to time and channel gains in different subchannels may have a big difference. In general, a mobile channel is characterized by fading, which may be flat or frequency selective. Usually, flat fading causes an error floor (or irreducible bit error rate), in digital communication systems, including WPDM. That means that the system performance cannot be improved by simply increasing signal power. A time-domain equalization technique Pilot Symbol Assisted Modulation (PSAM) is applied in WPDM to suppress the error-floor. The theoretical analysis and simulation results show that the PSAM helps WPDM suppress the error-floor successfully. The performance comparison between WPDM and FDM illustrate that WPDM wastes less energy and bandwidth in pilot symbols than conventional FDM [28]. The similar PSAM is also applied in two-ray frequency selective fading channels. The performance analysis and simulation demonstrate that when the time delay spread is small relative to the bit duration. e.g. less than or equal to 20%, WPDM performs well, although the error floor is not completely suppressed.

In frequency selective fading channels, since some of the subchannels may lie in the frequency bands which have small channel gains, a high bit error rate may be expected from these subchannels. However, in WPDM the errors do not spread over other subchannels. Thus, we may either allot little power on the subchannels with low channel gains or simply discard them. These discarded subchannels may be used by other users which may be in another location and may not have small channel gains. In TDM, since one bit occupies the whole available bandwidth, no such flexibility is available to enhance system performance.

1.3 Outline of thesis

We have explained the motivation of proposing WPDM as an bandwidth efficient multiplexing scheme in the present chapter. The concepts of the scheme will be developed. First, wavelet transform, wavelet packet, and their properties relevant to our application are reviewed in Chapter 2. In Chapter 3, system model of WPDM based on the orthogonal basis of wavelet packet and its equivalent discrete realization are described. The bandwidth efficiency issue is also discussed. Since wavelet functions often look irregular, concerns on how to recover the synchronizing signals may be raised. Hence, one of many possible schemes on symbol synchronization is briefly addressed in Chapter 4. Followed immediately is the discussion of the interference due to timing error. The mathematical model is proposed to characterize the interference. Based on the interference model, the optimum wavelet design criterion is proposed and the probability of error is analyzed. The theoretical analysis and simulation of WPDM in the presence of both impulsive noise and Gaussian noise are presented in Chapter 5. Moreover, a quantitative comparison between WPDM and TDM is given. Chapter 6 focuses on the performance of WPDM in fading channels. In flat fading channel, the performance of WPDM using PSAM is analyzed and simulated, and WPDM is compared with FDM. In two-ray selective fading channels, analysis and simulation of WPDM with PSAM is also presented. Finally, Chapter 7 concludes the thesis with a summary of what has been done, and future work is suggested to promote WPDM as a practically competitive bandwidth efficient multiplexing scheme.

Chapter 2

Wavelet and Wavelet Packet

The theory of wavelets and wavelet packets is by now a well-developed subject and can be found in a number of textbooks [33–40] and tutorial articles [41–46]. However, in order to keep the thesis self-contained, we briefly summarize the concepts of wavelets and the properties related to our application. Since multiresolution analysis provides a natural framework for the understanding of wavelet bases, we will start from multiresolution analysis.

A multiresolution analysis (MRA) [14,33] consists of a collection of embedded subspaces $V_{\ell 1}$ [1]

$$\cdots \subset V_{21} \subset V_{11} \subset V_{01} \subset V_{-11} \subset \cdots \tag{2.1}$$

in the space of finite energy signals, $\mathcal{L}_2(\mathcal{R})$, with properties:

$$\bigcap\nolimits_{\ell \in \mathcal{Z}} V_{\ell 1} = \{0\},$$

$$\bigcup\nolimits_{\ell \in \mathcal{Z}} V_{\ell 1} = \mathcal{L}_2(\mathcal{R}),$$

and

$$s(t) \in V_{\ell 1} \Longleftrightarrow s(2t) \in V_{\ell-1,1},$$

[1] Whilst the double subscript notation may seem redundant at this point, it simplifies the extension to wavelet packets.

where $\mathcal{Z}$ represents the set of integers, and $s(t)$ is a $\mathcal{L}_2$-function. Each subspace $V_{\ell 1}$ has an orthonormal basis $\{\phi_{\ell 1}(t-nT_\ell)\}_{n\in\mathcal{Z}}$, [2] generated by translations of a single function $\phi_{\ell 1}(t)$. Furthermore, the basis functions for different subspaces can be obtained from each other by dyadic dilation; i.e.,

$$\phi_{\ell 1}(t) = 2^{\ell/2}\phi_{01}(2^{-\ell}t), \tag{2.2}$$

and $T_\ell = 2^\ell T_0$. Since $V_{\ell 1} \subset V_{\ell-1,1}$, we can write $\phi_{\ell 1}(t)$ as a linear combination of the orthonormal basis functions of $V_{\ell-1,1}$; i.e.,[3]

$$\phi_{\ell 1}(t) = \sum_n h[n]\phi_{\ell-1,1}(t-nT_{\ell-1}), \tag{2.3}$$

where

$$h[n] = \langle \phi_{\ell-1,1}(t-nT_{\ell-1}), \phi_{\ell 1}(t)\rangle$$

and $\langle\cdot,\cdot\rangle$ denotes the $\mathcal{L}_2$ inner product.[4] A cursory examination of Eqs (2.2) and (2.3) reveals the close relation between $h[n]$ and $\phi_{\ell 1}(t)$; i.e.,

$$\phi_{\ell 1}(t) = \sqrt{2}\sum_n h[n]\phi_{\ell 1}(2t-nT_\ell). \tag{2.4}$$

Integration of both sides of Eq. (2.4) with respect to t

$$\int_{-\infty}^{\infty}\phi_{\ell 1}(t)dt = \sqrt{2}\sum_n h[n]\int_{-\infty}^{\infty}\phi_{\ell 1}(2t-nT_\ell)dt$$

yields

$$\sum_n h[n] = \sqrt{2},$$

which ensures uniqueness of $\phi_{\ell 1}(t)$.

The construction of $\phi_{\ell 1}(t)$ from $h[n]$ can be carried out using several algorithms [39,41, 48], one of which is the "cascade" algorithm in which the approximations are successively

[2] The translation factor for the space V_{01}, T_0, is set equal to one in most developments of MRA, but it will represent a fundamental symbol period in our application so we leave it explicit.

[3] We follow the signal processing convention to use square brackets for the argument of a discrete-time sequence [47].

[4] Due to the dilation isometry (Eq. 2.2), $h[n]$ is independent of ℓ.

refined by dyadic contraction and (tapped-delay line) filtering with the filter coefficients $h[n]$, starting from a unit energy "box" function of duration T_ℓ. That is,

$$\phi_{\ell 1}^{(j)}(t) = \sqrt{2} \sum_n h[n] \phi_{\ell 1}^{(j-1)}(2t - nT_\ell), \tag{2.5}$$

where $\phi_{\ell 1}^{(j)}$ is the jth approximation of the $\phi_{\ell 1}(t)$ and

$$\phi_{\ell 1}^{(0)}(t) = \begin{cases} 1/\sqrt{T_\ell} & 0 \leq t < T_\ell, \\ 0 & \text{otherwise.} \end{cases}$$

We terminate the iteration of the approximation when $\left\| \phi_{\ell 1}^{(j)}(t) - \phi_{\ell 1}^{(j-1)}(t) \right\|_2 \leq \epsilon_\phi$, where $\| \cdot \|_2$ denotes the $\mathcal{L}_2$ norm and ϵ_ϕ is a pre-set accuracy measure. The above procedure generates an approximation of the $\phi_{\ell 1}(t)$.

Indeed, in order for the set of functions $\{\phi_{\ell 1}(t - nT_\ell)\}_{n \in \mathcal{Z}}$ to form an orthonormal basis for the space $V_{\ell 1}$, $h[n]$ must satisfy the orthonormality constraint

$$\sum_n h[n] h[n - 2m] = \delta[m], \tag{2.6}$$

where $\delta[m]$ is the Kronecker delta [34,48].

Given such a sequence $h[n]$, we can find a sequence $g[n]$ satisfying

$$\sum_n g[n] h[n - 2m] = 0$$

such that the function

$$\phi_{\ell 2}(t) = \sum_n g[n] \phi_{\ell-1,1}(t - nT_{\ell-1}) \tag{2.7}$$

forms an orthonormal basis for the orthogonal complement of $V_{\ell 1}$ in $V_{\ell-1,1}$, which we will denote by $V_{\ell 2}$; i.e.,

$$V_{\ell-1,1} = V_{\ell 1} \bigoplus V_{\ell 2},$$

and

$$\langle \phi_{\ell 1}(t), \phi_{\ell 2}(t) \rangle = 0.$$

The sequence $g[n]$ can be chosen to be a reversed, modulated and shifted version of $h[n]$,

$$g[n] = (-1)^n h[1-n], \tag{2.8}$$

where $n \in \mathcal{Z}$. The resulting basis functions $\phi_{\ell 1}(t)$ and $\phi_{\ell 2}(t)$ are called the *scaling function* and *wavelet at scale ℓ*, respectively. To interpret Eqs (2.3) and (2.7) in the familiar frequency domain, we note that $\phi_{\ell 1}(t)$ occupies only half the bandwidth of $\phi_{\ell-1,1}(t)$, due to their dilation relation. Thus Eqs (2.3) and (2.7) can be conceived of as a low-pass (LP) and a high-pass (HP) discrete-time filtering (i.e., tapped delay line filtering [2, 49]) of the signal $\phi_{\ell-1,1}(t)$ to obtain $\phi_{\ell 1}(t)$ and $\phi_{\ell 2}(t)$, respectively. If the sequences $h[n]$ and $g[n]$ in a particular MRA are finite, they can be realized as FIR filters forming an orthonormal two-channel perfect reconstruction filter bank [34, 48]. This connection is the key to many applications of MRA [14].

In an analogous way to Eqs (2.3) and (2.7), we can recursively partition the spaces $V_{\ell m}$ ($\ell \geq 0$, $1 \leq m \leq 2^\ell$) as

$$V_{\ell,m} = V_{\ell+1,2m-1} \bigoplus V_{\ell+1,2m},$$

with the corresponding orthonormal bases $\{\phi_{\ell+1,m}(t - nT_\ell)\}_{n \in \mathcal{Z}}$ obtained in the following tree-structured manner [17]:

$$\phi_{\ell+1,2m-1}(t) = \sum_n h[n]\phi_{\ell m}(t - nT_\ell), \tag{2.9}$$

$$\phi_{\ell+1,2m}(t) = \sum_n g[n]\phi_{\ell m}(t - nT_\ell). \tag{2.10}$$

Here, the first subscript denotes the "level" in the tree structure induced by the recursive decomposition of V_{01} and the second subscript denotes the position of a node in a given level. We can grow or prune the tree in any desired fashion, and the functions $\phi_{\ell m}(t)$ at the "leafs" or *terminals* of a given tree structure provide a set of "wavelet packet basis functions" or simply a *wavelet packet*, as exemplified in Fig. 2.1. The corresponding *subband* structure in the frequency domain is shown in Fig. 2.2.

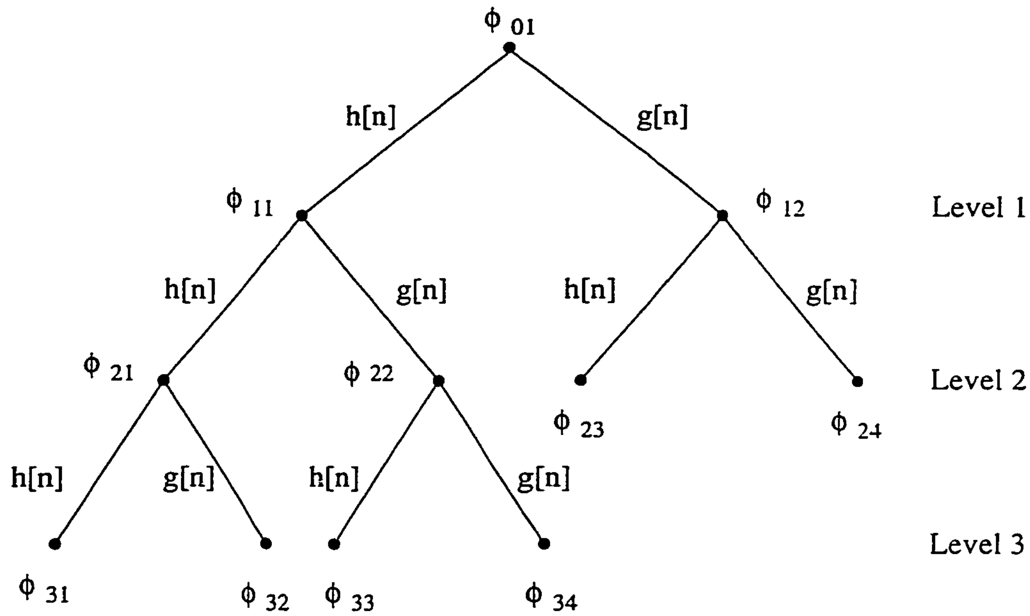


Figure 2.1: A typical wavelet packet tree structure.

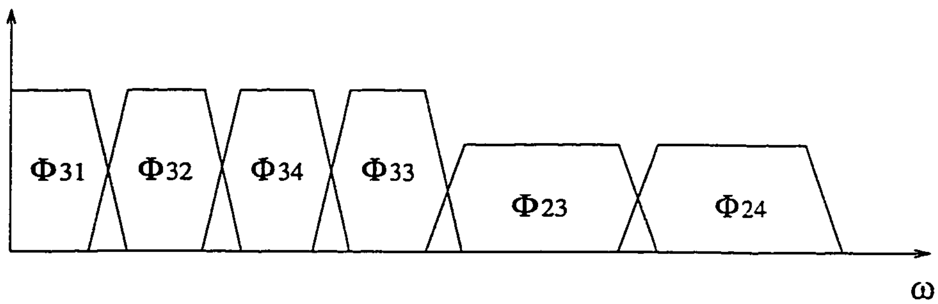


Figure 2.2: The subband structure of the tree, for representative filters.

It is worthwhile to summarize the property of such a wavelet packet: The orthonormal wavelet packet basis functions overlap in both time and frequency domains; each function has a duration several times longer than T_ℓ, which provides the wavelet packet with a potential of having a good frequency concentration, and therefore is a desirable property in communications.

Since $V_{\ell+1,2m}$ is the orthogonal complement of $V_{\ell+1,2m-1}$ in $V_{\ell m}$, we can write the function $\phi_{\ell m}(t)$ as a linear combination of translated versions of $\phi_{\ell+1,2m-1}(t)$ and $\phi_{\ell+1,2m}(t)$. Using Eqs (2.9) and (2.10), the coefficients of the linear combination can be shown to be reversed versions of the decomposition sequences $h[n]$ and $g[n]$ [14]; i.e.,

$$\phi_{\ell m}(t-nT_\ell) = \sum_k h[n-2k]\phi_{\ell+1,2m-1}(t-kT_{\ell+1}) + \sum_k g[n-2k]\phi_{\ell+1,2m}(t-kT_{\ell+1}). \quad (2.11)$$

Continuing this process, we can reconstruct $\phi_{01}(t)$ from the terminal functions of an arbitrary tree-structured decomposition:

$$\phi_{01}(t-nT_0) = \sum_{\ell\in\mathcal{L},\, m\in\mathcal{M}_\ell} \sum_k f_{\ell m}[n-2^\ell k]\phi_{\ell m}(t-kT_\ell), \quad (2.12)$$

where $\mathcal{L}$ denotes the set of levels containing the terminals of a given tree, $\mathcal{M}_\ell$ the set of indices of the terminals at the ℓth level, and $f_{\ell m}[n]$ is the equivalent sequence (filter) built from the combinations of $h[n]$, $g[n]$ and dilation (down-sampling) which lead from the root to the (ℓ, m)th terminal; i.e.,

$$\phi_{\ell m}(t-nT_\ell) = \sum_k f_{\ell m}[k-2^\ell n]\phi_{01}(t-kT_0). \quad (2.13)$$

For a given tree structure, the functions $\phi_{\ell m}(t)$ in Eq. (2.13) will be called the *constituent functions of* $\phi_{01}(t)$.

In addition to providing a framework for understanding and constructing wavelet packet bases, multiresolution analysis leads naturally to a fast scheme for the computation of the wavelet packet coefficients of a given function.

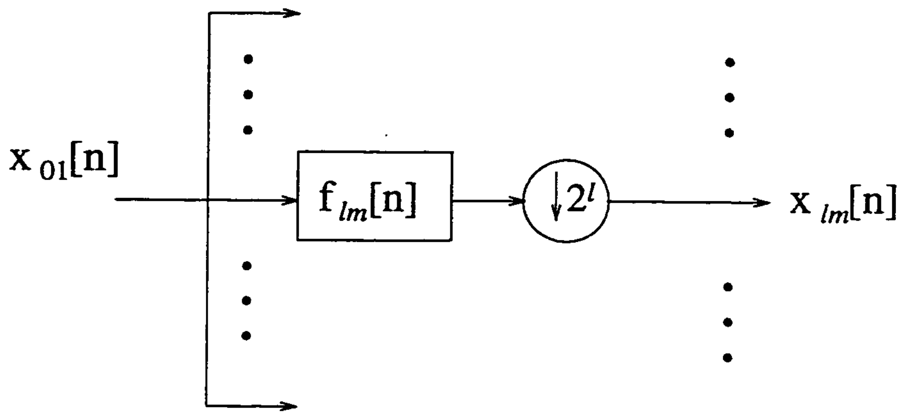


Figure 2.3: Schematic representation of wavelet packet decompostion.

Since $\{\phi_{\ell m}(t - nT_\ell)\}_{n\in\mathcal{Z}}$ is an orthonormal bases of the space $V_{\ell m}$, the projection of an $\mathcal{L}_2$-function $s(t)$ on $V_{\ell m}$ is $\sum_k x_{\ell m}[k]\phi_{\ell m}(t - kT_\ell)$, where $x_{\ell m}[k] = \langle s(t), \phi_{\ell m}(t - kT_\ell)\rangle$. Suppose that we have computed, or are given $x_{01}[n]$. By taking the inner product with $s(t)$ in both sides of Eq. (2.13), it follows that the wavelet coefficients at the terminal (ℓm) $(\ell > 0)$ can be computed from

$$x_{\ell m}[n] = \sum_k f_{\ell m}[k - 2^\ell n]x_{01}[k]. \tag{2.14}$$

That is, the $x_{\ell m}[n]$ are obtained by convolving the sequence $x_{01}[n]$ with $f_{\ell m}[-n]$ and then retaining only samples at $2^\ell n$. Eq. (2.14) is called the *wavelet packet decomposition* formula.

On the other hand, suppose the coefficients $x_{\ell m}[n]$ are given. Computing the inner products of $s(t)$ with both sides of Eq. (2.12) results in

$$x_{01}[n] = \sum_{\ell\in\mathcal{L},\, m\in\mathcal{M}_\ell} \sum_k f_{\ell m}[n - 2^\ell k]x_{\ell m}[k], \tag{2.15}$$

which is called the *wavelet packet reconstruction* formula.

Schematical representation of wavelet packet decomposition Eq. (2.14) and wavelet packet reconstruction Eq. (2.15) are presented in Fig. 2.3 and Fig. 2.4, respectively.

If $h[n]$ is chosen to be finite length, both wavelet packet decomposition and reconstruction are implemented by a bank of FIR (finite impulse response) filters, and there are no interconnections among the filters.

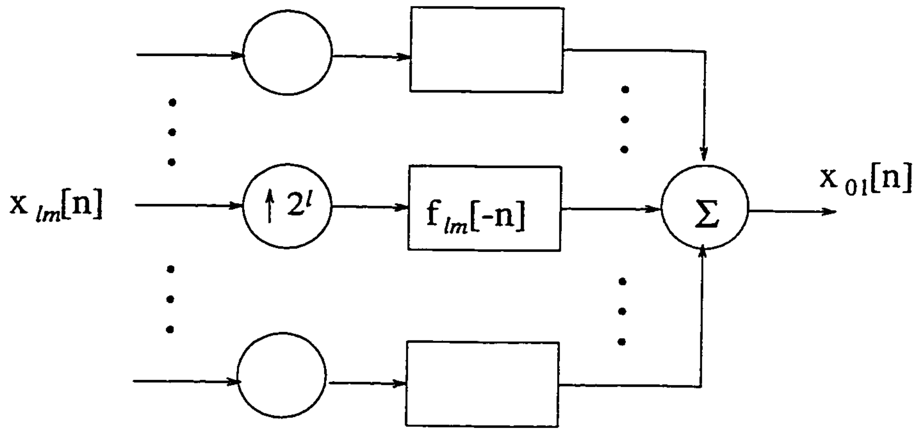


Figure 2.4: Schematic representation of wavelet packet reconstruction.

Chapter 3

Wavelet Packet Division Multiplexing

In this chapter, we will present how the orthogonal wavelet packet bases are applied in multiplexing and key features of such a multiplexing scheme. The analogies and differences between multiplexing based on the wavelet packet and other major multiplexing schemes are discussed. As one of the main advantages, the bandwidth efficiency of the new multiplexing system is compared with that of conventional FDM.

3.1 Application of wavelet packets in multiple signal transmission

In our application of wavelet packets to multiple signal transmission, we consider a TDM system in which there are $K_{\ell m}$ independent message signals interlaced with one another. Each message signal can be either binary or multilevel, independent of the number of levels in other signals. In the thesis, we treat the case of binary signaling in each message signals. Thus, in between two consecutive binary symbols of the same message there are $K_{\ell m} - 1$

other binary symbols, one from each of the other message signals. The combined sequence forms a composite sequence of binary symbols $x_{\ell m}[n]$ such that $x_{\ell m}[n] = \pm 1$. The system that we propose here seeks the representation of the binary symbols "1" and "−1" by $\phi_{\ell m}(t)$ and $-\phi_{\ell m}(t)$, respectively, so that the waveform-coded composite TDM sequence is given by

$$s_{\ell m}(t) = \sqrt{E_0} \sum_n x_{\ell m}[n] \phi_{\ell m}(t - nT_\ell). \tag{3.1}$$

Here E_0 is the energy in one symbol. The optimal receiver for $s_{\ell m}(t)$ in the additive white Gaussian noise (AWGN) channel is the well-known correlator followed by a sampler [1,3,50, 51]. In practice, the scaling function $\phi_{01}(t)$ will usually be chosen to have finite duration, so that the constituent functions $\phi_{\ell m}(t)$ are also of finite duration (and the filters $h[n]$ and $g[n]$ are FIR filters). Note, however, that the duration of $\phi_{\ell m}(t)$ is often larger than the bit interval, T_ℓ. The orthogonality of the set $\{\phi_{\ell m}(t - nT_\ell)\}_{n \in Z}$ ensures that such overlap will not cause (inter-symbol) interference.

Since all the constituent functions in a given tree structure are orthogonal to each other, we may employ all of these constituent functions to carry binary data from different TDM groups of users, whilst maintaining the simple correlator structure of the receiver. The configuration of this multiple signal transmission system is shown in the left half of Fig. 3.1, in which the tree structure has been chosen to have M terminals. At the (ℓ, m)th terminal (the mth node of the ℓth level), the group of binary messages $\{r_1, \ldots, r_{K_{\ell m}}\}$ is time multiplexed by a commutator, and the resulting binary sequence $x_{\ell m}[n]$ is waveform coded as in Eq. (3.1). Obviously, binary sequences waveform coded by constituent functions at different levels will have different bit intervals (or, equivalently, different bit rates), while sequences at the same level have equal bit intervals. At the output of the transmitter, we

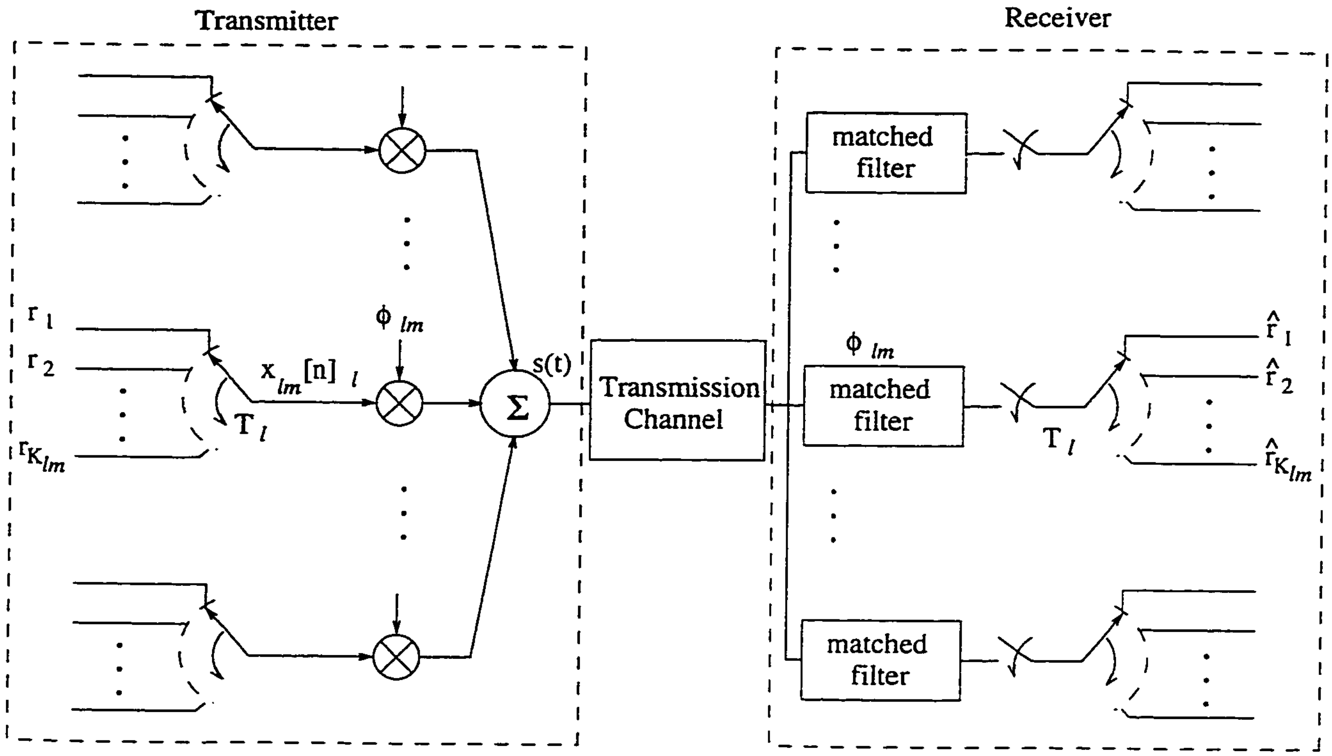


Figure 3.1: WPDM system model.

have a composite signal given by

$$s(t) = \sqrt{E_0} \sum_{\ell \in \mathcal{L},\, m \in \mathcal{M}_\ell} \sum_n x_{\ell m}[n] \phi_{\ell m}(t - nT_\ell). \tag{3.2}$$

The reception of such a composite signal can be carried out using a bank of correlating receivers, as shown on the right half of Fig. 3.1.

However, by using the efficient wavelet packet reconstruction formula Eq. (2.15), it can be seen that

$$s(t) = \sqrt{E_0} \sum_k x_{01}[k] \phi_{01}(t - kT_0), \tag{3.3}$$

where $x_{01}[k]$ is the equivalent sequence at the root of the tree,

$$x_{01}[k] = \sum_{\ell \in \mathcal{L},\, m \in \mathcal{M}} \sum_n f_{\ell m}[k - 2^\ell n] x_{\ell m}[n], \tag{3.4}$$

from which the original message signals can be recovered using wavelet packet decomposition

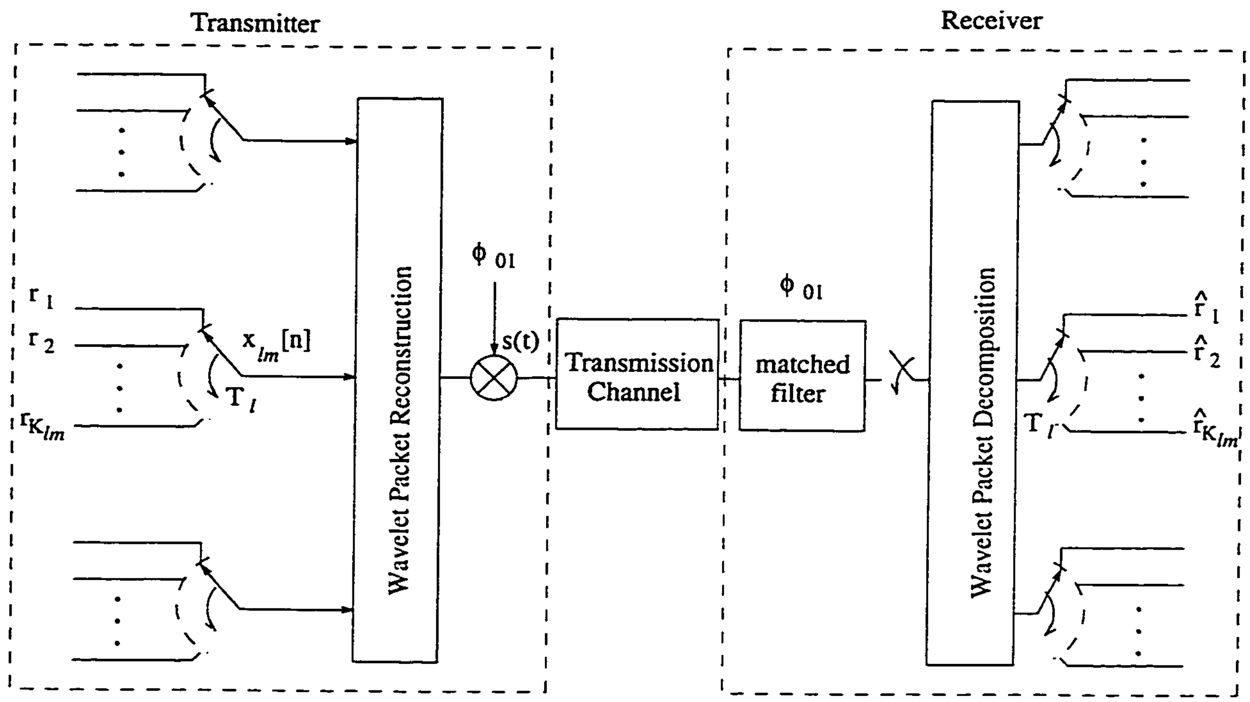


Figure 3.2: Equivalent transmission and reception systems for WPDM.

formula

$$x_{\ell m}[n] = \sum_k f_{\ell m}[k - 2^\ell n] x_{01}[k]. \tag{3.5}$$

Eqs (3.3) and (3.4) suggest an alternative realization of the WPDM system in which the bank of modulators in the transmitter is replaced by a multi-rate filter bank built from filters with impulse responses $f_{\ell m}[-n]$ (or appropriately shifted versions thereof), and a single modulator.

The receiver can also be alternatively realized, as suggested in Eqs (3.3) and (3.5), by a single correlator followed by a bank of discrete-time multi-rate filters with impulse responses $f_{\ell m}[n]$. This alternative realization of the receiver simply implements Mallat's algorithm for the calculation of a wavelet series [14], and the alternative realization of the transmitter implements the inverse algorithm. This equivalent configuration of the system is shown in Fig. 3.2.

The multiple signal transmission and reception system in Figs 3.1 and 3.2 is the so-

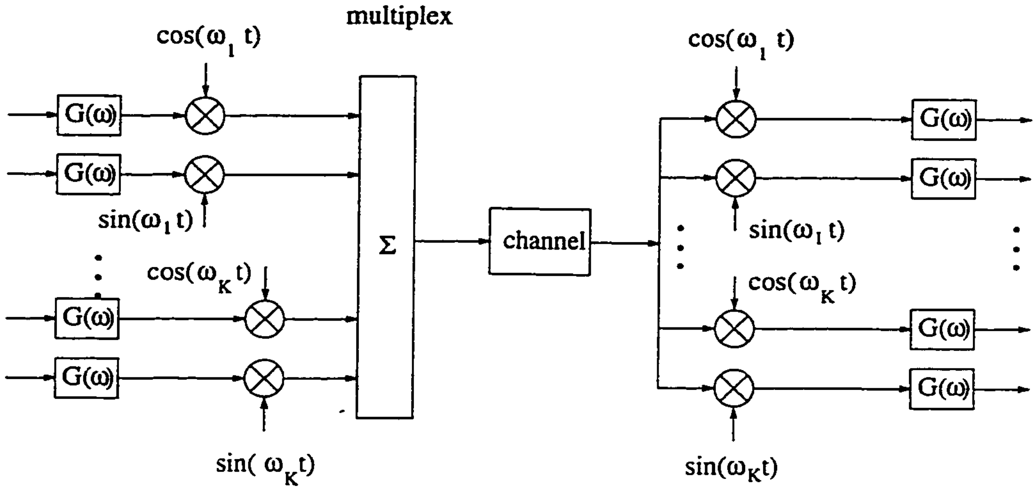


Figure 3.3: A basic OFDM system.

called wavelet packet division multiplexing (WPDM) in the thesis. It is noted that the transmitted signals in both realizations are identical, and hence so is the performance of the systems in the AWGN channel. It is also noted that since the WPDM system transmits orthogonal antipodal waveforms to represent the binary data, its performance in the AWGN channel is identical to any other transmission system which uses such waveforms.

3.2 Analogies and differences between WPDM and OFDM

OFDM was proposed to make fuller use of the available bandwidth by allowing subchannels overlapping, while maintaining the advantages of FDM systems such as combating impulsive noise and avoiding equalization. Since the subchannels are overlapped, time-domain waveform has to be carefully designed to satisfy the orthogonality constraint stated in Eq. (1.1). One of the OFDM system block diagrams is shown in Fig. 3.3 [4], where $G(\omega)$ was carefully designed to eliminate ISI and interchannel interference. As an example, $g_0(t)$, the time-domain function of $G(\omega)$, is chosen to be a rectangular function; i.e.,

$$g_0(t) = \begin{cases} 1 & 0 \leq t \leq T \\ 0 & \text{else} \end{cases}$$

where $1/T$ is the transmission rate of one subchannel. The functions

$$g_i(t) = g_0(t) \cos\left(\frac{2\pi}{T} it\right), i = 1, 2, \ldots, N_c - 1 \tag{3.6}$$

are orthogonal with each other, although they all lie in the duration $[0, T]$ and their spectrum is overlapped with subcarrier separation $2\pi/T$. Because of the window effect of $g_0(t)$, each of the subchannels takes a shape of $\frac{sin(x)}{x}$ in the frequency domain. With the N_c functions as the modulating waveforms for N_c subchannels, N_c bits are transmitted in duration $[kT, (k+1)T], k \in \mathcal{Z}$, each from one subchannel. In digital implementation, since the DFT or inverse DFT is performed on the N_c bits, the N_c bits are usually called a *block* [52]. The OFDM signal is, in fact, a block followed by another block, and there is no overlap between the adjacent blocks.

It has been seen that orthonormal wavelet packet basis functions may overlap in both time and frequency domains, and satisfy the constraints in Eq. (1.1) exactly. So they are natural choices to be used in bandwidth efficient multiplexing system.

In OFDM, in order to avoid the arrays of sinusoidal generators and coherent demodulators, digital implementation of OFDM was proposed. Fig. 3.4 shows one of the digital models, where DFT (or FFT) is used. Through MRA, the digital implementation model of WPDM was also derived as system model II.

In spite of the similarity between WPDM and OFDM, there are major differences between the two bandwidth efficient schemes. First, the spectrum of each subchannel in a WPDM system has a different shape, while all the subchannels of an OFDM system have the same spectral shape. The spectral comparison is illustrated in Fig. 3.5. More importantly,

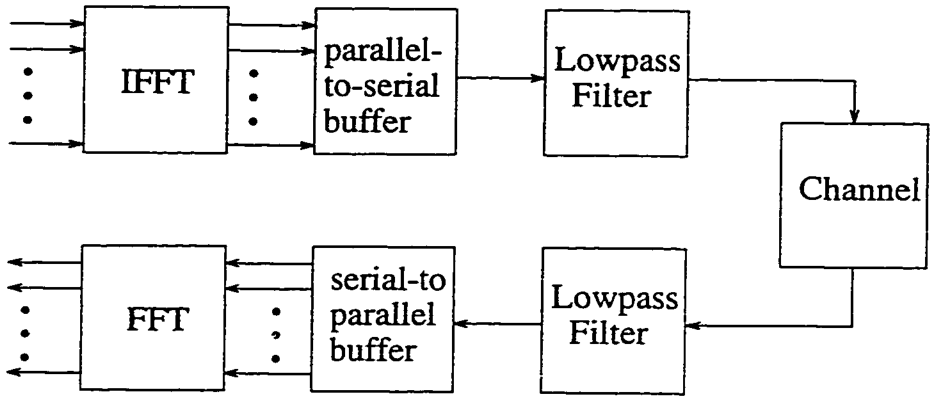
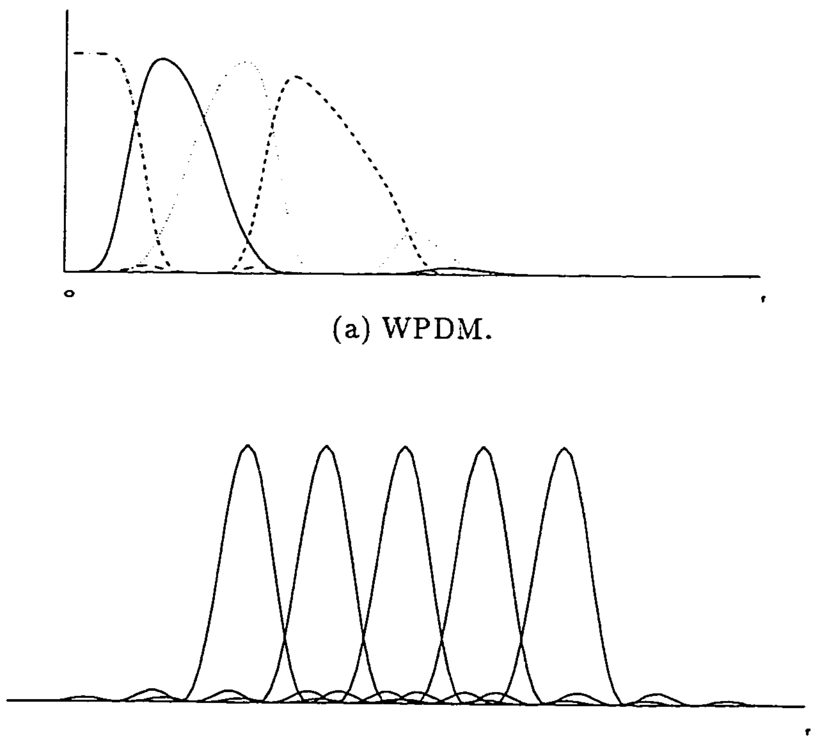


Figure 3.4: A digital implementation of OFDM



(a) WPDM.

(b) OFDM.

Figure 3.5: Spectral shape comparison between WPDM and OFDM.

WPDM is more bandwidth efficient than OFDM. When the OFDM signal passes through a bandlimited channel, the waveforms representing one block extend beyond the duration T and thereby causes interference between adjacent blocks. Usually, a rectangular function, such as $g_0(t)$, has a strong sidelobe, thus the interference due to the bandlimiting is severe. In order to alleviate the interference, a guard period is inserted between the blocks [7,52]. On the other hand, in WPDM, the duration of a wavelet packet basis function is not limited to $[0, T]$, and is usually much longer than T. Thus the waveforms can have such a good frequency concentration that a bandlimited channel will essentially not cause interference and no guard time is necessary. It is the guard time in OFDM that makes OFDM less bandwidth efficient than WPDM.

In addition, digital implementations of WPDM and OFDM have different structures. A WPDM system is implemented by a bank of FIR filters, and there are no interconnections among the filters of each subchannel. Hence, a simple and real-time implementation can be realized. The implementation of OFDM, however, requires inverse DFT and DFT transforms. Although there are fast algorithms for DFT and inverse DFT, the implementation of the transforms is more complex than a bank of FIR filters (no interconnection). Thus WPDM has a simpler implementation than OFDM.

3.3 Features of WPDM

– Bandwidth efficiency

As discussed in the last section, even compared with the bandwidth efficient OFDM, WPDM is more efficient in the bandwidth usage, since no guard time is necessary.

Since the constituent functions used in WPDM can overlap with one another in both time and frequency domains, we can expect the multiplexing is more efficient in bandwidth usage than conventional FDM systems. We will give a detailed discussion on this in Section

3.4.

– Generalized form of TDM, FDM and CDM

We can view the WPDM system as a combined form of TDM and FDM. If each user in a WPDM system is assigned a constituent function $\phi_{\ell m}(t)$ to carry the message data, then the WPDM system is reduced to a wavelet carrier (WC) FDM system. On the other hand, if only $\phi_{01}(t)$ is employed to carry the time-multiplexed binary data from different users, then the WPDM system is reduced to a WC TDM system. It should be emphasised that the WC TDM/WC FDM systems arrived at by taking the extreme cases of WPDM are different from conventional TDM/FDM systems in the sense that a conventional TDM/FDM system does not have overlapping time slots/frequency bands. However, in the special WC TDM/WC FDM systems, overlapping time slots/frequency bands are allowed. The signals from different users are separated by the orthogonality of the waveforms. The allowance of overlapping frequency bands results in an increase in the number of users being able to share the channel compared to an FDM system.

We can also interpret the WPDM scheme as a member of a class of generalized orthogonal code division multiplexing (CDM) schemes in which the codes are the equivalent filters $f_{\ell m}[k]$ and the chip waveform is $\phi_{01}(t)$. This generalized class includes the conventional orthogonal CDM schemes, but extends those schemes to allow for real-valued orthogonal codes which overlap in time, and orthogonal chip waveforms which have a duration longer than the chip interval.

– Dynamic assignment of channels

It has been seen that in one particular partition of V_{01}, the functions $\phi_{\ell m}(t)$ at different levels ℓ carry information bits at different rates, and therefore can be assigned to users to satisfy different transmission rate requirement. This useful property turns out to be a complicated solution if the system model I is employed and every waveform of $\phi_{\ell m}(t)$ has to be generated explicitly. However, by using the equivalent digital model, only discrete filter coefficients need to be changed to employ $\phi_{\ell m}(t)$ at different levels.

– Immunity to impulsive noise

It has been explained that WPDM is a generalized form of TDM and FDM. When WPDM is used as WC FDM, it has the advantages of an FDM system. Greater immunity to impulsive noise is one of the advantages. This property will be discussed in more detail in Chapter 5.

3.4 Bandwidth efficiency of WPDM

Similar to OFDM, WPDM is proposed as a bandwidth efficient multiplexing scheme. Indeed, this is one of the original motivations for WPDM. In order to appreciate the advantage and substantiate our intuition, we compare the required bandwidth of the WPDM system to that of the FDM system. In deriving the following results, we assume that the transmitter filters and data processors operate in baseband, since in practice signal shaping and data processing are usually performed in baseband [5].

We consider a multi-user communication system in which there are 2^ℓ user messages, each having a bit interval of T_ℓ seconds. In an FDM system, each message modulates a different sinusoidal carrier and requires a bandwidth of $(1+\alpha)/T_\ell$ Hz, where α is the "roll–off" factor of the frequency characteristic of the pulse used to represent the signal. If a pulse with an ideal LP characteristic (having a bandwidth from $-\frac{1}{2T_\ell}$ to $\frac{1}{2T_\ell}$ Hz) is used, then $\alpha = 0$. In practice, pulses with a raised cosine characteristic, for which $0 < \alpha < 1$, are generally used [1,3,50,51]. In addition, to avoid interference from the adjacent subchannels due to subcarrier frequency drifting, a guard band of bandwidth δ_g/T_ℓ is inserted between the frequency bands of the message signals. Thus, the total bandwidth required from the transmission channel is

$$W_F = \frac{2^\ell(1+\alpha+\delta_g)}{T_\ell} \text{ Hz.}$$

Now, consider the same 2^ℓ users employing the WPDM system in which each message modulates a finite duration constituent function from the ℓth level of the tree structure in Fig. 3.2. Although the total bandwidth of the constituent functions is theoretically infinite, we can define a γ-*effective bandwidth* of $\phi_{01}(t)$ to be β_γ/T_0 such that

$$\int_{-2\pi\beta_\gamma/T_0}^{2\pi\beta_\gamma/T_0} |\Phi_{01}(\omega)|^2 \, d\omega \; = \; \gamma \int_{-\infty}^{\infty} |\Phi_{01}(\omega)|^2 \, d\omega, \tag{3.7}$$

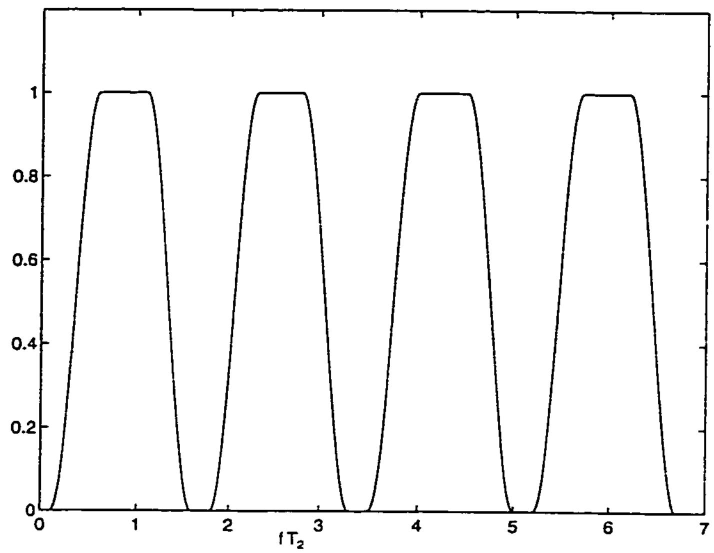
where $\Phi_{01}(\omega)$ is the Fourier transform of $\phi_{01}(t)$, for $0 < \gamma < 1$. It is customary to choose $\gamma = 0.99$ [53,54]. Of course, the γ-effective bandwidth in Eq. (3.7) is also the total γ-effective bandwidth of the 2^ℓ constituent functions. We transmit the 2^ℓ constituent functions in double side-band (DSB). The total γ-effective bandwidth needed to transmit the 2^ℓ messages in the WPDM system is

$$W_W = \frac{\beta_\gamma}{T_0} = \frac{2^\ell \beta_\gamma}{T_\ell} \text{ Hz.}$$

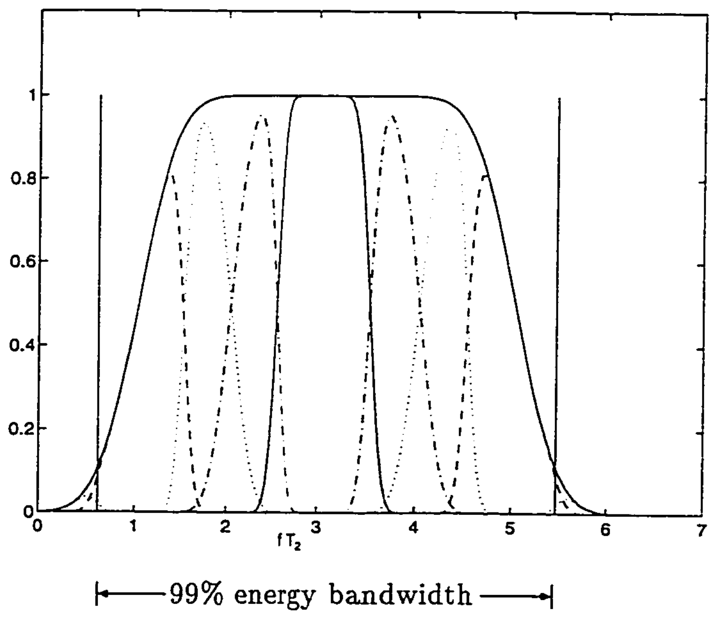
and hence the ratio of the total bandwidths of the two systems is given by

$$\eta = \frac{W_W}{W_F} = \frac{2\beta_\gamma}{1+\alpha+\delta_g}.$$

There are many orthogonal scaling functions for which $\beta_\gamma < 1$, among them, the well-known functions generated by Daubechies filters of length greater than or equal to 6 [34,48]. If we choose one for which the ratio $\eta < 1$, we obtain a bandwidth reduction over the FDM system. Table 3.1 shows the values of β_γ for the Daubechies filters of different lengths, as well as the values of the bandwidth ratio η for the representative values $\alpha = 0.5$ and $\delta_g = 0.2$. It can be observed from the table that if a Daubechies filter of length 14 is used, the WPDM system requires only 70% of the transmission bandwidth required by the corresponding FDM system — a substantial saving. The spectral energy characteristics of



(a) FDM with $\alpha = 0.5$ and $\delta_g = 0.2$.



(b) WPDM.

Figure 3.6: Spectral energy characteristics of a four user system with frequency normalized to the bit rate $1/T_2$.

the two systems in the case of four users are illustrated in Fig. 3.6. Fig. 3.6(a) shows the characteristics of the four user messages in the FDM scheme, and Fig. 3.6(b) shows the characteristics of the four user messages in the WPDM scheme (using Daubechies filters of length 14). The combined frequency characteristic is shown by the solid line and the 99% energy bandwidth is also indicated. In closing this argument, we point out that the nature of the bandwidth saving discussed above is independent of the structure of the WPDM tree.

It is worth of noting that although the Daubechies wavelets employed throughout the thesis illustrate the principle of WPDM system properly, they may not be good wavelets when the bandwidth efficiency is considered. The reason is that the Daubechies wavelets are aimed at giving the maximum regularity, not particularly for frequency concentration. A wavelet with better frequency concentration is given in [20].

N	4	6	8	10	12	14	16	18	20
β_γ	1.62	0.77	0.66	0.63	0.62	0.60	0.60	0.59	0.58
η	1.91	0.91	0.78	0.75	0.72	0.70	0.70	0.69	0.68

Table 3.1: Values of β_γ and η for scaling functions generated by Daubechies filters of length N with $\alpha = 0.5$, $\delta_g = 0.2$ and $\gamma = 0.99$.

Chapter 4

Timing Errors

4.1 Introduction

In digital communications, there is a hierarchy of synchronization problems to be considered [3,50,55]. First, the baseband signal, such as the multiplexed signal in the last chapter, often needs to be modulated on a high frequency carrier for transmission. Thus, there is the problem of *carrier synchronization* which concerns the generation of a reference carrier with a phase matching that of the transmitted signal. This reference carrier is used at the receiver to perform a coherent demodulation operation, recovering the baseband signal. Next comes the problem of synchronizing a receiver clock with the baseband sequence. This is commonly called *symbol synchronization.* Beside these, there are also word and frame synchronizations.

Carrier synchronization is necessary for any coherent reception communication system. Various methods [55–58] have been proposed, and for each method, the effect of synchronization error has been analyzed. These carrier synchronization methods and performance analyses can be directly applied to WPDM if carrier modulation is employed in it. Within the scope of this thesis, we simply do not consider carrier synchronization and always assume that carrier synchronization is perfect.

On the other hand, although the symbol synchronization in baseband exists in any digital communication system, waveforms used in WPDM are often irregular. This may may raise concern on how to carry out synchronization in WPDM. To relieve this concern, we note that WPDM is a particular format of Pulse Amplitude Modulation (PAM), for which timing recovery is a well solved problem. Thus, all the techniques to recover PAM timing signal can be applied in WPDM.

Since the synchronization errors in PAM timing recovery depend on the signaling waveforms, we will analyze the effect of timing errors on WPDM performance. There are usually two kind of synchronization errors. One is so called timing bias, which can be regarded as constant error over a group of symbols. Another one is timing jitter; i.e., timing error is a random variable at different instants. These two kind of errors will be analyzed under one framework in WPDM.

It is well known that synchronization errors cause intersymbol interference (ISI) and interchannel interference (crosstalk), which will be characterized in our analysis by interference filters. From the interference filters, the probability of errors will be derived, and an optimum wavelet design criterion will be proposed. As an example, we design one optimum wavelet to minimize the interference due to synchronization errors.

4.2 Symbol synchronization of WPDM

The model for a baseband PAM signal [55, 57] is

$$s_{PAM}(t) = \sum_k x[k]\phi_{PAM}(t - kT - \tau), \tag{4.1}$$

where $x[n]$ is the message sequence and $\phi_{PAM}(t)$ is the signaling pulse. We want to make an accurate determination of τ, from the operations on $s_{PAM}(t)$, to obtain the best sampling instant of $kT + \tau$.

Recall that the multiplexed signal in WPDM is

$$s(t) = \sqrt{E_0} \sum_k x_{01}[k] \phi_{01}(t - kT_0). \tag{4.2}$$

Assuming there is no channel distortion, the received signal is just the delayed version of $s(t)$; i.e.,

$$r(t) = s(t - \tau) = \sqrt{E_0} \sum_k x_{01}[k] \phi_{01}(t - kT_0 - \tau). \tag{4.3}$$

By comparing Eq. (4.1) and (4.3), it is obvious that WPDM signal has exactly the same expression as the PAM signals. Therefore the synchronization scheme of PAM can be directly applied in the symbol synchronization of WPDM.

In the timing recovery of PAM signals, the property that a PAM signal is a zero-mean cyclostationary process is used by assuming $\{x[n]\}$ is a zero-mean stationary sequence with independent elements. The square of the PAM signal possesses a periodic mean value

$$\begin{aligned} E\{s_{PAM}^2(t)\} &= E\{x^2[n]\} \sum_k \phi_{PAM}^2(t - kT - \tau) \\ &= \frac{E\{x^2[n]\}}{T} \sum_\ell \int_{-\infty}^{\infty} \Phi_{PAM}\left(\frac{\ell}{T} - f\right) \Phi_{PAM}(f) df \, \exp\left(j\frac{2\pi\ell}{T}(t - \tau)\right), \end{aligned} \tag{4.4}$$

where the Poisson Sum Formula [2] is used, $E\{x^2[n]\}$ represents the average power of $x[n]$, and $\Phi_{PAM}(f)$ is the Fourier transform of $\phi_{PAM}(t)$. It is a common practice to choose $|\Phi_{PAM}(f)| = 0$ for $|f| > \frac{1}{T}$ for a practical communication system, and therefore there are only three non-zero terms ($\ell = 0, \ell = \pm 1$) in Eq. (4.4).

This result suggests the use of a timing recovery circuit as shown in Fig. 4.1 with the bandpass filter tuned to the symbol rate $1/T$. Alternate zero crossings of the timing wave $s_{PAM}^2(t)$ are used as indications of the correct sampling instants. Assume that the BPF is a very narrow-band filter tuned at frequency $1/T$. The mean timing wave is sinusoidal with

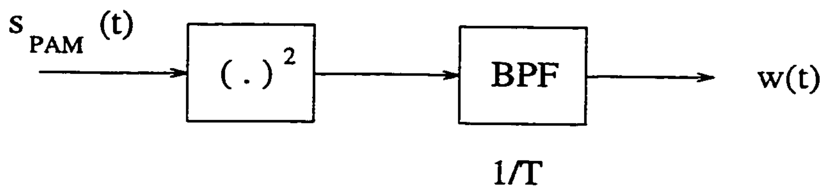


Figure 4.1: A simple timing recovery scheme.

a phase of $-2\pi\tau/T$, for a real $\Phi_{PAM}(f)$,

$$E\{s^2_{PAM}(t)\} = \frac{E\{x^2[n]\}}{T}\mathrm{Re}\left[\int_{-\infty}^{\infty} \Phi_{PAM}\left(\frac{1}{T} - f\right) \Phi_{PAM}(f)df \exp\left(j\frac{2\pi t}{T} - j\frac{2\pi\tau}{T}\right)\right]. \tag{4.5}$$

Let us define

$$C = \int_{-\infty}^{\infty} \Phi_{PAM}\left(\frac{1}{T} - f\right) \Phi_{PAM}(f)df. \tag{4.6}$$

Eq. (4.5) becomes

$$E\{s^2_{PAM}(t)\} = \frac{E\{x^2[n]\}C}{T}\mathrm{Re}\left[\exp\left(j\frac{2\pi t}{T} - j\frac{2\pi\tau}{T}\right)\right].$$

The zero crossings of the mean wave are at a fixed time offset relative to the desired sampling instants. This timing offset can be handled by counting logic in the clock circuitry.

The mean timing wave can be regarded as a kind of discriminator characteristic for measuring the parameter τ. Discrimination is enhanced by increasing the slope at the zero crossing. As this slope is proportional to C in Eq. (4.6), the shape of the pulse $\phi_{PAM}(t)$ affects timing recovery. We see that C depends on the amount of overlap between the functions $\Phi_{PAM}(f)$ and $\Phi_{PAM}(1/T - f)$, hence it depends on the amount by which the bandwidth of $\Phi_{PAM}(f)$ exceeds the $\frac{1}{2T}$ Nyquist bandwidth. With no excess bandwidth, $C = 0$ (which results in $E\{s^2_{PAM}(t)\} = 0$), and therefore this method of timing recovery fails. Although WPDM is a bandwidth efficient system, the function $\phi_{01}(t)$ in Eq. (4.3) has bandwidth in excess of the Nyquist frequency (Fig. 3.5b). Thus, the synchronization schemes based on the above principles can be used for the symbol synchronization of WPDM.

The actual zero crossings of $s^2_{PAM}(t)$ fluctuate about the desired sampling instants because the time wave depends on the actual realization of the entire data sequence and additive channel noise. The fluctuation is the cause of timing errors. In the next Section, we will analyze the effect on WPDM performance of zero crossing fluctuation or timing error.

4.3 Interference in transmission

The proposed multiple signal transmission and reception system (Figs 3.1 and 3.2) exploits the self- and mutual-orthogonality of the constituent functions so that the message bits can be separated by a synchronous correlator. Signal transmission schemes based on synchronous correlators are vulnerable to errors caused by timing discrepancy. In this section, without considering how timing errors are caused, we will analyze the effects of timing errors. Moreover, at early stages of analysis, channel additive white Gaussian noise (AWGN) is ignored. In the analysis, we make the assumption that the bandwidth of the communication channel is sufficiently large so as not to cause any significant distortion on the waveform of the transmitted signal.

Since the realizations in Figs 3.1 and 3.2 are equivalent, we may use either of them to analyze the performance of the WPDM system. Here we analyze the realization shown in Fig. 3.2, for which the transmitted signal $s(t)$ is given by Eq. (3.3). Let Δ_n denote the timing discrepancy at the nth symbol between the transmitter and the receiver. In an attempt to recover the transmitted signal, the output of the correlator in the receiver is given by

$$\begin{aligned}\hat{x}_{01}[n] &= \frac{1}{\sqrt{E_0}} \int s(t)\phi_{01}(t - nT_0 + \Delta_n)\, dt \\ &= \sum_k x_{01}[k] \int \phi_{01}(t - kT_0)\phi_{01}(t - nT_0 + \Delta_n)\, dt\end{aligned}$$

$$= \sum_k x_{01}[k] R_\phi(kT_0 - nT_0 + \Delta_n), \tag{4.7}$$

where $1/\sqrt{E_0}$ is introduced as a normalization factor, and $R_\phi(\tau)$ is the autocorrelation function of $\phi_{01}(t)$ defined as

$$R_\phi(\tau) = \int \phi_{01}(t)\phi_{01}(t+\tau)dt. \tag{4.8}$$

We assume that $R_\phi(t)$ is K times differentiable for some $K \geq 1$. Then we can employ Taylor's formula with remainder [59] on Eq. (4.7), giving

$$\hat{x}_{01}[n] = x_{01}[n] + \sum_k x_{01}[k] R'_\phi(kT_0 - nT_0)\Delta_n + O(\Delta_n^2), \tag{4.9}$$

where $R'_\phi(kT_0 - nT_0)$ denotes the derivative of $R_\phi(\tau)$ with respect to τ evaluated at $\tau = kT_0 - nT_0$, and the equality $R_\phi(kT_0 - nT_0) = \delta[k-n]$, $\delta[k]$ being the Kronecker delta, has been used. Ignoring the terms involving second and higher orders of Δ_n, we can regard the second term on the right hand side of Eq. (4.9) as the timing-discrepancy interference denoted by $I_{01}[n]$; i.e.,

$$I_{01}[n] = \sum_k x_{01}[k] R'_\phi(kT_0 - nT_0)\Delta_n. \tag{4.10}$$

In Eq. (4.10), whether Δ_n is a constant error or a timing jitter is not specified. However, constant timing error will have a different effect on system performance from timing jitter. Hence, we discuss these two cases separately.

4.3.1 Constant timing errors

In the case of constant timing error, Δ_n is a constant over n, therefore the subscript n in Δ_n can be omitted, and the interference term becomes

$$I_{01}[n] = \Delta \sum_k x_{01}[k] R'_\phi(kT_0 - nT_0). \tag{4.11}$$

Using the fact that $R_\phi(\tau)$ is an even function, we can rewrite this interference term as

$$I_{01}[n] \;=\; \Delta \sum_k \left(x_{01}[n+k] - x_{01}[n-k]\right) R'_\phi(kT_0). \tag{4.12}$$

After the correlator, the recovery of the estimates of the message signals, $\hat{x}_{\ell m}[n]$, requires the passing of $\hat{x}_{01}[n]$ through the appropriate path of the wavelet packet decomposer given in Eq. (3.5). The interference at the (ℓ, m)th terminal is simply $I_{01}[n]$ passed through the same path of the decomposer; i.e.,

$$\begin{aligned} I_{\ell m}[n] &= \sum_k f_{\ell m}[k - 2^\ell n] I_{01}[k] &(4.13)\\ &= \Delta \sum_k \sum_i \Big(x_{01}[k+i] - x_{01}[k-i]\Big) R'_\phi(iT_0) f_{\ell m}[k - 2^\ell n]. &(4.14)\end{aligned}$$

Substituting Eq. (3.4) into Eq. (4.14), we have

$$\begin{aligned} I_{\ell m}[n] &= \Delta \sum_{\lambda \in \mathcal{L},\, \mu \in \mathcal{M}_\lambda} \sum_j x_{\lambda\mu}[j] \sum_i R'_\phi(iT_0) \\ &\quad \times \sum_k f_{\lambda\mu}[k + i - 2^\lambda j] f_{\ell m}[k - 2^\ell n] - f_{\lambda\mu}[k - i - 2^\lambda j] f_{\ell m}[k - 2^\ell n]. \end{aligned}$$

To simplify the notation, we define the cross-correlation sequences

$$C_{\ell m}^{\lambda\mu}[n] = \sum_k f_{\lambda\mu}[k+n] f_{\ell m}[k], \qquad (4.15)$$

and the coefficients

$$J_{\ell m}^{\lambda\mu}[n] = \sum_i R'_\phi(iT_0) \left(C_{\ell m}^{\lambda\mu}[n+i] - C_{\ell m}^{\lambda\mu}[n-i]\right). \qquad (4.16)$$

This allows us to express the interference at the (ℓ, m)th terminal as

$$\begin{aligned} I_{\ell m}[n] &= \Delta \sum_{\lambda \in \mathcal{L},\, \mu \in \mathcal{M}_\lambda} \sum_j x_{\lambda\mu}[j] \sum_i R'_\phi(iT_0) \left(C_{\ell m}^{\lambda\mu}[i - 2^\lambda j + 2^\ell n] - C_{\ell m}^{\lambda\mu}[2^\ell n - i - 2^\lambda j]\right) \\ &= \Delta \sum_{\lambda \in \mathcal{L},\, \mu \in \mathcal{M}_\lambda} \sum_j x_{\lambda\mu}[j] J_{\ell m}^{\lambda\mu}[2^\ell n - 2^\lambda j]. &(4.17)\end{aligned}$$

We can interpret Eq. (4.17) as follows: The interference at the (ℓ, m)th terminal is the sum of multi-rate filtered versions of the constituent sequences $x_{\lambda\mu}[n]$ from each of the terminals. The filtering process consists of up-sampling by 2^λ, filtering with impulse response $J_{\ell m}^{\lambda\mu}[n]$, and then down-sampling by 2^ℓ. We will therefore refer to a set of coefficients $J_{\ell m}^{\lambda\mu}[n]$ as an interference filter. As expected, the interference due to timing discrepancy arises not only from message signals using the same constituent function (intersymbol interference – ISI), but also from message signals using other constituent functions as waveforms (cross-talk). If the user message bits are independent, then using Eq. (4.17), the variance of the interference at the (ℓ, m)th terminal can be shown to be directly proportional to the sum of $\sum_n \left|J_{\ell m}^{\lambda\mu}[n]\right|^2$ over all relevant λ and μ.

An important observation is that the total energy of the interference at the terminals is independent of the tree structure. We formalize this as the following proposition:

Proposition 4.3.1 *The total energy of the interference at the terminals of a WPDM system due to timing discrepancy is independent of the tree structure; i.e.,*

$$\sum_n |I_{01}[n]|^2 = \sum_{\ell \in \mathcal{L},\, m \in \mathcal{M}_\ell} \sum_n |I_{\ell m}[n]|^2 .$$

This energy balance between the initial and decomposed interferences is a direct result of Parseval's theorem. The wavelet packet decomposition is an orthonormal transformation and hence the norm is preserved.

Proposition 4.3.1 has a far-reaching implication: If our sole objective is to reduce the total energy of the timing-discrepancy interference then, regardless of the tree structure of the WPDM system, we can achieve this by choosing a $\phi_{01}(t)$ so that the energy of $I_{01}[n]$ is minimized. In Section 4.5, we will design an optimal $\phi_{01}(t)$ based on the minimization of the total interference energy. However, it is observed from Eq. (4.13) that the interference at a particular terminal depends on both $\phi_{01}(t)$ and the position of that terminal in the tree. Thus, while the total energy of the interference remains independent of the tree structure, the distribution of the interference energy to different terminals is dependent on the structure.

4.3.2 Timing jitter

Similar to the case of constant timing error, we want to find the interference due to timing jitter to each subchannel. The interference due to timing jitter at the terminal (ℓm), denoted by $I_{\ell m}^{(j)}[n]$ is obtained by passing $I_{01}[n]$ in Eq. (4.10) through the wavelet packet decomposer; i.e.,

$$I_{\ell m}^{(j)}[n] = \sum_u I_{01}[u] f_{\ell m}[u - 2^\ell n].$$

Substituting Eq. (4.10) in the above equation and making use of the wavelet packet reconstruction formula, we have

$$I_{\ell m}^{(j)}[n] = \sum_{\lambda \in \mathcal{L},\, \mu \in \mathcal{M}_\ell} \sum_v x_{\lambda\mu}[v] \sum_u \sum_k f_{\lambda\mu}[k - 2^\lambda v] f_{\ell m}[u - 2^\ell n] R'_\phi(uT_0 - kT_0)\Delta_u. \quad (4.18)$$

For simplicity, denote

$$Y_{\ell m}^{\lambda\mu}[n, v] = \sum_u \sum_k f_{\lambda\mu}[k - 2^\lambda v] f_{\ell m}[u - 2^\ell n] R'_\phi(uT_0 - kT_0)\Delta_u. \quad (4.19)$$

As a function of the random variable Δ_n, $Y_{\ell m}^{\lambda\mu}[n, v]$ is a random process. Substitution of Eq. (4.19) into (4.18) results in a simple expression of $I_{\ell m}^{(j)}[n]$,

$$I_{\ell m}^{(j)}[n] = \sum_{\lambda \in \mathcal{L},\, \mu \in \mathcal{M}_\ell} \sum_v x_{\lambda\mu}[v] Y_{\ell m}^{\lambda\mu}[n, v]. \quad (4.20)$$

From Eq. (4.20), the interference at terminal (ℓ, m) is composed of two parts: ISI and crosstalk, which are $\sum_v x_{\ell m}[v] Y_{\ell m}^{\ell m}[n, v]$ and $\sum_{\lambda\mu} \sum_v x_{\lambda\mu}[v] Y_{\ell m}^{\lambda\mu}[n, v]$ with $(\lambda, \mu) \neq (\ell, m)$, respectively. Comparison of Eq. (4.17) with Eq. (4.20) reveals that the interference under constant timing errors and timing jitters have the similar form of signal sequences $x_{\lambda\mu}[n]$ passing through filters. However, the interference filters in the two cases have different properties. Under constant timing errors, the impulse responses $J_{\ell m}^{\lambda\mu}[n]$ of the interference filters are deterministic, which in turn dictates the different interference energy at each terminal as stated in the Section 4.3.1. On the other hand, in the case of timing jitters, the filter $Y_{\ell m}^{\lambda\mu}[n]$ is a random process. Therefore, even if the patterns of $x_{\lambda\mu}[n]$ are known, the

$I_{\ell m}^{(j)}[n]$ is still a random variable. The statistical properties of $I_{\ell m}^{(j)}[n]$ can be derived from the properties of $x_{\lambda\mu}[n]$ and Δ_n.

By assuming that the timing jitter Δ_n has zero mean and variance $\sigma_{\Delta_n}^2$ (but no specific distribution is assumed), and making use of the fact that $x_{\lambda\mu}[n]$ are independent sequences, $I_{\ell m}^{(j)}[n]$ has zero mean and has variance

$$\sigma_{I_{\ell m}}^2 = \sigma_{\Delta_n}^2 \sum_{\lambda\in\mathcal{L},\,\mu\in\mathcal{M}_\ell} \sum_u \sum_{k_1} \sum_{k_2} f_{\lambda\mu}[k_1-2^\lambda v] f_{\lambda\mu}[k_2-2^\lambda v] f_{\ell m}^2[u] R'_\phi(uT_0-k_1T_0) R'_\phi(uT_0-k_2T_0),$$

which is independent of n. The interference $I_{\ell m}^{(j)}[n]$, therefore, is a wide-sense stationary process. It can also be verified that the same variance of interference occurs at every terminal (ℓ, m). Thus, unlike the case of constant timing errors, the interference under timing jitters is independent of terminal location. Because the wavelet packet decomposition is an orthogonal transform, Proposition 4.3.1 still holds in timing jitter case. Therefore, minimization of the total interference at the root $I_{01}[n]$ results in the minimization of interference at each and every terminal.

4.4 Probability of error

We now turn to the derivation of the probability of error in the reception of the binary messages under the influence of timing error interference. We examine the case where the channel is corrupted by stationary zero-mean additive white Gaussian noise (AWGN) $\nu(t)$, using the equivalent realization of Fig. 3.2.

4.4.1 Probability of error under constant timing error

In the case of constant timing error, since the interference at a particular terminal depends on the position of the terminal in the WPDM tree structure, as seen in Eq. (4.13), the probability of error at each terminal is structure dependent. The binary message signal

$x_{\ell m}[n]$ at the (ℓ, m)th terminal, which takes on one of two possible values, ± 1, is carried across the channel by the constituent function $\phi_{\ell m}(t - nT_\ell)$ and is recovered by the correlator and sampler at the receiver. A consequence of the orthogonality properties of the constituent functions is that the output of the (ℓ, m)th correlator, in the absence of noise and timing error, is simply

$$\begin{aligned}\hat{x}_{\ell m}[n] &= \int \sum_k x_{\ell m}[k] \phi_{\ell m}(t - kT_\ell)\, \phi_{\ell m}(t - nT_\ell)\, dt \\ &= \int x_{\ell m}[n]\, \phi^2_{\ell m}(t - nT_\ell)\, dt = \begin{cases} 1 & \text{if a ``1'' is sent,} \\ -1 & \text{if a ``}-1\text{'' is sent.} \end{cases}\end{aligned}$$

Therefore, assuming equal probabilities of ± 1, the optimal decision threshold is at zero. To incorporate the effects of timing error interference, recall from Eq. (4.17) that the interference at the (ℓ, m)th terminal is the sum of multi-rate filtered versions of the message signals from each terminal. For each instant n, the message bit $x_{\ell m}[n]$ can be modeled as a random variable which takes on the values $+1$ and -1 with equal probabilities. In order to determine the probability distribution of the resulting interference, we must determine how many message bits contribute to the interference. To this end, we note that if the filter $h[n]$ is of length N, the length of the equivalent filter from the root of the tree to the (ℓ, m)th terminal, $f_{\ell m}[n]$, is $N_\ell = (2^\ell - 1)(N - 1) + 1$ [39]. Using Eqs (4.15) and (4.16), and the fact that $R_\phi(\tau)$ is zero outside $\big(-(N-1)T_0, (N-1)T_0\big)$, the interference filter $J^{\lambda\mu}_{\ell m}[n]$ is zero outside an interval of length

$$N^\lambda_\ell = (2^\ell + 2^\lambda)(N - 1) - 1.$$

This means that from the (λ, μ)th terminal to the (ℓ, m)th terminal, there is a linear combination of N^λ_ℓ binary random variables, yielding a total of $2^{(N^\lambda_\ell)}$ discrete values. Another terminal at the level λ will contribute another $2^{(N^\lambda_\ell)}$ possible discrete values. Thus, the total number of possible discrete values of interference at the (ℓ, m)th terminal is given by

$$L_{\ell m} = \prod_{\lambda \in \mathcal{L}} 2^{M_\lambda (N_\ell^\lambda)},$$

where M_λ is the number of terminals at level λ. The magnitudes of these discrete interference values depend on the coefficients of the interference filters (Eq. 4.17), but because $x_{\lambda\mu}[n]$ takes on values of ± 1 with equal probabilities, these discrete interference values are equally likely and are symmetrically distributed about zero. If we denote the total interference due to timing discrepancy at the (ℓ, m)th terminal by the random variable $u_{\ell m}$, then its probability density function (pdf) is given by

$$p(u_{\ell m}) = \frac{1}{L_{\ell m}} \sum_{i=1}^{L_{\ell m}} \delta(u_{\ell m} - I_{\ell m;i}), \tag{4.21}$$

where $\delta(\cdot)$ denotes the Dirac delta function, and $I_{\ell m;i}$ denotes the value of $I_{\ell m}[n]$ when the ith combination of ± 1 is substituted into $x_{\lambda\mu}[j]$ in Eq. (4.17). We note that many of the $I_{\ell m;i}$ under different combinations of ± 1 are of equal values.

Now, let us consider the probability of decision error due to the combined effects of timing error interference and zero-mean white Gaussian channel noise having power spectral density $N_0/2$. The output noise of the correlator and sampler is a zero-mean Gaussian random variable with variance given by

$$\begin{aligned} E\{\nu_{\ell m}^2[n]\} &= \iint E\{\nu(t)\nu(\tau)\} \frac{1}{\sqrt{E_0}} \phi_{\ell m}(t - nT_\ell + \Delta) \frac{1}{\sqrt{E_0}} \phi_{\ell m}(\tau - nT_\ell + \Delta)\, dt\, d\tau \\ &= \frac{N_0}{2E_0} \iint \delta(t - \tau) \phi_{\ell m}(t - nT_\ell + \Delta) \phi_{\ell m}(\tau - nT_\ell + \Delta)\, dt\, d\tau \\ &= N_0/(2E_0), \end{aligned}$$

where $E\{\cdot\}$ denotes expected value. In the presence of timing error, the expected value of the output of the (ℓ, m)th correlator when a "1" is transmitted at the (ℓ, m)th transmitter terminal is $1 + u_{\ell m}$. Therefore, the probability of error due to channel noise under the influence of timing error is

$$\begin{aligned} P_{\ell m}(\varepsilon|u_{\ell m}) &= \frac{1}{\sqrt{\pi N_0/E_0}} \int_{-\infty}^{0} \exp\left(-\left(v-(1+u_{\ell m})\right)^2/(N_0/E_0)\right) dv \\ &= \frac{1}{2}\operatorname{erfc}\left(\frac{1+u_{\ell m}}{\sqrt{N_0/E_0}}\right), \end{aligned} \tag{4.22}$$

where $\operatorname{erfc}(\alpha) = \frac{2}{\sqrt{\pi}}\int_{\alpha}^{\infty} \exp(-z^2)\, dz$. Hence, the probability of error for the (ℓ, m)th terminal is given by

$$\begin{aligned} P_{\ell m}(\varepsilon) &= \int P_{\ell m}(\varepsilon|u_{\ell m}) p(u_{\ell m})\, du_{\ell m} \\ &= \frac{1}{L_{\ell m}} \sum_{i=1}^{L_{\ell m}} \operatorname{erfc}\left(\frac{1+I_{\ell m;i}}{\sqrt{N_0/E_0}}\right), \end{aligned} \tag{4.23}$$

where Eq. (4.21) has been used in the second step.

Eq. (4.23) yields an exact expression for the probability of decision error in the (ℓ, m)th terminal due to timing error and Gaussian noise. To use Eq. (4.23), the possible interference values $I_{\ell m;i}$ generated by all possible combinations of values of $x_{\lambda\mu}$ filtered by $J_{\ell m}^{\lambda\mu}$ have to be evaluated. Unfortunately, $L_{\ell m}$ may be a large number (e.g., for $\ell = 2$, $M_\ell = 4$ and $N = 14$, $L_{\ell m} = 2^{316}$), which may make Eq. (4.23) difficult to apply. We may, however, approximate Eq. (4.23) using the Central Limit Theorem [60]. Since the random variable $u_{\ell m}$ is a linear combination of a large number of binary random variables, we may approximate it by a random variable $\hat{u}_{\ell m}$ having a Gaussian pdf. Note that the mean of $u_{\ell m}$ is zero and the maximum and minimum values of $u_{\ell m}$ occur when all $x_{\lambda\mu}[j]$ in Eq. (4.17) are $+1$ and -1 respectively, with coefficients $|J_{\ell m}^{\lambda\mu}[n]|$. We may take these maximum and minimum values to be $\pm 3\zeta_{\ell m}$, where $\zeta_{\ell m}$ denotes the standard deviation of $\hat{u}_{\ell m}$. In this way, we have ensured that the probability of finding the approximated random variable $\hat{u}_{\ell m}$ within this range is over 99.7%. Substituting Eq. (4.22) and $p(\hat{u}_{\ell m})$ into Eq. (4.23), interchanging the order of integration, and factorizing the exponential part, the approximated probability of error can be written as

$$
\begin{aligned}
P_{\ell m}(\varepsilon) &= \int_{-\infty}^{\infty} \frac{1}{\zeta_{\ell m}\sqrt{2\pi}} \exp\left(-\hat{u}_{\ell m}^2/(2\zeta_{\ell m}^2)\right) \\
&\quad \int_{-\infty}^{0} \frac{1}{\sqrt{\pi N_0/E_0}} \exp\left(-(v-(1+\hat{u}_{\ell m}))^2/(N_0/E_0)\right) dv\, d\hat{u}_{\ell m} \\
&= \frac{1}{2}\mathrm{erfc}\left(\frac{1}{\sqrt{N_0/E_0+2\zeta_{\ell m}^2}}\right). \qquad (4.24)
\end{aligned}
$$

Eq. (4.24) indicates that a smaller variance of interference $\zeta_{\ell m}^2$ leads to a smaller probability of error in the (ℓ, m)th terminal.

4.4.2 Probability of error under timing jitter

It has been shown that the interference at the terminal (ℓ, m) is a random variable when Δ_n is a timing jitter. Since $I_{\ell m}^{(j)}[n]$ are functions of both $x_{\lambda\mu}[n]$ and Δ_n, to get its exact probability distribution is difficult. Under constant timing errors, the interference $I_{\ell m}[n]$ is approximated by Gaussian distribution as a result of the Central Limit theorem. With Δ_n being a random variable, which is usually assumed to be either uniformly distributed or Gaussian distributed, the assumption that $I_{\ell m}^{(j)}[n]$ is Gaussian distributed may be adopted reasonably. Furthermore, as in the case of constant timing errors, it is assumed that the channel AWGN is independent of $I_{\ell m}^{(j)}[n]$. Thus, the combination of AWGN and the interference $I_{\ell m}^{(j)}[n]$ is still Gaussian distributed with zero mean and variance being $N_0/(2E_0)+\sigma_{I_{\ell m}}^2$.

The probability of error can then be easily derived as

$$
\begin{aligned}
P_{\ell m}(\varepsilon) &= \int_{0}^{\infty} \frac{1}{\sqrt{2\pi(N_0/2E_0+\sigma_{I_{\ell m}}^2)}} \exp\left(-\frac{v^2}{N_0/E_0+2\sigma_{I_{\ell m}}^2}\right) dv \\
&= \frac{1}{2}\mathrm{erfc}\left(\frac{1}{\sqrt{N_0/E_0+2\sigma_{I_{\ell m}}^2}}\right) \qquad (4.25)
\end{aligned}
$$

A smaller variance of interference results in a smaller probability of error in the (ℓm)th terminal.

4.5 Optimum wavelet design

From the discussion in Section 4.3 we can see that timing discrepancy in the WPDM system gives rise to interference in the forms of ISI and cross-talk. Various standard techniques, such as adaptive equalization [1,3,50,51], may be used to reduce this interference, but in WPDM we have an extra degree of freedom: the design of the waveform $\phi_{01}(t)$. Once the scaling function $\phi_{01}(t)$ is determined, all the constituent functions to be used as waveforms are also determined, by Eqs (2.9) and (2.10). Hence we need only design $\phi_{01}(t)$. If we design $\phi_{01}(t)$ to minimize the energy of the timing-discrepancy interference at the root of the WPDM tree, such a function also minimizes the total interference energy, regardless of the wavelet packet tree structure (Proposition 4.3.1).

In order to formulate the optimization problem, recall that $\phi_{01}(t)$ is intimately related to the filter coefficients $h[n]$ via the dilation equation of Eq. (2.4) [34,48]. In practical implementations of a WPDM system, we would also like the resulting $\phi_{01}(t)$ to be of finite duration, which simply requires that $h[n]$ be an FIR filter [34,48]. Once the filter coefficients $h[n]$ are specified, the scaling function $\phi_{01}(t)$ can be obtained using any one of a number of possible algorithms [34,41,42,48], one of which is summarized in Chapter 2. From Eq. (4.10), we see that if the message bits are independent, the variance of the interference at the root of the tree is directly proportional to $\sum_n \left|R'_\phi(nT_0)\right|^2$, no matter Δ_n is constant timing error or timing jitter. Thus, minimization of the interference energy is equivalent to the minimization of $\sum_n \left|R'_\phi(nT_0)\right|^2$. In addition, $R'_\phi(nT_0)$ can be calculated directly from the set of filter coefficients $\{h[n]\}$ using the formulae [61]:

$$R'_\phi(nT_0) = 2R'_\phi(2nT_0) + 2\sum_{k=1}^{N/2} C_{11}^{11}[2k-1]\left[R'_\phi\left((2n-2k+1)T_0\right) + R'_\phi\left((2n+2k-1)T_0\right)\right],$$

and

$$\sum_{k\in\mathcal{Z}} kR'_\phi(kT_0) = -1, \tag{4.26}$$

where N is the length of the filter $h[n]$, and

$$C_{11}^{11}[n] = \sum_{k=0}^{N-k-1} h[k]h[k+n],$$

is the autocorrelation sequence of $h[n]$.

Thus, the problem is reduced to the design of a sequence $h[n]$ to minimize the objective functional

$$F = \sum_{n} \left| R'_{\phi}(nT_0) \right|^2. \tag{4.27}$$

The domain of $h[n]$ in the minimization of F must be constrained so that $h[n]$ satisfies the orthonormality condition of Eq. (2.6), which is re-written here as

$$\sum_{n=2k}^{N-1} h[n]h[n-2k] = \delta[k], k = 0, 1, \ldots, N/2 - 1,$$

for which a necessary condition is that N is even [62]. In addition, to ensure that the locally linear approximation of Eq. (4.9) holds, $R_{\phi}(\tau)$ must be at least first order differentiable, which requires the scaling function $\phi_{01}(t)$ to satisfy certain differentiability conditions. For FIR filters, the differentiability constraint on $R_{\phi}(\tau)$ is satisfied if we choose a filter satisfying the following 'regularity condition' [34,48] with $K_0 \geq 1$:

Proposition 4.5.1 *Let $h[n]$ be a sequence of length N starting at $n = 0$ which satisfies the orthogonality constraint in Eq. (2.6) and has a Z-transform of the form*

$$H(z) = \sqrt{2}\left(\frac{1+z^{-1}}{2}\right)^{K} P(z), \tag{4.28}$$

where $1 \leq K \leq N/2$ and $P(z)$ is a polynomial in z^{-1} of degree $N - K - 1$ with $|P(1)| = 1$. If $P(z)$ is bounded such that

$$|P(e^{j\omega})| < B, \quad \text{for all } \omega \in [0, \pi] \tag{4.29}$$

then $R_{\phi}(\tau)$ has (at least) $K_0 = 2K - 1 - \lceil 2\log_2 B \rceil$ continuous derivatives, where $\lceil x \rceil$ denotes the least integer $\geq x$.

Proof: The proof is based on the fact that a sufficient condition for a function $\psi(t)$ to have $k \in \mathcal{N}$ continuous derivatives is (e.g. [48])

$$\int |\Psi(f)|(1+|f|)^{k+1}\, df < \infty \tag{4.30}$$

where $\Psi(f) := \int \psi(t)e^{-j2\pi ft}\, dt$ is the Fourier transform of $\psi(t)$. Let

$$\beta := \sup_{\omega} |P(e^{j\omega})| = \max_{\omega} |P(e^{j\omega})| < B,$$

and for notational simplicity set $T_0 = 1$. Using a result of Tchamitchian (cited as Lemma 3.2 of Daubechies [48]) for $h[n]$ satisfying the conditions of the proposition, there exists a constant c such that

$$|\Phi_{01}(f)| \leq c(1+|f|)^{-K+\log_2 \beta}.$$

Let $\tilde{R}(f)$ denote the Fourier transform of $R(\tau)$. Since $\tilde{R}(f) = |\Phi(f)|^2$ we have that

$$|\tilde{R}(f)| \leq c^2(1+|f|)^{-2K+2\log_2 \beta}.$$

Hence

$$\int |\tilde{R}(f)|(1+|f|)^{k+1}\, df \leq c^2 \int (1+|f|)^{k+1-2K+2\log_2 \beta}\, df. \tag{4.31}$$

Using the fact that $\int (1+|f|)^{-x}\, df < \infty$ if $x > 1$, Eqs (4.30) and (4.31) imply that $R(\tau)$ has (at least) K_1 continuous derivatives, where

$$K_1 < 2K - 1 - 2\log_2 \beta.$$

Therefore, $R(\tau)$ has at least $K_2 \leq 2K - 1 - 2\log_2 B$ continuous derivatives. Consequently $R(\tau)$ has at least $K_0 = 2K - 1 - \lceil 2\log_2 B \rceil$ continuous derivatives, and hence the proposition.

For given N and K, the factorization of Eq. (4.28) can be ensured by having [63]

$$\sum_{n=0}^{N-1} (-1)^n n^k h[n] = 0, k = 0, 1, \ldots, K-1,$$

and for a given B, since

$$\left|H\left(e^{j\omega}\right)\right| = \left|\sum_{n=0}^{N-1} h[n]e^{-j\omega n}\right|,$$

the frequency response bound of Eq. (4.29) can be enforced by requiring that

$$\left|\sum_{n=0}^{N-1} h[n]e^{-j\omega n}\right| - 2^{-K+1/2}\left|(1+e^{j\omega})^K\right|(B-\epsilon) \le 0, \tag{4.32}$$

for all $\omega \in [0, 2\pi]$ and an appropriate small positive number ϵ. The constraint in Eq. (4.32) is difficult to apply when we seek a numerical solution, since ω is a continous variable, and the left side of (4.32) is a periodic function with period 2π. We discretize the ω in $[0, 2\pi)$ by the interval $2\pi/M$. Instead of Eq. (4.32), we impose the equivalent condition,

$$\left|\sum_{n=0}^{N-1} h[n]e^{-j2\pi nk/M}\right|^2 - 2^{-2K+1}\left|(1+e^{j2\pi k/M})^{2K}\right|(B-\epsilon)^2 \le 0, \tag{4.33}$$

for $k = 0, 1, \ldots, M-1$. Eq. (4.33) and (4.32) are equivalent at $\omega = 2\pi n/M, n = 0, 1, \ldots, M-1$. However, in between $\omega = 2\pi n/M$ and $\omega = 2\pi(n+1)/M$, $|P\left(e^{j\omega}\right)|$ might not be bounded by B. On the other hand, $H(z)$ cannot contain singularities on the unit circle, since $H(z)$ is a polynomial in z^{-1} (it is an FIR filter). This ensures that for an appropriate ϵ, the frequency response bound in Eq. (4.32) is not broken between the discrete constraint frequencies in Eq. (4.33).

From the above discussion, the criterion for the design of a scaling function $\phi_{01}(t)$ for WPDM which minimizes the total energy of timing-error interference can be summarized as follows. For the chosen N (even), the length of the FIR filter $h[n]$, K, and B,

1. The objective function is

$$\min \sum_k |R'_\phi(kT_0)|^2 \quad \text{with respect to} \quad h[n]. \tag{4.34}$$

2. The orthogonality constraint is

$$\sum_k h[k]h[k-2n] = \delta[n], \qquad n \in \mathcal{Z}. \tag{4.35}$$

3. The regularity constraint is

$$\sum_k (-1)^k k^n h[n] = 0, n = 0, 1, \ldots, K, \tag{4.36}$$

and

$$\left|\sum_{n=0}^{N-1} h[n]e^{-j2\pi nk/M}\right|^2 - 2^{-2K+1}\left|(1+e^{j2\pi k/M})^{2K}\right|(B-\epsilon)^2 \leq 0, \tag{4.37}$$

The optimization is a typical constrained optimization problem. Since how to obtain an optimal solution is not the main theme of the thesis, we simply use one of the well-developed packages called *constr* in MATLAB [64] to get a numerical solution. In *constr*, Sequential Quadratic Programming(SQP) methods are used, which allow one to closely mimic Newton's method for constrained optimization just as is done for unconstrained optimization. At each major iteration an approximation is made of the Hessian of the Lagrangian function using a quasi-Newton updating method. This is then used to generate a Quadratic Problem(QP) sub-problem whose solution is used to form a search direction for a line search procedure. An overview of SQP is found in [65], and the implementation of *constr* is clearly explained in [64, 66].

The procedure for the design of a scaling function $\phi_{01}(t)$ using the above criterion is as follows:

1. Choose N (even), the length of the FIR filter $H(z)$. Choose K $(1 \leq K \leq N/2)$, the number of zeros of $H(z)$ at $z = -1$, and $B > 1$ such that $K_0 = 2K - 1 - \lceil 2\log_2 B \rceil \geq 1$, the number of continuous derivatives of $R_\phi(\tau)$, satisfies $K_0 \geq 1$.

2. For a set of filter coefficients $h[n], n = 0, \ldots, N-1$, $R'_\phi(nT_0)$ can be calculated using Eq. (4.26), and thus the cost functional F of Eq. (4.27) can be evaluated.

3. A new functional G which is the sum of F and the $(N/2 + K)$ equality constraints of Eq. (4.35) and (4.36) and the inequality constraints of (4.37) multiplied by their

corresponding Langrange multipliers can then be formed.

4. Employ the *constr* routine in MATLAB to find a stationary point of G and hence obtain the optimum set of filter coefficients $\tilde{h}[n]$.

5. The optimum scaling function $\tilde{\phi}(t)$ can then be obtained.

When a solution $\tilde{h}[n]$ is produced from *constr*, the gradient is checked so as to make sure that $\tilde{h}[n]$ is at least a local minimum. It is well-known that it is very difficult to obtain global optimization for a non-quadratic functional of multiple parameters. We employ a simple strategy to get a solution as close as possible to the global minimum. The strategy is that we proceed to local minima from a lot of random starting points, and then compare the functional at these local minima. The local minimum which gives the smallest objective functional is chosen as the best solution.

4.5.1 A design example

We employ the above procedure to design an optimum scaling function for the WPDM system in the following example:

Example 4.5.1 In this example, we show how we designed a scaling function $\phi_{01}(t)$ to minimize the total interference energy due to timing discrepancy in the WPDM system. For the design, we chose $N = 14$, $K = 2$ and $B = \sqrt{2} - 10^{-6}$ so that $R_\phi(t)$ was guaranteed to be at least first order differentiable. A procedure based on the steps outlined above was employed to obtain the optimum filter coefficients $\{\tilde{h}[n]; n = 0, 1, \ldots, 13\}$ subject to the orthonormality and regularity constraints. The values of these optimal coefficients are shown in Table 4.1. The optimum scaling function $\tilde{\phi}_{01}(t)$ was then obtained from the set of filter coefficients by the "cascade" algorithm [48] we summarized in Chapter 2. The above procedure generated an approximation of the optimum $\tilde{\phi}_{01}(t)$ "off-line", which is shown in Fig. 4.2. Its energy spectrum and autocorrelation function are shown in Fig. 4.3 and Fig. 4.4, respectively.

Table 4.2 (last column) shows the comparison of the total interference due to timing error between the optimum scaling function ($N = 14$) and two scaling functions which have been widely used in other signal processing applications — the function generated by the Daubechies filter of length $N = 14$ [34, 48] and the Lemarié-Battle scaling function based on cubic splines [67, 68] (approximated by finite-duration function of length $31T_0$). In the table, timing error Δ is normalized to 1. It can be observed from Table 4.2 that the use of the optimum scaling function yields the smallest amount of interference due to timing error.

To illustrate the distribution of the interference energy in the case of constant timing error, we also show in the first four columns of Table 4.2 the values of the ISI energies $\sum_n |J_{11}^{11}[n]|^2$ and $\sum_n |J_{12}^{12}[n]|^2$ at the first and second terminals of the first level of the WPDM tree, as well as the respective cross-talk terms $\sum_n |J_{11}^{12}[n]|^2$ and $\sum_n |J_{12}^{11}[n]|^2$. It is confirmed that the interference energy distribution is structure dependent, although $\sum_n |J_{11}^{12}[n]|^2 = \sum_n |J_{12}^{11}[n]|^2$.

Table 4.3 shows how the interference energy is distributed under timing jitters. As expected, the distribution of the interference energy is indeed independent of tree structure, since $\sum_n |J_{11}^{11}[n]|^2 + \sum_n |J_{11}^{12}[n]|^2 = \sum_n |J_{12}^{12}[n]|^2 + \sum_n |J_{12}^{11}[n]|^2$; i.e.,different terminals have the same amount of interference energy.

It should also be noted that the mapping from signal waveform to energy is many-to-one, and therefore that $\tilde{\phi}(t)$ is not unique. A different initial point in the optimization may lead to a different optimum $\tilde{\phi}(t)$ which also minimizes the interference energy.

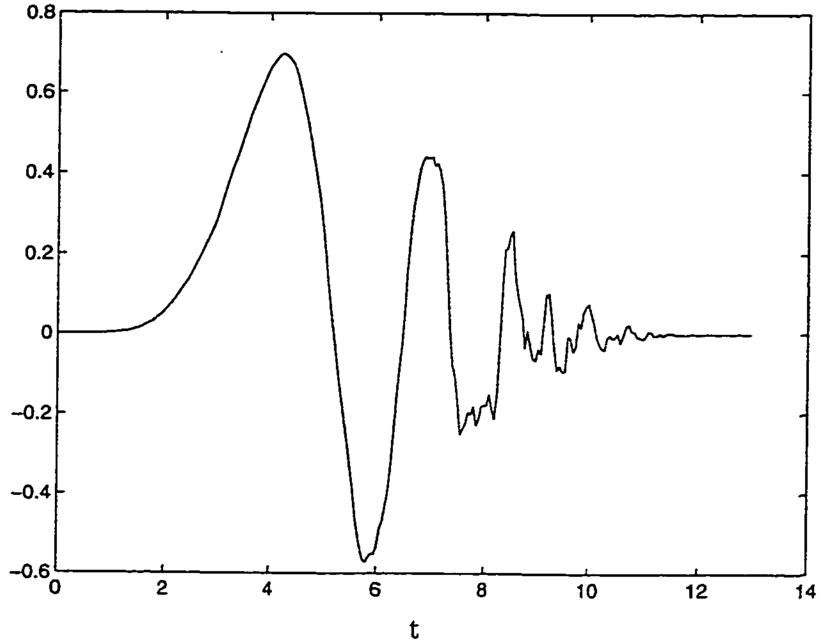


Figure 4.2: Optimum scaling function $\tilde{\phi}_{01}(t)$ from Example 4.5.1.

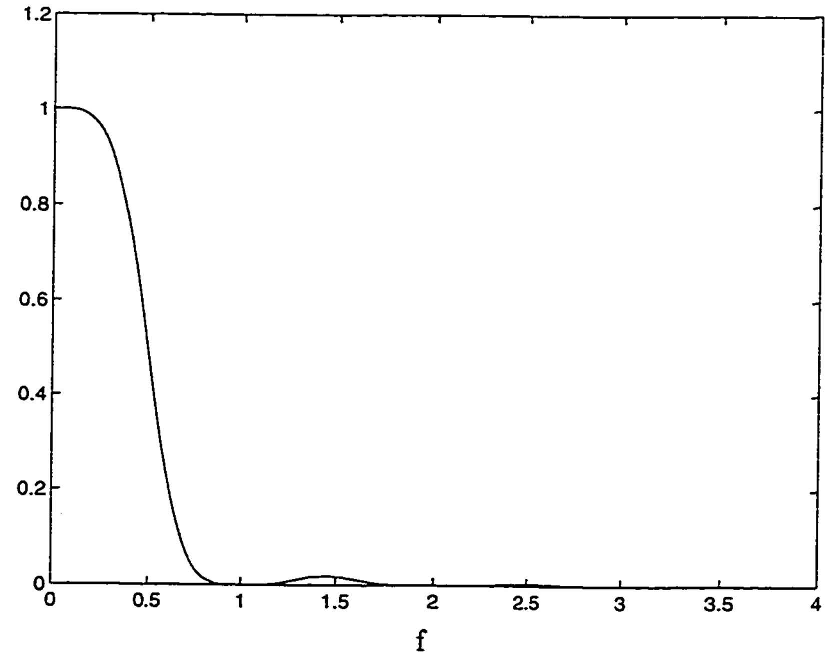


Figure 4.3: Energy spectrum of $\tilde{\phi}_{01}(t)$ from Example 4.5.1.

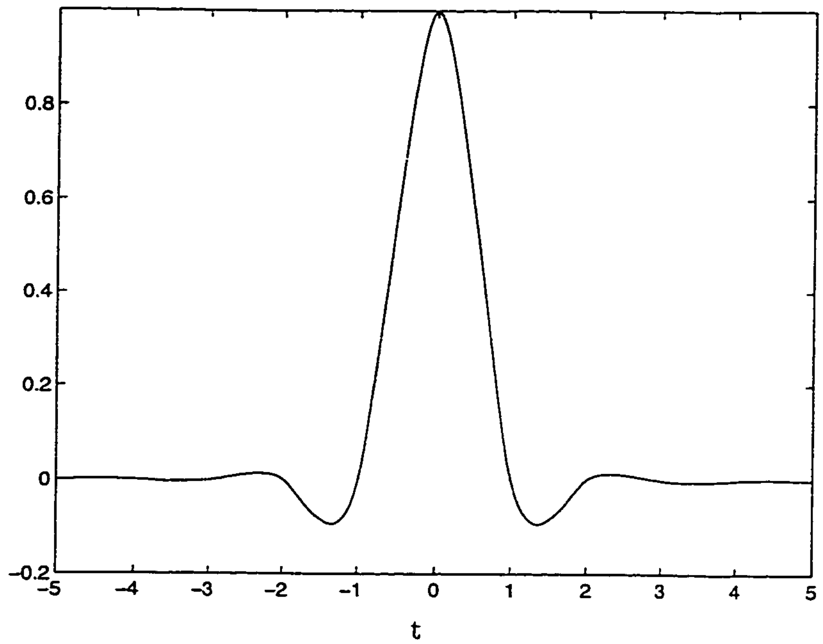


Figure 4.4: The autocorrelation function of the optimum scaling function $\tilde{\phi}_{01}(t)$ from Example 4.5.1.

$\tilde{h}[0]$	$\tilde{h}[1]$	$\tilde{h}[2]$	$\tilde{h}[3]$	$\tilde{h}[4]$	$\tilde{h}[5]$	$\tilde{h}[6]$
0.021474	0.061336	0.144399	0.269674	0.427230	0.513130	0.364040

$\tilde{h}[7]$	$\tilde{h}[8]$	$\tilde{h}[9]$	$\tilde{h}[10]$	$\tilde{h}[11]$	$\tilde{h}[12]$	$\tilde{h}[13]$
-0.090059	-0.463195	-0.042816	0.299628	-0.034431	-0.086468	0.030273

Table 4.1: Coefficients of the optimum filter $\tilde{h}$ in Example 4.5.1.

	$\sum\|J_{11}^{11}\|^2$	$\sum\|J_{11}^{12}\|^2$	$\sum\|J_{12}^{11}\|^2$	$\sum\|J_{12}^{12}\|^2$	total
optimum	0.261	0.542	0.542	0.742	2.086
Daubechies	0.422	0.464	0.464	2.024	3.374
Lemarié-Battle	0.548	0.304	0.304	3.228	4.384

Table 4.2: Interference energies for several scaling functions from Example 4.5.1 with constant timing error.

4.6 Numerical results

We now examine the transmission error rates for WPDM system under the influence of timing error and Gaussian channel noise in the following example.

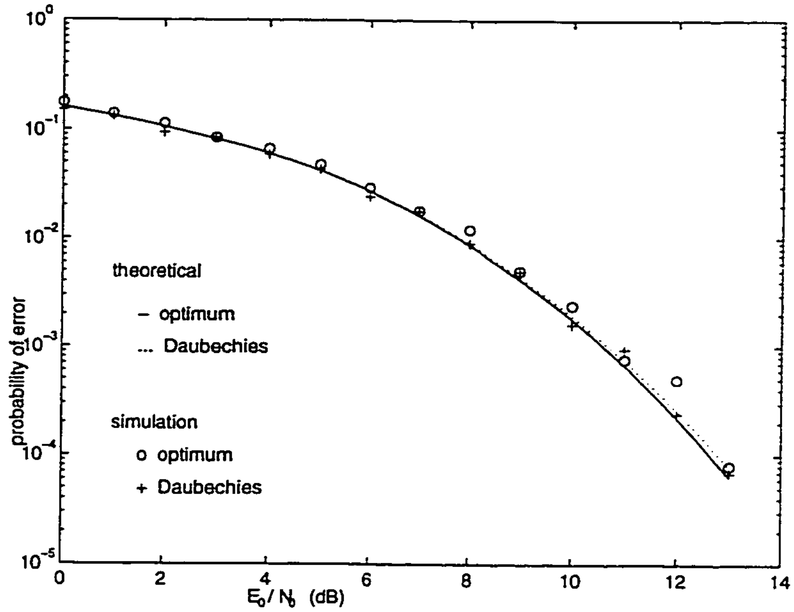
The optimum scaling function design in Section 4.5 seeks to minimize the total energy of the interference. Thus, we may expect the average probability of error over all the terminals using the optimum scaling function to be lower than that using other scaling functions. This is also illustrated in the example.

Example 4.6.1 Consider a WPDM system consisting of two branches, using $\phi_{11}(t-nT_1)$ and $\phi_{12}(t-nT_1)$ as carriers for the binary message signals $x_{11}[n]$ and $x_{12}[n]$. The transmission of the modulated signal through a communication channel is disrupted by spectrally white zero-mean Gaussian noise the spectral density of which is $N_0/2$. At the receiver, the signal is demodulated by a correlator and sampler at the root of the WPDM tree which has a timing error Δ_n relative to the transmitted carrier scaling function. The output of the correlator and sampler, $\hat{x}_{01}[n]$, is then passed through a wavelet packet decomposer to obtain the sequences $\hat{x}_{11}[n]$ and $\hat{x}_{12}[n]$.

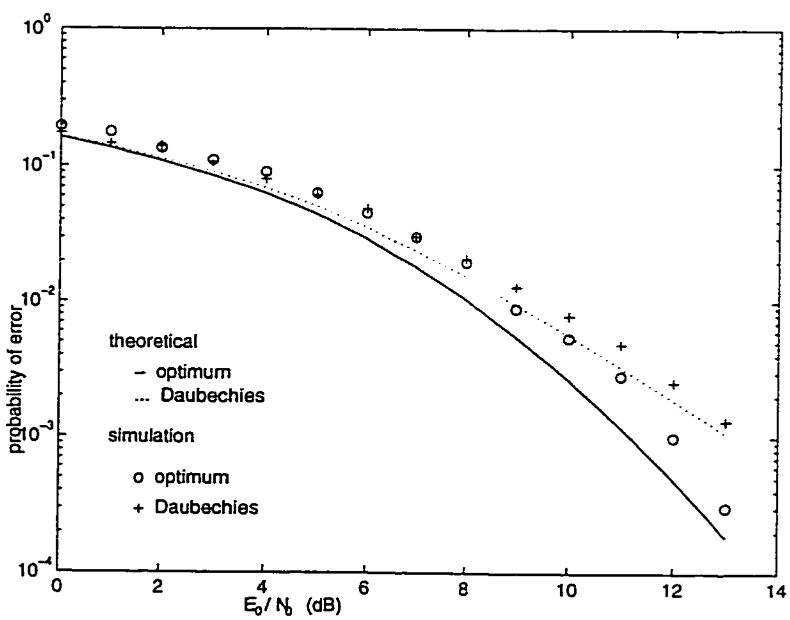
The performance of such a system, in terms of the bit error rate under different bit energy to N_0 ratios (E_0/N_0), was evaluated both theoretically and by simulation.

For a constant timing discrepancy of $\Delta = 0.15T_0$, the results obtained by employing the optimum scaling function designed in Example 4.5.1 and the Daubechies scaling functions of the same duration, are plotted in Figs 4.5(a) and 4.5(b) for the the message signals at the $(1,1)$ and $(1,2)$ terminals respectively.[1] The theoretical error rates given by Eq. (4.24) are also plotted, for reference. It can be observed that for the message signal $x_{11}[n]$, the performance of the WPDM scheme using the optimum scaling function $\tilde{\phi}_{01}(t)$ is almost

[1]For easier viewing of the graphs, the results obtained using the Lemarié-Battle scaling function (which are inferior to the Daubechies functions due to higher interference) are omitted.



(a) (1, 1) terminal.



(b) (1, 2) terminal.

Figure 4.5: The probability of error for a range of E_0/N_0 for the two-channel scenario studied in Example 4.6.1.

identical to that obtained by the Daubechies function, both in simulation and theoretically. Also, the theoretical analyses agree very well with the simulation results.

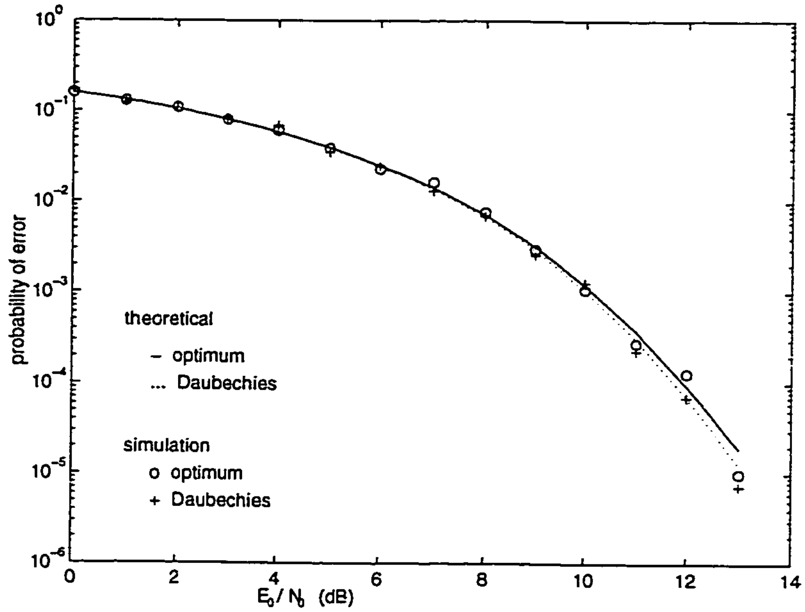
For the message $x_{12}[n]$, the theoretical performance using $\tilde{\phi}_{01}(t)$ is almost the same as the corresponding performance for $x_{11}[n]$, but the theoretical performance using Daubechies functions has greatly deteriorated. This can be easily explained by the difference in the amount of interference at the $(1,2)$ terminal as shown in Table 4.2. With much higher interference, using the Daubechies function will certainly result in inferior performance. That the use of $\tilde{\phi}_{01}(t)$ results in superior performance to the Daubechies functions is also confirmed by the simulations. However, there is a discrepancy between the theoretical evaluation of performance and that from simulation in each case. The discrepancy can be attributed to the terms containing second and higher-order powers of Δ being ignored in Eq. (4.9). The second and higher-order terms in Eq. (4.9) contain second and higher-order derivatives of $R_{\phi}(\tau)$ and therefore tend to have little or no low-frequency content. (Differentiation in the time domain is equivalent to multiplication by ω in the frequency domain.) These higher-order terms are passed through a LP filter $h[n]$ and a HP filter $g[n]$ (just as the first-order terms are) to form parts of the interference at the $(1,1)$ and $(1,2)$ terminals, respectively. Due to the HP nature of these higher-order interference terms, their contribution to the interference in the $(1,2)$ terminal is much more significant than that to the $(1,1)$ terminal. Thus, ignoring these terms in the linear approximation will lead to a larger discrepancy between the theoretical and simulated results at the $(1,2)$ terminal than at the $(1,1)$ terminal. In Fig. 4.5(b) the difference in performance between using the optimum scaling function $\tilde{\phi}_{01}(t)$ and the Daubechies function become significant only when the E_0/N_0 is relatively high ($\geq$ 8dB). This is because at low E_0/N_0, the performance is dominated by errors due to the Gaussian channel noise.

Figs 4.6(a), (b), (c) and (d) show a comparison of the performance of a second level

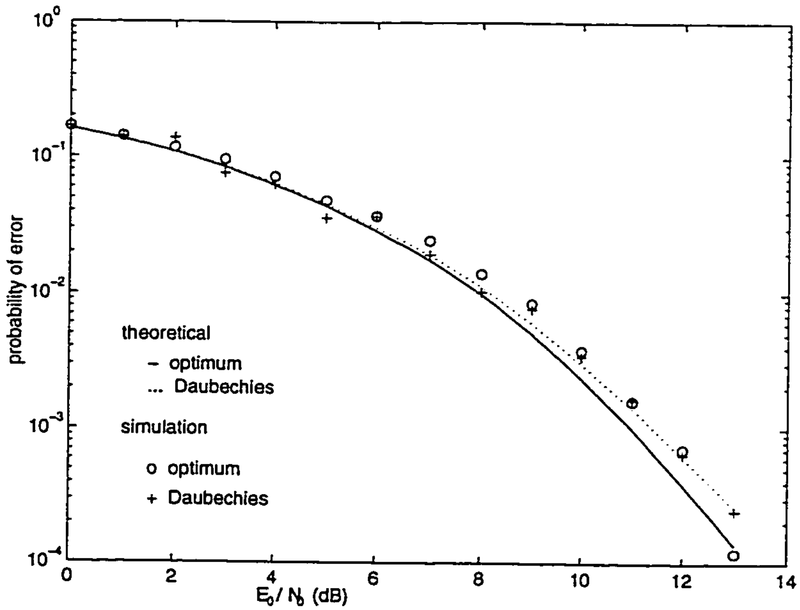
($\ell = 2$, $M_2 = 4$) WPDM tree structure using the constituent functions of the optimum scaling function $\tilde{\phi}_{01}(t)$ and the corresponding Daubechies function.

(The theoretical performance is evaluated using Eq. (4.24).) Again, at the first two terminals (i.e., the $(2,1)$ and $(2,2)$ terminals), the difference between the performance of the optimal and Daubechies schemes are rather small. However, at the $(2,3)$ and $(2,4)$ terminals, the use of the constituent functions of $\tilde{\phi}_{01}(t)$ results in superior performance to the use of the constituent functions of the Daubechies scaling function.

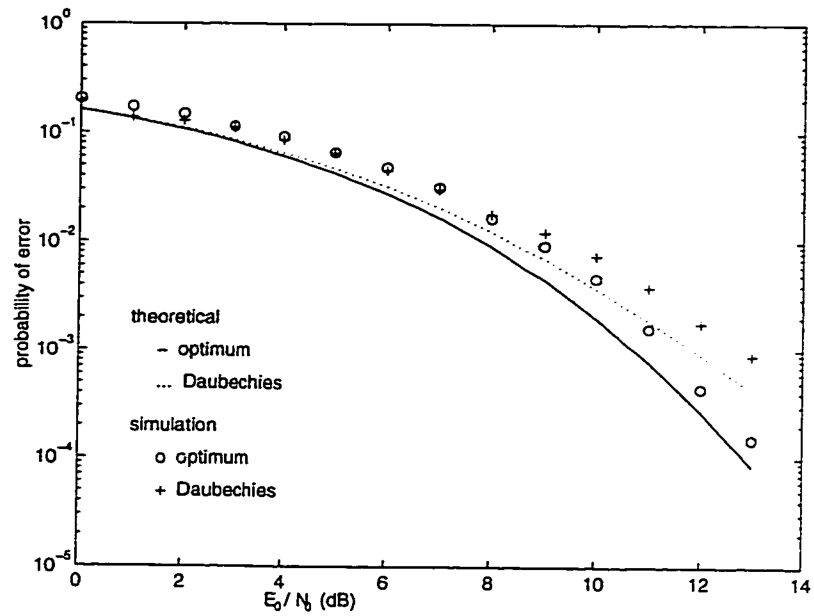
For a timing jitter of uniformly distributed in $[-0.15T_0, 0.15T_0]$, the performance by exploring the same optimum scaling function $\tilde{\phi}(t)$ and the Daubechies scaling function of $N = 14$ is shown in Fig. 4.7. Several observations can be made from Fig. 4.7. First, as under constant timing errors, the optimum scaling function results in a lower probability of error than the Daubechies scaling function with the same compact length. Second, in contrast with the situation of constant timing error, in both (1,1) and (1,2) terminals, the theoretical analysis closely agrees with the simulation results. One of the reasons is that the variance of Δ_n (unformly distributed in $[-0.15T_0, 0.15T_0]$) is smaller than Δ^2 ($\Delta = 0.15T_0$). With smaller timing error, less error is introduced when we make the approximation in Eq. (4.7). However, there is a more important reason here. The ignored high order terms of Δ_n, which cause 1-2dB discrepancy at the terminal (1,2) between theoretical and simulation results in the case of constant timing errors, are no longer having the property of a high-pass filter, since Δ_n is a random variable. This implies that the ignored high order terms of Δ_n have the same effect on terminal (1,1) as on (1,2) terminal. Third, from both the theoretical and simulation results, the same probability of error present in the terminals (1,1) and (1,2). This further confirms that the distribution of interference energy under timing jitter is structure independent.



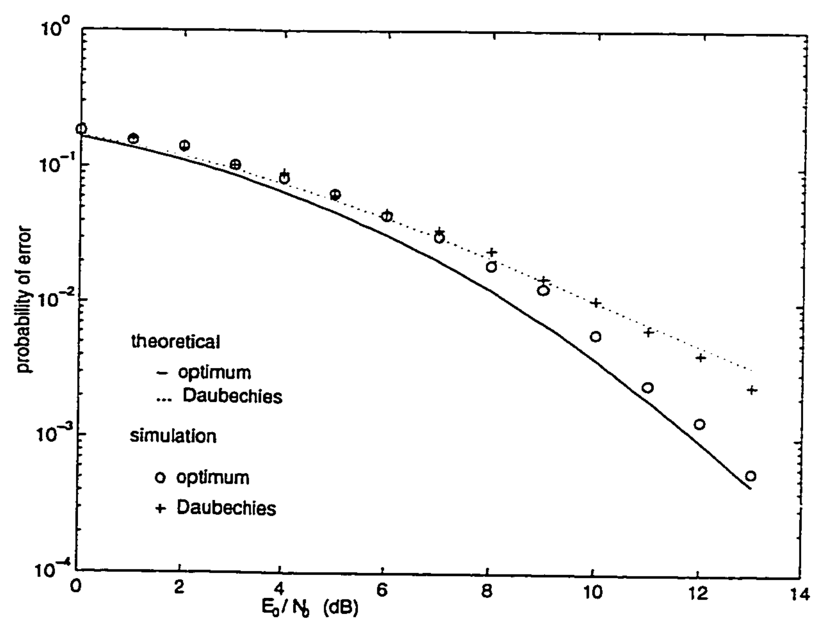
(a) (2, 1) terminal.



(b) (2, 2) terminal.

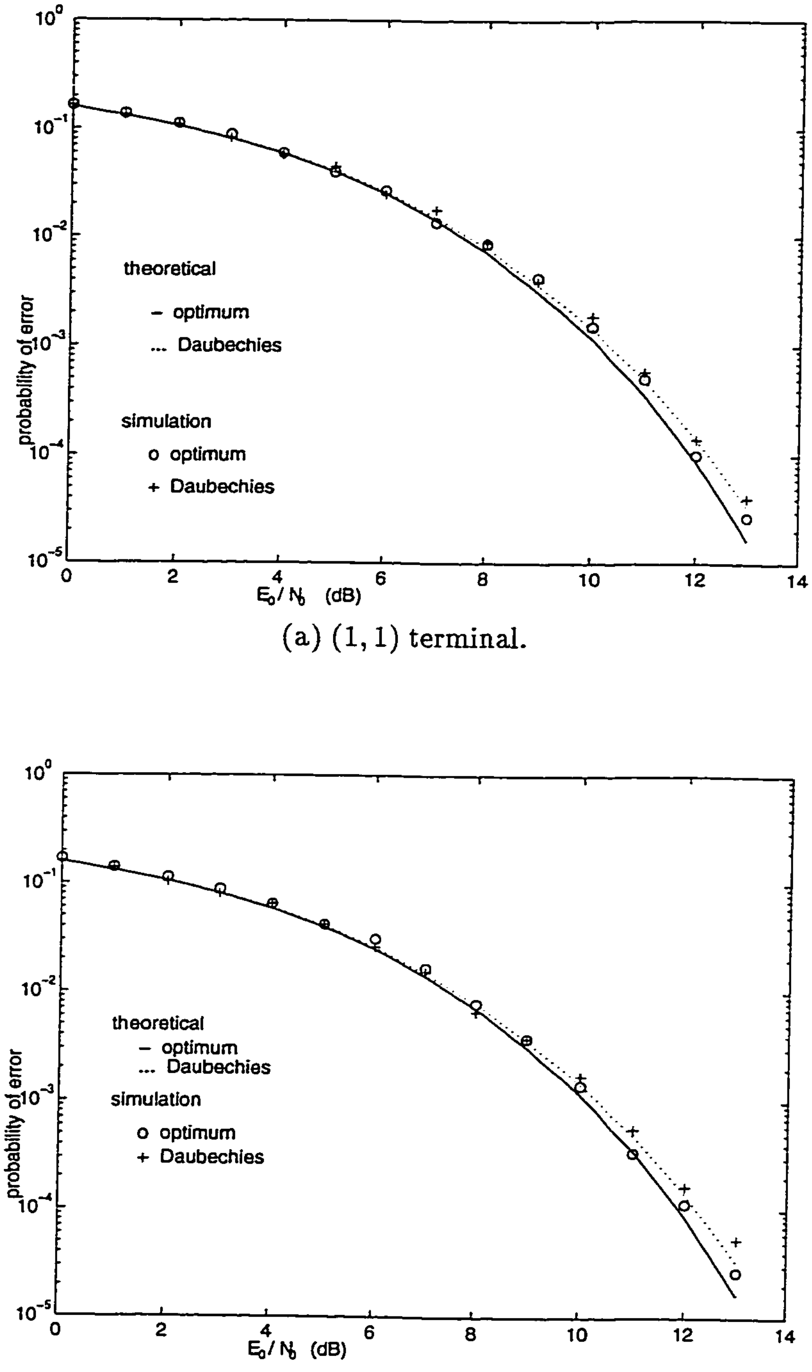


(c) $(2,3)$ terminal.



(d) $(2,4)$ terminal.

Figure 4.6: The probability of error for a range of E_0/N_0 for the four channel scenario studied in Example 4.6.1.



(a) (1, 1) terminal.

(b) (1, 2) terminal.

Figure 4.7: The probability of error for a range of E_0/N_0 for the two-channel scenario studied in Example 4.6.1 under timing jitters.

4.7 Summary

In this chapter, we have examined the problem of timing error (both constant timing errors and timing jitters) in WPDM and observed that the interference due to a timing discrepancy between the transmitter and receiver is made up of two components, ISI and crosstalk, which can be respectively modeled as the message bits from the given channel and the adjacent channels passed through corresponding filters. It is also found that constant timing errors and timing jitters have different properties. In the case of constant timing errors, the interference due to timing error is wavelet packet structure dependent. On the other hand, the distribution of the interference energy due to timing jitters is independent of tree structure. In spite of the difference, the simple interference models led us to propose a simple optimum design of the coding waveforms (i.e., the scaling function and its constituents) in which the total timing-error interference is minimized. Interestingly, an optimum wavelet serves both constant timing errors and timing jitters. We have also obtained expressions for the probability of error under the effects of timing error and additive white Gaussian channel noise. Simulations have confirmed these expressions and shown that employing the optimum scaling function as the fundamental waveform results in a lower probability of error in transmission than other commonly used scaling functions.

	$E\{\sum \lvert Y_{11}^{11}\rvert^2\}$	$E\{\sum \lvert Y_{11}^{12}\rvert^2\}$	$E\{\sum \lvert Y_{12}^{11}\rvert^2\}$	$E\{\sum \lvert Y_{12}^{12}\rvert^2\}$
optimum	0.326	0.717	0.476	0.567
Daubechies	0.422	1.265	0.463	1.224

Table 4.3: Interference energies for scaling functions from Example 4.5.1 with timing jitters.

Chapter 5

WPDM in Impulsive Noise Channels

5.1 Impulsive noise

Impulsive noise, along with thermal noise, crosstalk, and intermodulation, constitutes four categories of noise in communications [69, 70]. Impulsive noise can emanate from natural sources such as lightning, or from man-made sources such as switching transients [70]. The out-of-band intermodulation products due to saturated amplifiers were also found to have a similar effect on the probability of error performance to that of the impulsive interference [31]. Impulsive noise is sporadic, may occur in bursts, and has broad spectrum. Usually when a noise burst is said to be impulsive, its spectrum is flat across the passband of the RF filter at the receiver. Therefore, an impulse burst causes a response in the RF filter which is essentially the same as the impulse response of the filter [29–32].

In most cases, impulsive noise can be neglected for telephone speech transmission. However, for data transmission, it can be a major source of system performance degradation. Therefore, evaluating the probability of error of various digital communication systems in the presence of impulsive noise has received considerable attention. Bello and Esposito proposed an approach of analyzing the error rates and applied it to a BPSK system [29] and a DPSK system [30]. In both cases, impulsive noise was the only noise source under consideration. Oshita and Feher subsequently extended the analysis to an environment which included both impulsive and Gaussian noise [31,32].

In previous chapters, Wavelet Packet Division Multiplexing (WPDM) was introduced as a bandwidth efficient multiplexing scheme. When it is used in the extreme case as WC/FDM, WPDM has the advantages of a parallel transmission system, one of which is the great immunity to impulsive noise. Intuitively, one noise burst, which can destroy one bit in TDM, will be spread over all the subchannels of a WPDM system. For each channel, the spread effects of the impulsive noise may not be strong enough to cause an error. In this chapter, we evaluate the system performance of WPDM in the presence of impulsive and Gaussian noise. An expression for the probability of error for WPDM will be derived. The derivation extends Oshita and Feher's work on the performance of digital communication systems when impulsive noise and Gaussian noise are the causes of error [29–32]. (Part of the analysis was reported in [27].)

5.2 Channel models

To transmit the multiplexed baseband signal $s(t)$, carrier modulation or RF modulation is often necessary. We assume double side-band suppressed carrier (DSB-SC) modulation with carrier frequency f_0 is performed on $s(t)$ before transmission; i.e., the transmitted signal is

$$s_0(t) = s(t)\cos(2\pi f_0 t). \tag{5.1}$$

The system to be analyzed in impulsive noise channels can be generally represented as the cascade of two processing operations, the first of which is RF filtering, as shown in Fig.

5.1. In the figure, $p(t)$ is impulsive noise, $\nu(t)$ is AWGN with spectral density $N_0/2$, and $s_0(t)$ is defined in Eq. (5.1).

The impulsive noise $p(t)$ has a broad spectrum and is essentially flat over the RF filter bandwidth. Each noise burst gives rise to an output which takes the shape of impulse response of the BPF. Therefore, the low-pass equivalent of $p(t)$ can be modeled as

$$p_c(t) = p(t)\exp(-j2\pi f_0 t) = \sum_{i=-\infty}^{\infty} a_i \exp(j\theta_i)\delta(t - \tau_i'), \tag{5.2}$$

where a_i and τ_i' are the random amplitude and occurrence time of the ith noise burst, respectively, and θ_i is the phase of the ith noise impulse relative to the carrier. Usually, a_i is assumed to have a log-normal distribution and the pulse arrival is assumed to have Poisson distribution [71]. Furthermore, the average arrival rate ν_a of noise bursts in one symbol duration T_0 is assumed to be small. For instance, for a wide variety of practical applications, it has been shown that $\nu_a T_0$ is in the order of 10^{-3} [29]. For a Poisson distribution with arrival rate $\nu_a T_0$, we know that the interval t_p between two events has an exponential distribution

$$G_{t_p}(t) = \nu_a T_0 \exp(-\nu_a T_0 t).$$

Given an impulse burst happening at t_1, it can be computed that the probability of the next burst occurring in $(t_1, t_1 + \Delta t)$ is

$$\begin{aligned} P &= \int_0^{\Delta t} G_{t_p}(t)dt \\ &= \int_0^{\Delta t} \nu_a T_0 \exp(-\nu_a T_0 t)\; dt \\ &= 1 - \exp(-\nu_a T_0 \Delta t). \end{aligned}$$

For a small $\nu_a T_0$, such as $\nu_a T_0$ being on the order of 10^{-3}, the probability of the two bursts occurring in $20T_0$ is small.

The receiver receives the transmitted signal together with channel Gaussian noise and impulsive noise. The low-pass equivalent of the received signal is:

$$r_p(t) = s(t) + \nu_c(t) + p_c(t), \tag{5.3}$$

where $\nu_c(t)$ is the complex envelop of AWGN $\nu(t)$.

5.3 Performance analysis

Under perfect carrier synchronization in the demodulator (frequency and phase synchronization) in Fig. 5.1, the received baseband signal of our interest is

$$r_i(t) = s(t) + \nu_I(t) + \sum_i a_i \cos\theta_i h_p(t - \tau_i') \tag{5.4}$$

where $\nu_I(t)$ is the in-phase component of AWGN, whose spectral density is N_0 [3], and $h_p(t)$ is the equivalent low pass of the impulse response of the RF filter. The assumption of a symmetrical bandpass filter makes $h_p(t)$ real. Note that the BPF may distort $s(t)$. However, by assuming the bandwidth of the BPF to be wide compared with that of $s(t)$, the effect of the BPF on $s(t)$ is negligible. For Gaussian noise, since we are interested in the part in signal bandwidth, the shaping effect of the BPF can also be ignored.

In order to simplify the analysis of the probability of error due to impulsive and Gaussian noise, the following reasonable assumptions are made [29–32].

1. The random variables a_i, τ_i', θ_i and $\nu_I(t)$ are statistically independent.

2. The random variables τ_i' and θ_i are uniformly distributed over the intervals T_0 and $[0, 2\pi)$, respectively. Let $G_{\tau'}(\tau_i')$ and $G_\theta(\theta_i)$ be the probability density functions of τ_i' and θ_i, respectively. Then,

$$G_{\tau'}(\tau_i') = \begin{cases} \frac{1}{T_0}, & kT_0 \le \tau_i' \le (k+1)T_0 \\ 0, & \text{otherwise,} \end{cases}$$

and

$$G_\theta(\theta_i) = \frac{1}{2\pi}, \quad 0 \le \theta_i \le 2\pi$$

3. The $\nu_a T_0$ is assumed to be much smaller than unity, and the probability of more than one impulse in T_0 is negligible.

To recover the information sequence $x_{\ell m}[n]$, the demodulated signal $r_p(t)$ passes through the wavelet packet demultiplexer, for which we use the structure in Fig. 3.2. That is, $r_p(t)$ first passes through a matched filter matching to the scaling function $\phi_{01}(t)$ and a sampler to get estimated data sequence $\hat{x}_{01}[n]$, which is then decomposed by discrete wavelet packet decomposer to get the estimated $\hat{x}_{\ell m}[n]$.

In WPDM, the duration of the support length of the compactly supported scaling function $\phi_{01}(t)$ is equal to $(N-1)T_0$ where N is the length of the FIR filters employed in the wavelet packet multiplexer. By defining the effective support length D_{eff} as the region beyond which $|\phi_{01}(t)|$ is very close to 0, D_{eff} of the Daubechies wavelets is about 3 or 4 times of T_0, which is much less than $(N-1)T_0$. Thus, after the matched filter and sampler, one impulse noise burst results in a short sequence containing 3 or 4 non-trivial values depending on D_{eff}. Passing through the short sequence through wavelet packet decomposer, we obtain the interferece from the impulsive burst at each terminal. The interference has just small numbers of non-trivial values, because the wavelet packet decomposer is an FIR followed by a down-sampler. Since the length of the FIR filter in the wavelet packet decomposer is usually longer than the length of the short sequence resulting from one impulsive noise burst, the input of the FIR filter might include the effect of more than one impulsive noist bursts. However, we have shown that for small $\nu_a T_0$, the probability of two impulsive noise bursts occurring in $20T_0$ is small. Here we wish to emphasize that $20T_0$ is long enough for most FIR filters, especially when the effective length of FIR filters is considered. Thus, we may simply ignore the effect of two impulsive noise bursts and assume that one data symbol at terminal (ℓm) is affected either by AWGN alone or by AWGN and one impulse burst.

The estimated data sequence $\hat{x}_{01}[n]$ is therefore given by

$$\begin{aligned}
\hat{x}_{01}[n] &= \frac{1}{\sqrt{E_0}}\int r_p(t)\phi_{01}(t-nT_0)\,dt \\
&= x_{01}[n] + \frac{\nu_{01}[n]}{\sqrt{E_0}} + \frac{F_{01}[n]a_i\cos\theta_i\, c[n,\tau_i']}{\sqrt{E_0}},
\end{aligned} \tag{5.5}$$

where $\frac{1}{\sqrt{E_0}}$ is introduced as a normalization factor to ease our analysis,

$$\nu_{01}[n] = \int \nu_I(t)\phi_{01}(t-nT_0)dt, \tag{5.6}$$

$$c[n,\tau_i'] = \int_{nT_0}^{(N-1+n)T_0} h_p(t-\tau_i')\phi_{01}(t-nT_0)dt, \tag{5.7}$$

and $F_{01}[n]$ is simply a flag to indicate whether an impulsive noise burst affects the nth symbol of $x_{01}[n]$,

$$F_{01}[n] = \begin{cases} 1 & \text{if a noise impulse occurs in } D_{eff} \text{ of } \phi_{01}(t-nT_0)\,, \\ 0 & \text{if no noise impulse occurs in } D_{eff} \text{ of } \phi_{01}(t-nT_0)\,. \end{cases} \tag{5.8}$$

For each noise burst, there exist several n which satisfy $F_{01}[n]=1$. Instead of treating each $F_{01}[n]=1$, we will consider the effect of one impulse noise burst collectively. In Eq. 5.5, $\nu_{01}[n]$ is the linear combination of a random Gaussian process $\nu_I(t)$, and therefore is still a Gaussian random variable with zero mean and variance N_0.

For notational convenience, we use the following normalization:

$$\begin{aligned}
\overline{h}_p(t) &= T_0 h_p(tT_0), \\
\overline{\phi}_{01}(t) &= T_0\phi_{01}(tT_0), \\
\tau_i &= \tau_i'/T_0.
\end{aligned}$$

Thus, Eq. (5.7) can be written as

$$c[n,\tau_i] = \frac{1}{T_0}\int_n^{N-1+n} \overline{h}_p(t-\tau_i)\overline{\phi}_{01}(t-n)\,dt,$$

with τ_i being uniformly distributed over $[0, 1)$,

$$G_\tau(\tau_i) = \begin{cases} 1 & k \le \tau_i \le (k+1) \\ 0 & \text{otherwise} \end{cases}$$

Furthermore, we normalize a_i and $\nu_{01}[n]$ with respect to E_0. Let

$$\overline{a}_i = \frac{a_i}{\sqrt{E_0}T_0},$$

and

$$\overline{\nu}_{\ell m}[n] = \frac{\nu_{\ell m}[n]}{\sqrt{E_0}}.$$

Using the normalizations, we re-write Eq. (5.5) as

$$\hat{x}_{01}[n] = x_{01}[n] + \overline{\nu}_{01}[n] + F_{01}[n]\overline{a}_i \cos\theta_i c[n, \tau_i].$$

An estimate of $x_{\ell m}[n]$ at terminal (ℓm) is obtained from $\hat{x}_{01}[n]$ by passing it through the decomposition filters; i.e.,

$$\begin{aligned} \hat{x}_{\ell m}[n] &= \sum_k f_{\ell m}[k - 2^\ell n]\hat{x}_{01}[k] \\ &= x_{\ell m}[n] + \overline{\nu}_{\ell m}[n] + F_{\ell m}[n]\overline{a}_i \cos\theta_i \, q_{\ell m}[n, \tau_i], \end{aligned} \tag{5.9}$$

where

$$\begin{aligned} \overline{\nu}_{\ell m}[n] &= \sum_k f_{\ell m}[k - 2^\ell n]\overline{\nu}_{01}[k], \\ q_{\ell m}[n, \tau_i] &= T_0 \sum_k f_{\ell m}[k - 2^\ell n]c[k, \tau_i], \end{aligned} \tag{5.10}$$

and $F_{\ell m}[n]$ is a flag to signal the effects of $F_{01}[n]$ and may be defined as

$$F_{\ell m}[n] = \begin{cases} 1 & \text{if } \sum_k f_{\ell m}[k - 2^\ell n]F_{01}[k] \ne 0 \\ 0 & \text{otherwise.} \end{cases} \tag{5.11}$$

In Eq. (5.10), $q_{\ell m}[n, \tau_i]$ should be considered from a short sequence $c[k, \tau_i]$ which is generated from one single impulse noise burst.

To analyze the probability of error at the (ℓm)th terminal, the joint effect of the Gaussian noise and the impulsive noise is examined. Ignoring the probability of having more than one impulsive bursts in the duration T_0, we can consider the effect under two separate cases of having no impulsive burst and having one impulsive burst, the probability of which are designated P_0 and P_1, respectively. From the point of view of each $x_{\ell m}[n]$, the probability of error can be approximated by P_0 multiplied by the probability of error due to the Gaussian noise alone plus P_1 multiplied by the probability of error due to Gaussian noise and the impulse burst [31, 32]. On the other hand, from the point of view of noise burst, one impulsive burst generates several $q_{\ell m}[n, \tau]$, which take different values and affect several $x_{\ell m}[n]$ in different ways. Because the occurrence of one impulse burst is independent of time n, if we take the short sequence $q_{\ell m}[n, \tau]$ as a whole, the effect of impulse noise is independent of n. In other words, the noise at terminal (ℓm) is stationary, and the probability of error is

$$P_{\ell m}(e) = P_0 \int_y P_{G_{\overline{v}}}(e|y) G_{\overline{v}}(y)\, dy + P_1 \sum_k \int R_{\ell m}^{[k]}(\overline{a}_i) G_{\overline{a}}(\overline{a}_i)\, d\overline{a}_i, \tag{5.12}$$

where $G_{\overline{v}}(y)$ is the probability density function (pdf) of Gaussian noise $\overline{v}_{\ell m}[n]$, $G_{\overline{a}}(\overline{a}_i)$ is the pdf of $\overline{a}_i$, $P_{G_{\overline{v}}}(e|y)$ and $R_{\ell m}^{[k]}(\overline{a}_i)$ denote the probability of error when y and $\overline{a}_i$ are given, respectively. The summation with respect to k is taken over the length of $q_{\ell m}[k, \tau]$ resulting from one impulse noise burst. The relationship between $R_{\ell m}^{[k]}(\overline{a}_i)$ and $q_{\ell m}[k, \tau]$ will be clear later. We now analyze each term in Eq. (5.12).

The impulse occurrence times τ_i' have been assumed to be a result of a Poisson point process. For a positive integer k, if P_k denotes the probability of k points occurring in an interval of length T_0, then [72]

$$P_k = e^{-\nu_a T_0} \frac{(\nu_a T_0)^k}{k!}.$$

By the assumption 3 that the average arrival rate ν_a is slow with respect to T_0, we have

$$P_k = \frac{(\nu_a T_0)^k}{k!} + O\left((\nu_a T_0)^{k+1}\right).$$

Thus $P_1 = \nu_a T_0$, and $P_0 = 1 - \nu_a T_0$ when the arrival rate is slow.

Since a wavelet packet decomposition is a linear and orthogonal decomposition, and each filter satisfies the normality constraints, $\overline{\nu}_{\ell m}[n]$ is still a Gaussian random variable with zero mean and variance N_0/E_0; i.e.,

$$G_{\overline{\nu}}(y) = \frac{1}{\sqrt{2\pi N_0/E_0}} \exp\left(-\frac{y^2}{2N_0/E_0}\right).$$

If binary transmission is assumed for each user, $x_{\ell m}[n] = \pm 1$, the probability of error due to the Gaussian noise alone is

$$\begin{aligned} \int_y P_{G_{\overline{\nu}}}(e|y)G(y)\,dy &= \int_{-\infty}^{-1} G_{\overline{\nu}}(y)\,dy \\ &= \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{E_0}{2N_0}}\right), \end{aligned} \tag{5.13}$$

where $\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty \exp(-z^2)dz$ is the complementary error function.

In Eq. (5.12), $R_{\ell m}^{[k]}(\overline{a}_i)$ is the probability of error conditioned on $\overline{a}_i$. By making use of the assumptions 1 and 2, we have

$$\begin{aligned} R_{\ell m}^{[k]}(\overline{a}_i) &= \int\int\int P_{\ell m}(e|\overline{a}_i, \tau_i, \theta, y) G_{\tau,\theta,\overline{\nu}}(\tau_i, \theta_i, y) d\tau_i\, d\theta_i\, dy \\ &= \int\int\int P_{\ell m}(e|\overline{a}_i, \tau_i, \theta, y) G_{\tau}(\tau_i) G_{\theta}(\theta_i), G_{\overline{\nu}}(y) d\tau_i\, d\theta_i\, dy \\ &= \frac{1}{2\pi}\int_0^1\int_0^{2\pi}\int_y P_{\ell m}(e|\overline{a}_i, \tau_i, \theta_i, y) G_{\overline{\nu}}(y)\,dy\,d\theta_i\,d\tau_i. \end{aligned} \tag{5.14}$$

where $G_{\tau,\theta,\overline{\nu}}(\tau_i, \theta_i, y)$ is the joint probability density function of the random variables.

For notational simplicity, we omit the subscript i in $\overline{a}_i, \tau_i$, and θ_i in the following analysis.

Since $\overline{\nu}_{\ell m}[k]$ and $\overline{a}\cos\theta\, q_{\ell m}[k,\tau]$ ($\theta \in [0, 2\pi)$) are symmetric about 0, we can suppose that $x_{\ell m}[k] = 1$ without loss of generality. For $x_{\ell m}[k] = 1$, an error occurs if the combined impulsive and Gaussian noise is less than -1; i.e.,

$$P_{\ell m}(e|\overline{a}, \tau, \theta, y) = \begin{cases} 1 & \text{if } \ 1 + \nu_{\ell m}[k] + \overline{a}\cos\theta\, q_{\ell m}[k,\tau] < 0 \\ 0 & \text{if } \ 1 + \nu_{\ell m}[k] + \overline{a}\cos\theta\, q_{\ell m}[k,\tau] > 0 \end{cases} \tag{5.15}$$

Denoting the innermost integral in Eq. (5.14) with respect to y by I_y, we have

$$\begin{aligned} I_y &= \int_{-\infty}^{\infty} P_{\ell m}(e|\overline{a}, \tau, \theta, y) G_{\overline{\nu}}(y) dy \\ &= \int_{-\infty}^{-1-\overline{a}q_{\ell m}[k,\tau]\cos\theta} G_{\overline{\nu}}(y) dy \\ &= \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1 + \overline{a}q_{\ell m}[k,\tau]\cos\theta)\right). \end{aligned} \tag{5.16}$$

We represent the integral in Eq. (5.14) with respect to θ by I_θ, i.e.,

$$\begin{aligned} I_\theta &= \int_0^{2\pi} I_y d\theta \\ &= \int_0^{2\pi} \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1 + \overline{a}q_{\ell m}[k,\tau]\cos\theta)\right) d\theta \\ &= \int_0^{\pi} \text{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1 + \overline{a}q_{\ell m}[k,\tau]\cos\theta)\right) d\theta. \end{aligned}$$

Let $z = \cos\theta$ so that

$$I_\theta = \int_{-1}^{1} \frac{1}{\sqrt{1-z^2}}\text{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1 + \overline{a}q_{\ell m}[k,\tau]z)\right) dz. \tag{5.17}$$

Integrating Eq. (5.17) by parts, we have

$$\begin{aligned} I_\theta &= \pi\text{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1 - \overline{a}q_{\ell m}[k,\tau])\right) \\ &\quad -2\int_{-1}^{1} \frac{1}{\sqrt{2\pi N_0/E_0}}\overline{a}q_{\ell m}[k,\tau]\arccos(z)\exp\left(-\frac{(1+\overline{a}q_{\ell m}[k,\tau]z)^2}{2N_0/E_0}\right) dz. \end{aligned} \tag{5.18}$$

Substituting Eq. (5.18) into Eq. (5.14), we obtain

$$\begin{aligned} R_{\ell m}^{[k]}(\overline{a}) &= \frac{1}{2}\int_0^1 \text{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1 - \overline{a}q_{\ell m}[k,\tau])\right) d\tau \\ &\quad -\frac{1}{\pi}\int_0^1 d\tau \int_{-1}^{1} \frac{1}{\sqrt{2\pi N_0/E_0}}\overline{a}q_{\ell m}[k,\tau]\arccos(z)\exp\left(-\frac{(1+\overline{a}q_{\ell m}[k,\tau]z)^2}{2N_0/E_0}\right) dz \end{aligned} \tag{5.19}$$

The relationship between $R_{\ell m}^{[k]}(\overline{a})$ and $q_{\ell m}[k,\tau]$ is established.

Now, let us deal with the integration with respect to τ. From Eq. (5.10), $q_{\ell m}[n,\tau] = T_0 \sum_k f_{\ell m}[k-2^{\ell}n]c[k,\tau]$, while $c[k,\tau]$ is given by Eq. (5.7); i.e.,

$$c[n,\tau] = \frac{1}{T_0}\int_n^{N-1+n} \overline{h}_p(t-\tau)\overline{\phi}_{01}(t-n)dt.$$

In order to analytically evaluate the integral in the right hand of Eq. (5.7), simple forms of $\overline{h}_p(t)$ are often used [29–32]. However, observing that the scaling function $\phi_{01}(t)$ is often defined implicitly, and that any closed-form expressions can be quite complicated, the numerical form for $\phi_{01}(t)$ is used instead. Therefore in Eq. (5.7), even if the simple form of $\overline{h}_p(t)$ is assumed, the integral is unlikely to give an analytical solution. Here, we seek a numerical evaluation of Eq. (5.19) using numerical values of $\phi_{01}(t)$.

The integration with respect to τ over $[0,1)$ can be approximated by its sum with τ represented by a set of discrete time instants $\{0,\tau_s,2\tau_s,\ldots,(N_\tau-1)\tau_s\}$, when τ_s is sufficiently small. Thus, Eq. (5.19) becomes

$$\begin{aligned} R_{\ell m}^{[k]}(\overline{a}) &= \frac{1}{2N_\tau}\sum_{i=0}^{N_\tau-1}\operatorname{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1-\overline{a}q_{\ell m}[k,i\tau_s])\right) \\ &\quad -\frac{1}{\pi N_\tau}\sum_{i=0}^{N_\tau-1}\int_{-1}^{1}\frac{1}{\sqrt{2\pi}}\sqrt{\frac{E_0}{2N_0}}\overline{a}q_{\ell m}[k,i\tau_s]\arccos(z)\exp\left(-\frac{(1+\overline{a}q_{\ell m}[k,i\tau_s]z)^2}{2N_0/E_0}\right)dz \end{aligned} \tag{5.20}$$

Substituting Eqs. (5.20) and (5.13) into (5.12), we have

$$\begin{aligned} P_{\ell m}(e) &= \frac{1}{2}(1-\nu_a T_0)\operatorname{erfc}\left(\sqrt{\frac{E_0}{2N_0}}\right) \\ &\quad +\frac{\nu_a T_0}{2N_\tau}\sum_k\sum_{i=0}^{N_\tau-1}\int_0^\infty \operatorname{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1-\overline{a}q_{\ell m}[k,i\tau_s])\right)G_{\overline{a}}(\overline{a})d\overline{a} \\ &\quad -\frac{\nu_a T_0}{\pi N_\tau}\sum_k\sum_{i=0}^{N_\tau-1}\int_0^\infty G_{\overline{a}}(\overline{a}) \\ &\quad \left[\int_{-1}^{1}\frac{1}{\sqrt{2\pi N_0/E_0}}\overline{a}q_{\ell m}[k,i\tau_s]\arccos(z)\exp\left(-\frac{(1+\overline{a}q_{\ell m}[k,i\tau_s]z)^2}{2N_0/E_0}\right)dz\right]d\overline{a} \end{aligned} \tag{5.21}$$

In order to calculate the probability of error in the presence of impulsive noise for certain user in a WPDM system, we need only to compute Eq. (5.21).

The probability of error due to impulse noise and Gaussian noise at different terminals may be different, since it is a function of $q_{\ell m}[n, i\tau_s]$, which is calculated from $c[n, i\tau_s]$ by using different $f_{\ell m}[n]$ filters. To compare the overall performance of various multiplexing schemes, we define an average probability of error for WPDM as a whole. Assuming the usage of channel at the root terminal (01) being 1, the usage of the channel at a (ℓm) terminal is $1/2^{\ell}$. Thus, the average probability of error is

$$\overline{P}(e) = \sum_{\ell m} \frac{1}{2^{\ell}} P_{\ell m}(e), \tag{5.22}$$

where (ℓm) represents each and every terminal which is used to carry the signals.

It is easily shown that an expression similar to Eq. (5.21) can be derived to evaluate the probability of error of a TDM system with PAM being low-pass equivalent signals. (In later examples, for each user, BPSK is assumed to be the modulation scheme, of which the low-pass equivalent is a bi-polar PAM.) In TDM, the received signal $\hat{x}_{01}[n]$ does not go through further filtering, and the decisions are made solely on $\hat{x}_{01}[n]$. Thus, in a TDM scheme which provides the same bit rates as a given WPDM scheme, assuming that the coding waveform $\phi_{tdm}(t)$ for each user in TDM is a unit energy rectangular function of duration T_0

$$\phi_{tdm}(t) = \begin{cases} \frac{1}{\sqrt{T_0}}, & kT_0 \leq t \leq (k+1)T_0 \\ 0, & \text{otherwise} \end{cases}$$

we have

$$q[n, \tau] = c[n, \tau] = \int h_p(t-\tau)\phi_{tdm}(t - nT_0)dt = \frac{1}{\sqrt{T_0}} \int_{nT_0}^{(n+1)T_0} h_p(t-\tau)dt.$$

We can see that one impulse burst only affects one symbol in TDM, but affects several symbols of different users in WPDM.

Analogous to the derivation for WPDM, the expression for TDM can be derived as

$$\begin{aligned} P_{tdm}(e) &= \frac{1}{2}(1-\nu_a T_0)\mathrm{erfc}\left(\sqrt{\frac{E_0}{2N_0}}\right) \\ &+\frac{\nu_a T_0}{2N_\tau}\sum_{i=0}^{N_\tau-1}\int_0^\infty \mathrm{erfc}\left(\sqrt{\frac{E_0}{2N_0}}(1-\overline{a}q[n,i\tau_s])\right)G_{\overline{a}}(\overline{a})d\overline{a} \\ &-\frac{\nu_a T_0}{\pi N_\tau}\sum_{i=0}^{N_\tau-1}\int_0^\infty G_{\overline{a}}(\overline{a}) \\ &\left[\int_{-1}^{1}\frac{1}{\sqrt{2\pi N_0/E_0}}\overline{a}q[n,i\tau_s]\arccos(z)\exp\left(-\frac{(1+\overline{a}q[n,i\tau_s]z)^2}{2N_0/E_0}\right)dz\right]d\overline{a} \end{aligned} \tag{5.23}$$

Furthermore, since the performance of each user in a TDM is uniform, the average probability of error for a TDM system is the same as the probability of error of a single user.

5.4 Numerical results

In this section, we evaluate the performance of both WPDM and TDM systems in several impulsive noise scenarios.

Following [29–32], it is assumed that the amplitude of the impulsive noise $\overline{a}$, has a lognormal distribution with the pdf

$$G_{\overline{a}}(\overline{a}) = \frac{1}{\sigma_{\overline{a}}(2\pi)^{1/2}\overline{a}}\exp\left(-\frac{(\log\overline{a}-\mu)^2}{2\sigma_{\overline{a}}^2}\right). \tag{5.24}$$

The skewness parameter B of the lognormal distribution is defined by

$$B = 20\log_{10}\left(\sqrt{E\{\overline{a}^2\}/E\{\overline{a}\}^2}\right) = 4.343\sigma_{\overline{a}}^2. \tag{5.25}$$

As in [29–32], we define the signal-to-impulsive noise ratio ρ as the signal power to the impulsive noise power in the information bandwidth at the root terminal $\{01\}$; i.e.,

$$\rho = \frac{E_0^2/2}{\nu_a E\{a^2\}/(2T_0)}$$

$$= \frac{1}{\nu_a T_0 E\{\overline{a}^2\}}. \tag{5.26}$$

The pdf of a lognormal distributed random variable is uniquely determined when ρ and B are given. Given ρ, the noise power is

$$E\{\overline{a}^2\} = \frac{1}{\rho \nu_a T_0}.$$

Given B, $\sigma_{\overline{a}}^2$ is known

$$\sigma_{\overline{a}}^2 = \frac{B}{4.343},$$

and $E\{\overline{a}\}$ can be calculated from Eq. (5.25); i.e.,

$$\begin{aligned} E\{\overline{a}\} &= 10^{-B/20}\sqrt{E\{\overline{a}^2\}} \\ &= 10^{-B/20}\frac{1}{\sqrt{\rho \nu_a T_0}}. \end{aligned}$$

In lognormal distribution,

$$E\{\overline{a}\} = \exp\left(\mu + \frac{\sigma_{\overline{a}}^2}{2}\right).$$

Thus, μ in Eq. (5.24) is calculated as

$$\begin{aligned} \mu &= \log\left(E\{\overline{a}\}\right) - \frac{\sigma_{\overline{a}}^2}{2} \\ &= -\left(\log 10/20 + 1/(4.343 * 2)\right) B - \log(\rho \nu_a T_0). \end{aligned}$$

The probability of error is calculated as a function of ρ for some combinations of the parameters $\nu_a T_0$, SNR, and the skewness parameter B.

Although in our approximation method $\overline{h}_p(t)$, the low-pass equivalent of the impulse response of the RF filter, does not have to have a simple form, for simplicity, we choose $\overline{h}_p(t)$ as

$$\overline{h}_p(t) = \frac{1}{C}\exp\left(-\frac{t}{C}\right), \tag{5.27}$$

where $1/C$ is the 3dB bandwidth of the BPF. This choice was also made in [29–32].

In the first two examples, the parameters are: $\nu_a T_0 = 10^{-2}$, $E_0/N_0 = 13$ dB, $B = 1$, and $1/C = 20$. In the third example, $B = 10$ and the other parameters remain the same as in the first two examples to see the effect of skewness factor B.

Example 5.4.1 Consider a two user case. In WPDM system, the first level functions $\{\phi_{11}(t), \phi_{12}(t)\}$ of a Daubechies wavelet packet are used to modulate information sequences of two users each with transmission rate of $1/2T_0$. At the root terminal (01), the overall transmission rate is $1/T_0$, which is kept as a constant in all the examples. Arbitrarily, the length of $h[n]$ is chosen to be $N = 14$; i.e.,the support length of $\phi_{01}(t)$ is $[0, 13T_0]$.

With $P_{11}(e)$ and $P_{12}(e)$ computed from Eq. (5.21), the average probability of error is

$$P_e = P_{11}(e)/2 + P_{12}(e)/2, \tag{5.28}$$

because each user has the same transmission rate and uses the channel with equal probability of 1/2. The results of the numerical calculation are shown in Fig. 5.2. The corresponding simulation results are shown on the same graph.

In order to compare the performance of different multiplexing systems, the probability of error of a TDM system shared by two users is also computed and simulated. The overall transmission rate is $1/T_0$. The transmission rate for each user in the TDM system is $1/(2T_0)$ and the bit duration is T_0. In the system, BPSK modulation is assumed. The results are included in Fig. 5.2.

Example 5.4.2 In this example, we use the second level functions $\{\phi_{2i}(t), i = 1, 2, 3, 4\}$ in the same wavelet packet tree used in Example 5.4.1 to modulate the data sequences of four users. Similarly, with $P_{21}(e)$, $P_{22}(e)$, $P_{23}(e)$, and $P_{24}(e)$ calculated from Eq. (5.21), the average probability of error is

$$P_e = P_{21}(e)/4 + P_{22}(e)/4 + P_{23}(e)/4 + P_{24}(e)/4, \tag{5.29}$$

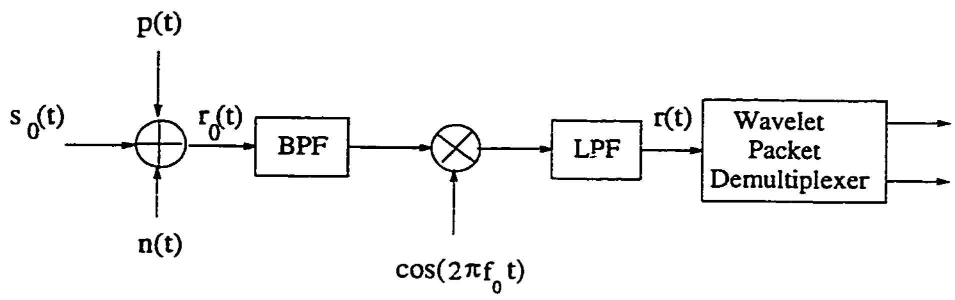


Figure 5.1: Receiver model in impulsive noise channels

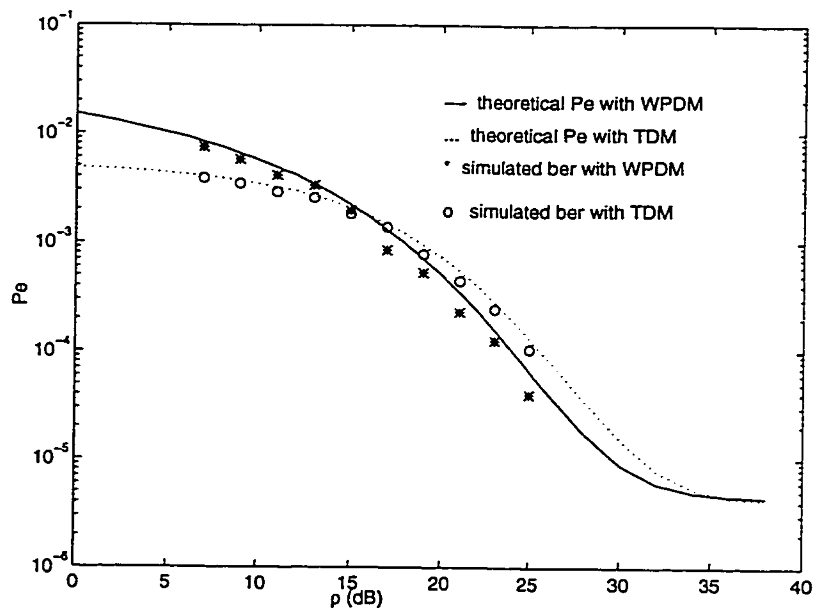


Figure 5.2: Theoretical and simulated probability of error of the WPDM and TDM systems. (The skewness factor $B = 1$.)

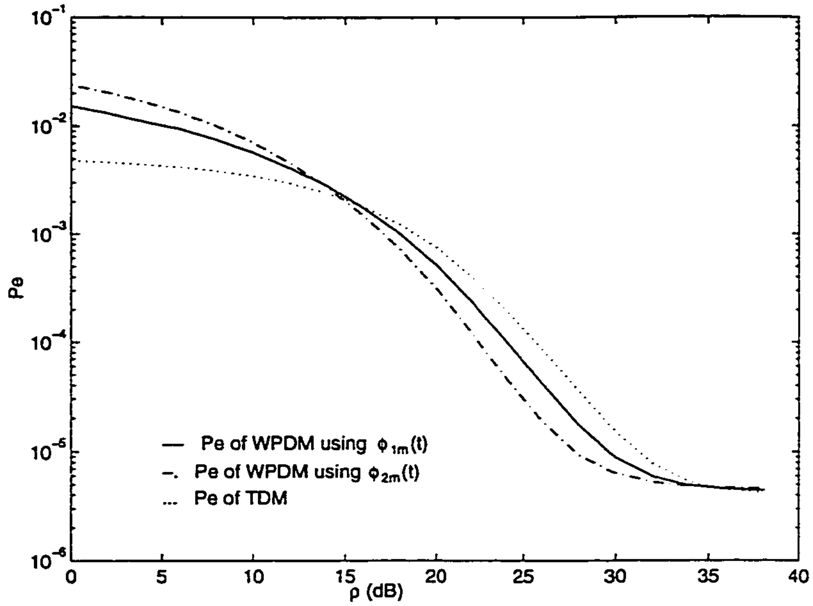


Figure 5.3: Probability of error of the TDM system and WPDM systems with different wavelet packet functions.

where the coefficients 1/4 result from the equal usage of the channel by four users. Fig. 5.3 shows the calculation results for the WPDM system.

In order for a convenient comparison, the calculated P_e curves in Example 5.4.1 are also plotted in Fig. 5.3. For the TDM system used for comparison, since the overall rate is $1/T_0$, each user has a transmission rate $1/4T_0$.

It can be seen from Fig. 5.2 that our theoretical derivation agrees with the simulation results in both WPDM and TDM systems. When the signal-to-impulsive noise ratio is very high, e.g., $\rho \geq 35$dB, the system performance is dominated by AWGN. That is why a floor effect is observed on the P_e vs. ρ graphs. When ρ is moderate, e.g., 15dB $\leq \rho < 35$dB, the WPDM system performs better than the TDM system. This is expected, since an impulsive noise burst, which is strong enough to destroy one bit in a TDM system, is spread over all the channels in WPDM system. For each of the channels, the spread noise burst may not be strong enough to cause an error. On the other hand, when ρ is low, the noise burst is strong and likely to destroy one bit in TDM system. In WPDM system, the strong noise burst, though spreading over all the channels, may be still strong to cause errors for every channel. In practical situations, an impulse burst cannot be very strong, since the dynamic range of the amplifiers at the front end of the receiver cannot be infinite. Thus in most cases, WPDM outperforms TDM when impulsive noise is in concern.

In a WPDM system in which more users share the channel, i.e., users employing $\phi_{\ell m}(t)$ with larger ℓ, the system should provide greater immunity to the impulsive noise, since a noise burst is spread over more channels. This can be validated by the results of Example 5.4.2 shown in Fig. 5.3. From Fig. 5.3, it can be observed that for both WPDM systems in Example 5.4.1 and Example 5.4.2, there is a range of ρ in which WPDM system behaves better than TDM. The range in Example 5.4.2 is wider than that in Example 5.4.1, and at the same signal to impulsive noise ratio, ρ, the probability of error of the WPDM system in Example 5.4.2 is lower than that in Example 5.4.1. By extending the above results, we can expect a wider range of ρ, in which the performance of WPDM is better than TDM, when a WPDM system with a large number of users is employed.

To see the effect of skewness factor B, we present one more example, in which the parameters are $\nu_a T_0 = 10^{-2}$, $E_0/N_0 = 13$ dB, $B = 10$, and $1/C = 20$.

Example 5.4.3 The system setup is exactly same as in Example 5.4.1; i.e.,, two users are sharing the channel, and the wavelet used in WPDM is the Daubechies wavelet with $N = 14$. The average probability of error in impulsive and Gaussian noise channel is calculated from Eq. (5.21) and Eq. (5.28) and the evaluation result is shown in Fig. 5.4.

The calculated result on the probability of error for a TDM system is also included in Fig. 5.4. The TDM system is as in Example 5.4.1.

It can be observed from Fig. 5.4 and Fig. 5.3 that when ρ is small, the bigger B corresponds to better performance for both WPDM and TDM; when ρ is big, the bigger B

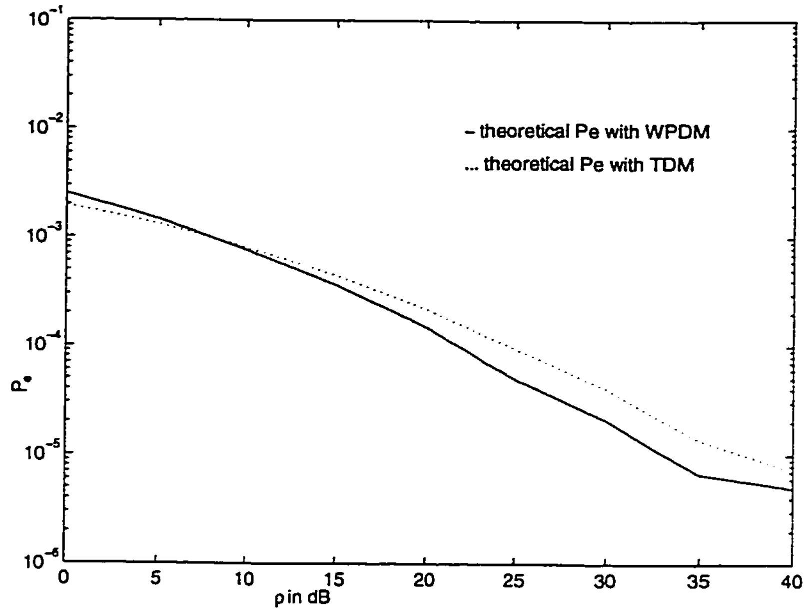


Figure 5.4: Probability of error of the WPDM system and TDM systems. (The skewness factor is $B = 10$.)

relates to higher probability of error and the systems approach the error floor slowly. The observation is directly related with the property of a lognormal distribution. As shown in Fig. 5.5, the impulse noise burst with the bigger skewness factor has higher probability to have week noise bursts, which are not as destructive as strong noise bursts.

Let us assume when the amplitude of a noise burst greater than a_w, the burst may cause error. (a_w should be a big number, for instance, $a_w = 2$.) For a small ρ (e.g. $\rho = 10$dB), we can see that the probability of the amplitude greater than a_w for a smaller B is much higher than that for a bigger B. Thus, both WPDM and TDM performs better in the impulsive noise channel with bigger B than in the similar channel with smaller B for a small ρ. On the other hand for a given big ρ (e.g. $\rho = 30$dB), though the noise burst more concentrates on smaller amplitude, the probability of the amplitude greater than a_w for bigger B is higher that that for a smaller B (refer to Fig. 5.5(c)). Therefore we have our observation that for a high signal to impulsive noise ratio ρ, bigger B corresponds to higher probability of error.

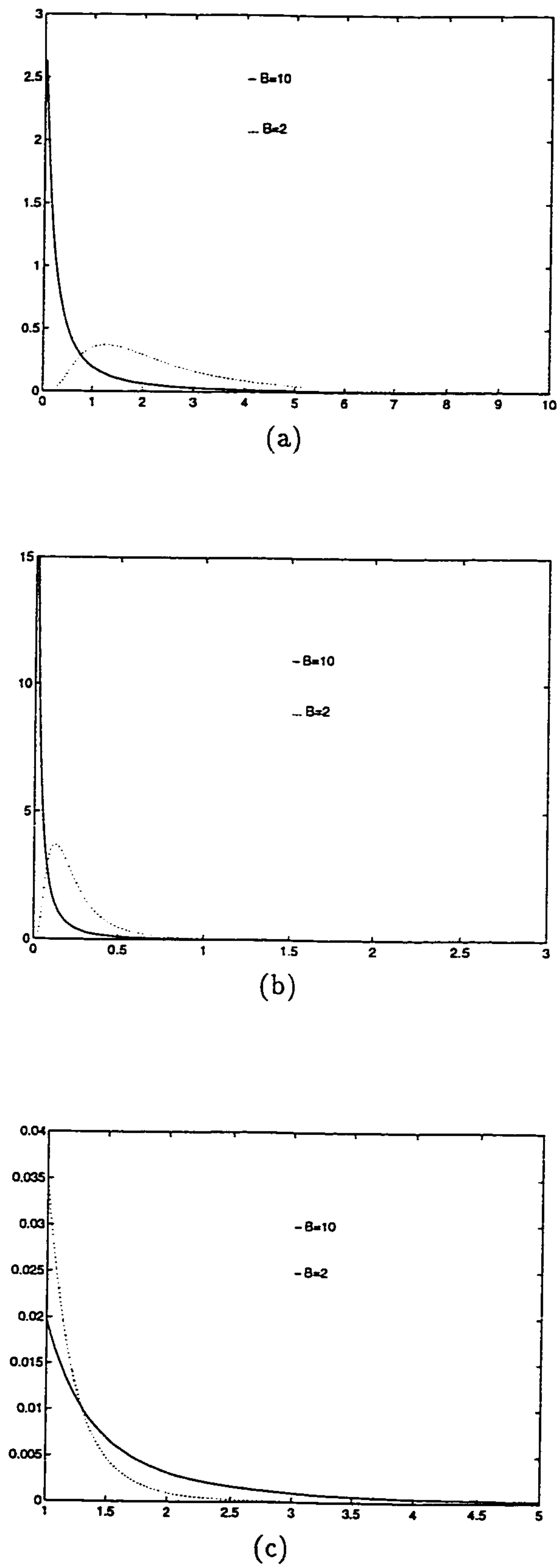


Figure 5.5: The pdfs of lognormal distribution. Assume $\nu_a T_0 = 0.01$. In the figure, B is the skewness factor. (a) $\rho = 10$dB. (b) $\rho = 30$dB. (c) Enlarged version of (b).

Furthermore, comparison of Fig. 5.4 and Fig. 5.3 reveals that bigger skewness factor corresponds to a wider range of ρ in which WPDM performance better than TDM; for example, for $B = 10$, the range ρ is from 7dB extending to beyond $\rho > 40$dB in contrast with the range $15 \leq \rho \leq 35$dB for $B = 1$. The reason for this still lies in the relationship between skewness factor B and the pdf of a lognormal distribution. In scenarios in which B is high (e.g. $B = 10$), for a given ρ there is an increased probability of noise bursts with small amplitudes, to which the WPDM schemes are more robust than the TDM schemes.

5.5 Summary

The performance of WPDM system in the presence of both Gaussian and impulse noise has been analyzed. While the algebraic analysis makes significant progress [29–32], the lack of simple closed-form expressions for general wavelet packet basis functions makes further algebraic analysis difficult. Progress has been made by using numerical approximate techniques. As illustrated in the examples, the calculation of the approximations closely matches simulated performance. This makes our derivation more widely applicable than the approach in [29–32] in the sense that the simple form assumption of $h_p(t)$ is not necessary. The performance of WPDM system has been compared with TDM system and it has been found that WPDM provides greater immunity to impulsive noise of moderate amplitude than TDM, especially when the functions $\phi_{\ell m}(t)$ with larger ℓ are used in the WPDM system. The intuitively pleasing result is due to the fact that in a WPDM scheme, an individual noise burst is spread over subchannels. A noise burst which is strong enough to cause an error in one bit of TDM scheme may be sufficiently spread in the WPDM scheme so as not to cause an error. The result also substantiates that a parallel transmission system is more robust than a serial transmission system in impulsive noise channel.

Chapter 6

WPDM Performance in Fading Channels

6.1 Introduction

In mobile and indoor wireless communication channels, there may be more than one path from transmitter to receiver [25,73]. Such multiple paths (multipaths) arise from multiple scattering by buildings and other structures in the vicinity of a receiver. Multipath scattering may contribute constructively or destructively to the received signal, thus producing random amplitude and phase variations in the received signal as the vehicle moves in the multipath field. The variation of the amplitude in the received signal is usually termed *fading*. Experimentally, we transmit an extremely short pulse (ideally an impulse) over a fading channel at time instants t_1 and t_2, respectively. The received signals at $t_1 + t_d$ and $t_2 + t_d$ (t_d is an arbitrary delay) would have different strengths. Thus unlike in an AWGN channel, the signal at $t_2 + t_d$ cannot be predicted even if we already know the signal at $t_1 + t_d$. Hence, the effect of a fading channel is that of temporal de-correlation. The temporal de-correlation effect is called *flat fading*, or time selective fading. Here "flat" means that at certain instant the mobile channel has a constant gain and linear phase response over a bandwidth which is greater than the bandwidth of the transmitted signal, and the distortions to all the frequencies in the signal bandwidth are even.

Depending upon the differential delays over all the paths, there can also be distortion, which is referred to as *frequency selective multipath.* If we transmit two sinusoids having the same strength at frequencies f_1 and f_2 ($f_2 \neq f_1$), the received signal strengths are different; i.e., we cannot predict the signal for one frequency based on the received signal for the other frequency. Thus in contrast to flat fading, the effect of a frequency selective channel is that of frequency de-correlation and the distortion to the signal varies with frequency. When frequency selective fading occurs, the received signal includes multiple versions of the transmitted waveform which are attenuated and delayed in time. Hence the channel induces ISI.

The harsh channels in mobile communications makes the system design very challenging, but the spectrum is scarce and expensive. For instance, in recent spectral auctions in USA, licenses in the $2GH_z$ frequency band cost companies over nine billion dollars. The limited spectrum must be used efficiently to get a reasonable return on the investment, and at the same time the increasing demand for wireless communications has to be satisfied.

It has been shown that WPDM is a band efficient parallel data transmission system. A parallel transmission system may have advantage in fading channels intuitively. In a parallel approach, a subchannel occupies a narrow bandwidth and the symbol duration is much longer than that in a serial approach. Thus in a specific physical channel environment, it may turn out that for a parallel system, such as FDM or WPDM, the channel may be modeled as flat fading, but for a serial transmission system such as TDM, the channel may have to be modeled as frequency selective fading. It is well known that for frequency selective fading, equalizers are usually needed to obtain acceptable performance. Therefore, with a parallel approach, equalizers may be avoided and a certain level of performance may be still obtained. This is one of the motivations to propose WPDM to be used in wireless communications.

However, whether WPDM will be adopted is determined by the performance of WPDM in various mobile communication channels. The evaluation of the performance is the main theme of this chapter.

It is well known that fading produces irreducible error or error floors in many digital transmission systems [50,73–75]. That is, increasing the signal power does not lower the probability of error. In order to suppress the error floors caused by fading, channel sounding (channel measurement) techniques are usually employed to compensate for fading distortion. One of the commonly used channel sounding techniques is Pilot Symbol Assisted Modulation (PSAM) [75–77]. In PSAM, known pilot symbols are inserted periodically to allow the receiver to estimate the time-varying channel gain and phase shift. The estimates of the channel may then be used to compensate the channel distortion to the transmitted signal. The PSAM is a time-domain equalization technique which is effective in combating flat fading. For frequency selective fading, frequency domain equalization, such as inserting pilot tones, may mitigate the fading effects. More specifically, the RAKE receiver [50] can be used to take advantage of the signals in different paths to improve the bit error rate. However, in this thesis, we confine our study to the application of PSAM in frequency selective fast fading channels and examine the extent to which PSAM helps to combat frequency selective fading.

In this chapter, we first develop a statistical characterization of the fading channel, and illustrate the principles of PSAM. Then the performance of WPDM system is derived over both flat fading and frequency-selective fading channels with PSAM. In time-selective fading, or flat fading channels, the error floor is successfully suppressed with PSAM and a comparison illustrates that WPDM wastes less energy and bandwidth in pilot symbols than FDM. In frequency-selective fading channels, when the time-delay spread is small relative to the bit duration, e.g. less than or equal to 20%, WPDM performs well, although the error floor is not completely suppressed.

6.2 Fading channels

Multipath fading channel is usually modeled by a time variant filter with the impulse response $\alpha(\tau;t)$ [50]. Let us assume the transmitted signal to be

$$s_0(t) = \mathrm{Re}\{s(t)e^{j2\pi f_0 t}\},$$

where $s(t)$ is a baseband signal. Assume the wavelet packet multiplexed signal in Eq. (3.3) is further modulated by a double sideband modulation (DSB), $s(t)$ would be the multiplexed signal.

Then the received bandpass signal may be expressed as

$$y(t) = \int_{-\infty}^{\infty} \alpha(\tau;t)s_0(t-\tau)\,d\tau, \tag{6.1}$$

where $\alpha(\tau;t)$ denotes the random attenuation of the signal components at delay τ and at time instant t. The equivalent low pass impulse response of the channel is

$$u(\tau;t) = \alpha(\tau;t)e^{-j2\pi f_0 \tau}\ ,$$

which is a complex valued random process. From the autocorrelation functions of $u(\tau;t)$, we will see how the fast/slow and frequency flat/selective fading are defined.

First, define the autocorrelation function of $u(\tau;t)$ as

$$\psi_u(\tau_1,\tau_2;t_1,t_2) = \frac{1}{2}E\{u^*(\tau_1;t_1)u(\tau_2;t_2)\}.$$

If $u(\tau;t)$ is assumed to be wide-sense stationary, the autocorrelation is just a function of time difference $\Delta t = t_2 - t_1$. Making use of the fact that in most radio transmission media,

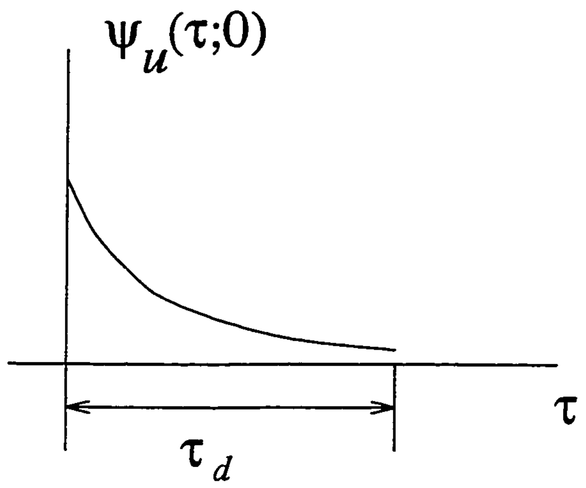


Figure 6.1: One of the typical delay power spectra.

because of the uncorrelated scattering, the attenuation and phase shift at two different delays τ_1 and τ_2 are uncorrelated [50,73], the $\psi_u(.)$ can be obtained as

$$\psi_u(\tau_1, \tau_2; \Delta t) = \psi_u(\tau_1; \Delta t)\delta(\tau_1 - \tau_2). \tag{6.2}$$

If we let $\Delta t = 0$, $\psi_u(\tau; 0)$ is simply the average power output of the channel as a function of the time delay τ. Typically, the function $\psi_u(\tau; 0)$ may appear as shown in Fig. 6.1. The range of values of τ over which $\psi_u(\tau; 0)$ is essentially nonzero is called the *multipath spread* of the channel denoted by τ_d.

In the frequency domain, the time varying channel can be similarly characterized. Taking the Fourier transforms of $u(\tau; t)$ with respect to τ, we obtain the time variant transfer function $U(f; t)$, i.e.,

$$U(f; t) = \int_{-\infty}^{\infty} u(\tau; t) e^{-j2\pi f \tau} \, d\tau. \tag{6.3}$$

Similarly, we define the autocorrelation of $U(f; t)$

$$\psi_U(f_1, f_2; \Delta t) = \frac{1}{2} E\{U^*(f_1; t) U(f_2; t + \Delta t)\}. \tag{6.4}$$

Substitution of Eq. (6.3) into (6.4) and using Eq. (6.2) results in the relationship

$$\psi_U(f_1, f_2; \Delta t) = \int_{-\infty}^{\infty} \psi_u(\tau_1; \Delta t) e^{-j2\pi(f_1 - f_2)\tau_1} \, d\tau_1,$$

or

$$\psi_U(\Delta f; \Delta t) = \int_{-\infty}^{\infty} \psi_u(\tau_1; \Delta t) e^{-j2\pi\Delta f \tau_1}\, d\tau_1.$$

with $\Delta f = f_1 - f_2$. When $\Delta t = 0$, the above relationship becomes

$$\psi_U(\Delta f; 0) = \int_{-\infty}^{\infty} \psi_u(\tau_1; 0) e^{-j2\pi\Delta f \tau_1}\, d\tau_1. \tag{6.5}$$

Since $\psi_U(\Delta f; 0)$ is an autocorrelation function in the frequency variable, it measures the frequency coherence of the channel. As a result of Eq. (6.5) and the property of Fourier transforms, the reciprocal of the maximum multipath spread τ_d is a measure of the coherent bandwidth $(\Delta f)_c$ of the channel. The bigger the maximum multipath spread is, the smaller the coherent bandwidth is.

When $(\Delta f)_c$ is small in comparison to the transmitted signal, the channel is said to be *frequency selective*. On the other hand, if $(\Delta f)_c$ is large in comparison with the bandwidth of the transmitted signal, the channel is said to be frequency-nonselective or flat.

We now focus our attention on the time variations of the channel as measured by the parameter $\psi_U(\Delta f; \Delta t)$. The time variations in the channel are evidenced as a Doppler broadening. Let us consider a situation in Fig. 6.2 when the transmission signal consists of only one carrier wave with frequency f_0. Assume that the wave is coming in at an azimuthal angle θ with respect to the vehicle motion. The time domain signal can be written as

$$r(t) = A_0 \cos\left(2\pi t \left(f_0 - \frac{v f_0}{c} \cos\theta\right)\right), \tag{6.6}$$

where A_0 is a constant amplitude, v is the vehicle speed and c is the speed of light ($c = 3*10^8$ m/s). We will look at the signal in Eq. (6.6) from two points of view.

First, the envelope of the received signal in Eq. (6.6) is considered to be constant while its frequency f_r is given by

$$f_r = f_0 - \frac{v f_0}{c} \cos\theta.$$

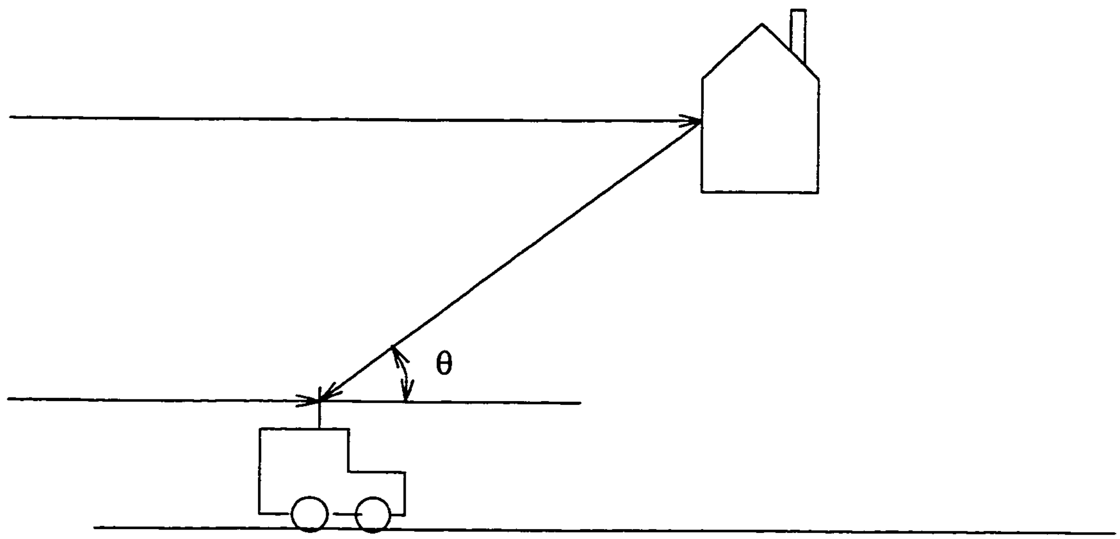


Figure 6.2: Doppler effect in mobile communications.

The received frequency differs from the transmitted frequency by

$$f_d = \frac{v f_0}{c} \cos\theta.$$

This displacement in frequency is called the *Doppler shift.* It can be seen that if the automobile is moving toward the direction of arrival of the wave, the Doppler shift is positive (i.e., the apparent received frequency is increased), and if the mobile is moving away from the direction of arrival of the wave, the Doppler shift is negative (i.e. the apparent received frequency is decreased). Multipath components arriving from different directions contribute to Doppler spread of the received signal, thus increasing the signal bandwidth.

Second, we may express Eq. (6.6) in terms of in-phase $r_I(t)$ and quadrature $r_Q(t)$ components as

$$r(t) = r_I(t)\cos(2\pi f_0 t) + r_Q(t)\sin(2\pi f_0 t)$$

where

$$r_I(t) = A_0 \cos(2\pi f_d t)$$

and

$$r_Q(t) = A_0 \sin(2\pi f_d t).$$

Eq. (6.6) is interpreted as the complex envelope of the received signal varying in Doppler frequency f_d. In other words, the complex envelope signal spectrum is a sinusoid at frequency f_d. Multiple components arriving from various direction cause the complex envelope of the received signal to vary at various frequencies from 0 to maximum Doppler frequency. By the property of Fourier transform, the spectrum of the received signal is the convolution of the spectrum of the complex envelope and the transmitted signal. Therefore, the received signal is broadened by the Doppler effect.

To extend the relationship of the Doppler effects with the time variations of the channel, we define the Fourier transform of $\psi_U(\Delta f; \Delta t)$ with respect to Δt,

$$\Psi_U(\Delta f; \lambda) = \int_{-\infty}^{\infty} \psi_U(\Delta f; \Delta t) e^{-j2\pi\lambda\Delta t} \, d\Delta t.$$

Setting Δf to zero, $\Psi_U(.)$ becomes

$$\Psi_U(0; \lambda) = \int_{-\infty}^{\infty} \psi_U(0; \Delta t) e^{-j2\pi\lambda\Delta t} \, d\Delta t. \tag{6.7}$$

$\Psi_U(0; \lambda)$ gives the signal power intensity at the Doppler frequency λ. If the channel is time-invariant, $\psi_U(0; \Delta t) = 1$ and $\Psi_U(0; \lambda)$ becomes $\delta(\lambda)$. That is, with no time variations in the channel, there is no spectral broadening. The range of values of λ over which $\Psi_U(0; \lambda)$ is essentially nonzero is called the *Doppler spread* f_D. Since $\Psi_U(0; \lambda)$ is related to $\psi_U(0; \Delta t)$ which measures the time coherence of the channel, the reciprocal of f_D is a measure of the time coherence $(\Delta t)_c$ of the channel.

A slowly changing channel has a large coherence time or equivalently a small Doppler spread. Then the channel is called *slow fading channel.* On the other hand, a channel with a big Doppler spread implies that the channel loses time correlation quickly; i.e., the channel changes fast. Such a channel is called *fast fading channel.*

6.2.1 Flat fading channels

As explained in Section 6.2, when an information-bearing signal is transmitted through the channel, if the coherence bandwidth $(\Delta f)_c$ of the channel is large in comparison with the bandwidth of the transmitted signal, the channel is called a flat fading channel. Because of the inverse proportionality relationship between $(\Delta f)_c$ and the multipath spread τ_d, a flat fading channel essentially imply $\tau_d \approx 0$. Referring to Eq. (6.1), the equivalent low-pass received signal after a flat fading channel is

$$r(t) = u(0;t)s(t) + \nu_c(t). \tag{6.8}$$

For notational simplicity, in the context of flat fading channel, we denote $u(0;t) = u(t)$. Thus, the received signal is simply the transmitted signal multiplied by a complex-valued random process, which represents the time-variant characteristics of the channel, corrupted by the AWGN. When $u(t)$ is modeled as a zero-mean complexed-valued Gaussian process, the envelope of the channel response at any instant has a Rayleigh distribution and the phase is uniformly distributed in the interval $(0, 2\pi)$. In this case, the channel is said to be a Rayleigh fading channel. The rapidity of the fading is determined from the correlation function $\psi_U(0;\Delta t)$ or the Doppler spread f_D.

We may generally express $\psi_U(0;\Delta t)$ as

$$\psi_U(0;\Delta t) = \sigma_u^2 \tilde{R}_u(\Delta t),$$

where σ_u^2 is the average signal power in $u(t)$ and $\tilde{R}_u(\Delta t)$ has unit power. Although the analysis is not limited to a specific autocorrelation function, in this chapter, we assume that $u(t)$ is characterized by Jake's power spectrum $\Psi_U(0;\lambda)$ [73]

$$\Psi_U(0;\lambda) = \begin{cases} \frac{\sigma_u^2}{2\pi}\frac{1}{\sqrt{f_D^2-\lambda^2}} & |\lambda| < f_D \\ 0 & |\lambda| > f_D \end{cases} \tag{6.9}$$

with f_D being the maximum Doppler frequency experienced by the moving vehicle. From Eq. (6.7), the autocorrelation function of $u(t)$ is

$$\psi_U(0;\Delta t) = \sigma_u^2 J_0(2\pi f_D \Delta t), \tag{6.10}$$

where $J_0(t)$ is the zeroth order Bessel function of the first kind.

6.2.2 Frequency selective fading channels

When the coherence bandwidth $(\Delta f)_c$ is not large in comparison to the signal bandwidth, i.e., the time delay spread is a significant fraction of the symbol duration, frequency selective fading would characterize the channel effect. We have illustrated that the channel is usually modeled as a time-varying linear filter, and the low-pass equivalent of the output of such a channel is

$$r(t) = \int_0^{\tau_d} u(\tau;t)s(t-\tau)d\tau \tag{6.11}$$

where $u(\tau;t)$ is the channel gain, observed at t, of scatters with delay τ, and τ_d is maximum delay. As in flat fading, we assume $u(\tau;t)$ is a complex-valued Gaussian process in t, so that the channel is a Rayleigh fading channel.

In this thesis, we will consider a special case of frequency selective fading: two-ray multipath fading. Although the model itself is simple, it is widely used to approximate the channel characteristics [77,78]. In a two-ray multipath, the channel impulse response $u(\tau;t)$ is

$$u(\tau;t) = u_1(t)\delta(\tau) + u_2(t)\delta(\tau-\tau_1) \tag{6.12}$$

where $u_1(t)$ and $u_2(t)$ are uncorrelated complex-valued Gaussian processes, and are assumed to have the same kind of power spectral density as described in Eq. (6.9).

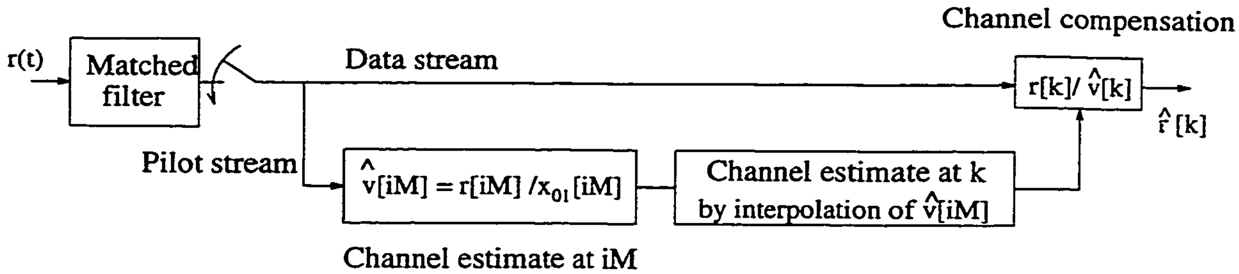


Figure 6.3: Illustration of the principle of PSAM

6.3 Pilot Symbol Assisted Modulation (PSAM)

Pilot Symbol Assisted Modulation (PSAM) is an equalization technique used to estimate channel distortion, in which prespecified symbols are inserted periodically into the information sequence $x_{01}[n]$ prior to modulation. Both pilot symbols and data symbols experience the same fading distortion after they pass through the channels. At the receiver, the received signal is split into two streams. One stream consists of faded pilot symbols, and the other one consists of data symbols. Since the pilot symbols are known, an estimate of the channel distortion on pilot signals can be extracted. If the channel does not change too quickly, the estimate then may be interpolated to form an estimate of the channel distortion on data symbols, so that channel distortion to data stream can be compensated for and better performance can be expected. Fig. 6.3 is the diagram illustrating the principle of PSAM.

We assume every Mth symbol is the prespecified pilot symbol. Without loss of generality, take $x_{01}[iM], i \in \mathcal{Z}$ to be pilot symbols. In order to prevent the transmission of a discrete tone the pilot symbols are usually randomized.

6.3.1 Review of receiver processing with PSAM

Suppose that the equivalent low-pass received signal is expressed as

$$r(t) = v(t)s(t) + \tilde{\nu}(t), \tag{6.13}$$

where $v(t)$ is a low pass complex-valued Gaussian process, $\tilde{\nu}(t)$ is a noise process, and $s(t)$ is the equivalent low-pass of the transmitted signal whose modulating pulse is assumed to be $\phi_{01}(t)$. In the next section, we will relate $v(t)$ and $\tilde{\nu}(t)$ with the timing variant filters and AWGN in flat fading channel and frequency selective fading channel.

Multiplicative fading broadens the spectrum of the transmitted signal, and may cause interference to the adjacent channels. Sufficient guardband will limit such interference. On the other hand, the fading may also destroy the orthogonality in the signal and result in ISI. However, it is usually a reasonable assumption that fading is slow enough not to cause appreciable distortion of the signal pulse. For example, at the carrier frequency 800 MHz, for a vehicle speed of 100km per hour, the Doppler frequency is about 80Hz. When compared with the moderate signal transmission rate 24.3 kbaud/s, which is the air interface rate in the IS-54 and IS-136 digital cellular standards, the complex envelope due to fading is so slow that for several tens of symbols the fading can be considered as constant attenuation. Therefore little ISI resulted. It is stated that 5% Doppler shift results in only 1% maximum ISI in the case of 25% rolloff spectral raised cosine pulses [76]. Under this assumption, a suboptimum reception employing the matched filter could be used.

At the receiver, by assuming the perfect carrier demodulation, the demodulated signal is $r(t)$. Passing $r(t)$ through a filter matched to the transmitted pulse, and sampling the output of the matched filter at exact instants nT_0 (i.e., perfect symbol synchronization has been assumed), we have

$$\begin{aligned} r_v[n] &= \frac{1}{\sqrt{E_0}} \int r(t)\phi_{01}(t - nT_0)\, dt \\ &= v[n]x_{01}[n] + \tilde{\nu}_0[n], \end{aligned} \tag{6.14}$$

where $v[n]$ are samples of the fading process $v(t)$ at $t = nT_0$, which makes use of the fact

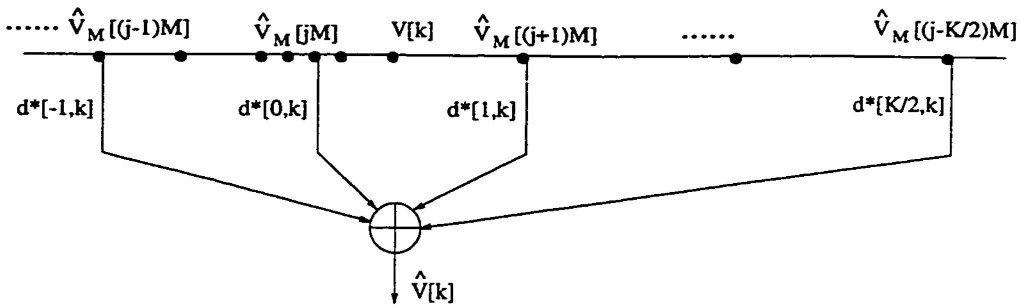


Figure 6.4: Diagram for interpolating the channel estimation

that $v(t)$ is essentially constant in the integration interval, and

$$\tilde{\nu}_0[n] = \frac{1}{\sqrt{E_0}} \int \tilde{\nu}(t)\phi_{01}(t - nT_0)\, dt.$$

To carry out fading estimation at $n = mM$, $r_v[mM]$ is simply divided by the known symbol $x_{01}[mM]$. The fading estimates at $n = mM$ are therefore

$$\begin{aligned} \hat{v}_M[mM] &= \frac{r_v[mM]}{x_{01}[mM]} \\ &= v[mM] + \frac{\tilde{\nu}_0[mM]}{x_{01}[mM]}. \end{aligned} \tag{6.15}$$

If the noise $\tilde{\nu}(t)$ is zero, the channel distortion at mM due to multiplicative fading $v(t)$ is accurately found. With nonzero noise $\tilde{\nu}(t)$, the fading distortion can only be estimated with the noise $\frac{\tilde{\nu}_0[mM]}{x_{01}[mM]}$ as shown in Eq. (6.15). Given $\tilde{\nu}(t)$, bigger $x_{01}[mM]$ would result in smaller estimation errors. Hence, if $x_{01}[n]$ takes multiple amplitude, the pilot samples should take the biggest amplitude to reduce the channel estimate error. However, such a choice requires more signal power to be spent on the pilots.

Since the fading process is band-limited, the fading distortion on the data symbols can be interpolated from the K nearest pilot samples $\hat{v}_M[mM]$. In [76], a method was proposed using a Wiener filter to obtain estimates $\hat{v}[k]$ for $jM - \lfloor M/2 \rfloor \leq k \leq jM + \lfloor M/2 \rfloor$, where j is an integer. As shown in Fig. 6.4, let $d^*[i,k]$ be the coefficients corresponding to $\hat{v}_M[(j+i)M]$ for the estimation at the kth position. Then we can express $\hat{v}[k]$ as

$$\hat{v}[k] = \sum_{i=-\lfloor K/2 \rfloor}^{\lfloor K/2 \rfloor} d^*[i,k]\hat{v}_M[(j+i)M], \tag{6.16}$$

where the symbol $\lfloor x \rfloor$ represents the operation of taking the integer part of x.

Denote

$$\mathbf{d}[k] = \left[d\left[-\frac{K}{2}M, k\right], \cdots\ d[-M,k]\ d[0,k]\ d[M,k]\ \cdots d\left[\frac{K}{2}M, k\right]\right]^T,$$

and

$$\hat{\mathbf{v}}_M = \left[\hat{v}_M\left[jM - \frac{K}{2}M\right], \cdots\ \hat{v}_M[jM-M]\ \hat{v}_M[jM]\ \hat{v}_M[jM+M]\ \cdots \hat{v}_M\left[jM + \frac{K}{2}M\right]\right]^T.$$

Eq. (6.16) can be written as

$$\hat{v}[k] = \mathbf{d}^\dagger[k]\hat{\mathbf{v}}_M,$$

where the symbol $\dagger$ represents the conjugate transpose. The coefficients $d[i,k]$ ($d^*[i,k]$ is the conjugate of $d[i,k]$) can be determined by minimizing the variance of the estimation error

$$e[k] = v[k] - \hat{v}[k].$$

To minimize $E\{e[k]e^*[k]\}$, we differentiate it with respect to $d[i,k]$ and set the derivative to zero. By assuming that the fading process is a wide sense stationary process, an optimum estimate is achieved when the coefficient vector $\mathbf{d}$ satisfies the normal equation

$$\mathbf{Q}\,\mathbf{d}[k] = \mathbf{w}[k],$$

where the elements of $\mathbf{Q}$ and $\mathbf{w}$ are given by

$$\begin{aligned}
Q_{ik} &= \frac{1}{2}E\{\hat{v}[iM]\hat{v}^*[kM]\} \\
&= \sigma_v^2 \tilde{R}_v(\,(iM-kM)T_0\,) + \frac{1}{E\{|x_{01}[iM]|^2\}} R_{\tilde{\nu}_0}((i-k)MT_0) \qquad (6.17) \\
w_i[k] &= \frac{1}{2}E\{v^*[k]\hat{v}[iM]\,\} \\
&= \sigma_v^2 \tilde{R}_v(\,(iM-k)T_0\,). \qquad (6.18)
\end{aligned}$$

Optimum $d[i,k]$ can be calculated from the normal equation

$$\mathbf{d}[k] = \mathbf{Q}^{-1}\mathbf{w}[k],$$

and the variance of interest is

$$\sigma_{\hat{v}}^2[k] = \mathbf{w}^{\dagger}[k]\mathbf{Q}^{-1}\mathbf{w}[k], \tag{6.19}$$

$$\sigma_e^2[k] = \sigma_v^2[k] - \sigma_{\hat{v}}^2[k]. \tag{6.20}$$

For the information symbols, the channel fading compensation is carried out by dividing each received symbol $r_v[k]$ by the corresponding fade estimate $\hat{v}[k]$. Thus, the compensated samples are

$$\hat{r}_v[k] = \frac{r_v[k]}{\hat{v}[k]}. \tag{6.21}$$

The above equalization procedure can be used in many digital communications systems. We will apply PSAM in WPDM when the channel of interest is a fading channel.

6.4 Performance analysis of WPDM in fading channels

6.4.1 Performance of WPDM in flat fading channels with PSAM

First, we analyze the performance of a WPDM system with $\phi_{\ell m}(t)$ being employed by users to transmit binary sequences of ± 1.

The equivalent low pass signal of the output of a flat fading channel is

$$r(t) = u(t)s(t) + \nu_c(t), \tag{6.22}$$

where $u(t)$ is the fading process with a power spectrum given in Eq. (6.9), $\nu_c(t)$ is the lowpass equivalent of channel AWGN $\nu(t)$, and $s(t)$ is the equivalent low pass signal of the transmitted signal. In WPDM,

$$s(t) = \sqrt{E_0}\sum_n x_{01}[n]\phi_{01}(t - nT_0),$$

with $x_{01}[n]$ being reconstructed from the signal sequences $x_{\ell m}[n]$, which is the modulating sequences of $\phi_{\ell m}(t)$.

Comparison of Eq. (6.13) with Eq. (6.22), reveals that

$$v(t) = u(t),$$

and

$$\tilde{\nu}(t) = \nu_c(t).$$

Therefore, the PSAM fading compensation process discussed in last section can be directly used to compensate the distortion caused by the flat fading. The received signal $r(t)$ is first passed through a matched filter matching to $\phi_{01}(t)$ and sampled at the exact instants nT_0. Recalling Fig. 6.3, the resulting signal $r_u[n]$ is split into two streams. From the stream consisting of the distorted pilot signal, the channel distortion at the instants of the pilot symbols, $\hat{u}[kM]$, is extracted and interpolated to form the estimate of the channel distortion on the information symbols, $\hat{u}[k]$. In the other data stream, the channel distortion is compensated by the estimate. As in Eq. (6.21), the compensated samples are

$$\hat{r}_u[k] = \frac{r_u[k]}{\hat{u}[k]} = \frac{u[k]x_{01}[k] + \nu[k]}{\hat{u}[k]}, \tag{6.23}$$

where

$$\nu[n] = \frac{1}{\sqrt{E_0}} \int \nu_c(t)\phi_{01}(t - nT_0)dt. \tag{6.24}$$

By assuming that the spectral density of the AWGN is $N_0/2$, we can calculate the variance of the complex $\nu[n]$ to be $\sigma_\nu^2 = N_0/E_0$.

It has been shown in the last section that the variance of the estimated channel distortion $\hat{u}[k]$ is

$$\sigma_{\hat{u}}^2[k] = \mathbf{w}^\dagger[k]\mathbf{Q}^{-1}\mathbf{w}[k],$$

where $\mathbf{Q}$ and $\mathbf{w}$ are defined in Eq. (6.17) and (6.18) with v replaced by u such as

$$\begin{aligned}
Q_{ik} &= \frac{1}{2}E\{\hat{u}[iM]\hat{u}^*[kM]\} \\
&= \sigma_u^2\bar{R}_u(\,(iM-kM)T_0\,) + \frac{N_0}{E\{|x_{01}[iM]|^2\}\,E_0}\delta[i-k] \\
w_i[k] &= \frac{1}{2}E\{u^*[k]\hat{u}[iM]\,\} \\
&= \sigma_u^2\bar{R}_u(\,(iM-k)T_0\,)\,.
\end{aligned}$$

The variance of the estimation error $e[k]$ ($e[k] = u[k] - \hat{u}[k]$) is as defined in Eq. (6.20); i.e.,

$$\sigma_e^2[k] = \sigma_u^2[k] - \sigma_{\hat{u}}^2[k]. \tag{6.25}$$

Passing the compensated data samples $\hat{r}_u[k]$ through wavelet packet decomposition filters, we obtain the estimated $x_{\ell m}[n]$,

$$\hat{x}_{\ell m}[n] = \sum_k \hat{r}_u[k] f_{\ell m}[k-2^\ell n]. \tag{6.26}$$

Substitution of Eq. (6.23) into (6.26) results in

$$\begin{aligned}
\hat{x}_{\ell m}[n] &= \sum_k \frac{u[k]x_{01}[k]+\nu[k]}{\hat{u}[k]} f_{\ell m}[k-2^\ell n] \\
&= \sum_k \frac{(\hat{u}[k]+e[k])\,x_{01}[k]+\nu[k]}{\hat{u}[k]} f_{\ell m}[k-2^\ell n] \\
&= x_{\ell m}[n] + \sum_k \frac{e[k]x_{01}[k]+\nu[k]}{\hat{u}[k]} f_{\ell m}[k-2^\ell n].
\end{aligned} \tag{6.27}$$

Assume $x_{\ell m}[n]$ are real valued. Then it can be seen that the overall noise of interest to $x_{\ell m}[n]$ is the real part of the second term in Eq. (6.27). Let us denote it as $I_{\ell m}[n]$.

$$I_{\ell m}[n] = \mathrm{Re}\left\{\sum_k \frac{e[k]x_{01}[k]+\nu[k]}{\hat{u}[k]} f_{\ell m}[k-2^\ell n]\right\}.$$

Under the assumption that the fading process $u(t)$ is a Gaussian process, $\hat{u}[n]$ is Gaussian distributed. The channel estimate error $e[n]$ is also Gaussian distributed, since $e[n]$ is a linear combination of Gaussian distributed random variables. If $x_{01}[n] = \pm 1$, $e[n]x_{01}[n]$ has a Gaussian distribution. (The proof of $e[n]x_{01}[n]$ being Gaussian distributed can be found in Appendix A.) However, in many cases, such as multilevel QAM [50] and the WPDM described in the previous chapters, $x_{01}[n]$ takes on multiple amplitudes so that it is difficult to find the pdf of $e[n]x_{01}[n]$ even if it is possible. This in turn complicates the derivation of the pdf of $I_{\ell m}[n]$.

To derive the probability density function (pdf) of $I_{\ell m}[n]$, we make use of the slowly varying property of $\hat{u}[n]$, and express $I_{\ell m}[n]$ as

$$I_{\ell m}[n] \simeq \mathrm{Re}\left\{\frac{\sum_k (e[k]x_{01}[k] + \nu[k])\, f_{\ell m}[k - 2^\ell n]}{\hat{u}[2^\ell n]}\right\}, \tag{6.28}$$

as long as the varying rate of $\hat{u}[n]$ is slow compared with the length of $f_{\ell m}[n]$. In many practical situations when WPDM system is used, this requirement can be easily satisfied.

For notational simplicity, denote

$$\tilde{e}_{\ell m}[n] = \sum_k (e[k]x_{01}[k] + \nu[k])\, f_{\ell m}[k - 2^\ell n].$$

Thus,

$$I_{\ell m}[n] = \mathrm{Re}\left\{\frac{\tilde{e}_{\ell m}[n]}{\hat{u}[2^\ell n]}\right\}.$$

The complex variable $\hat{u}[n]$ can be written as

$$\hat{u}[n] = |\hat{u}[n]| \exp(j\theta_{\hat{u}}[n]) \tag{6.29}$$

The circular symmetry of $\hat{u}[n]$ implies that $\theta_{\hat{u}}[n]$ is uniformly distributed over $[0, 2\pi)$. Substituting Eq. (6.29) into Eq. (6.28), we have

$$\begin{aligned} I_{\ell m}[n] &= \mathrm{Re}\left\{\frac{\tilde{e}_{\ell m}[n]}{|\hat{u}[2^\ell n]| \exp(j\theta_{\hat{u}}[2^\ell n])}\right\} \\ &= \mathrm{Re}\left\{\frac{\tilde{e}_{\ell m}[n] \exp(-j\theta_{\hat{u}}[2^\ell n])}{|\hat{u}[2^\ell n]|}\right\} \end{aligned} \tag{6.30}$$

Because of the circular symmetry of $e[n]$ and $\nu[n]$, $\tilde{e}_{\ell m}[n]$ has circular symmetry; i.e., $\tilde{e}_{\ell m}[n] = |\tilde{e}_{\ell m}[n]| \exp(j\theta_{\tilde{e}}[n])$ with $\theta_{\tilde{e}}[n]$ uniformly distributed over $[0, 2\pi)$. Since $\theta_{\hat{u}}[n]$ and

$\theta_{\tilde{e}}[n]$ are independent, $\tilde{e}_{\ell m}[n]\exp(-j\theta_{\hat{u}}[2^\ell n])$ has the same distribution property as $\tilde{e}_{\ell m}[n]$. Thus $I_{\ell m}[n]$ can be expressed as

$$I_{\ell m}[n] = \frac{\mathrm{Re}\left\{\tilde{e}_{\ell m}[n]\right\}}{|\hat{u}[2^\ell n]|},$$

without change of the distribution property of $I_{\ell m}[n]$.

Let us define

$$e_{\ell m}[n] = \mathrm{Re}\left\{\tilde{e}_{\ell m}[n]\right\},$$

and

$$e_d[n] = \frac{1}{|\hat{u}[2^\ell n]|}. \tag{6.31}$$

Then $I_{\ell m}[n]$ is the product of two independent random variables $e_{\ell m}[n]$ and $e_d[n]$.

It can be seen from Eq. (6.28) that $e_{\ell m}[n]$ is the sum of the K independent random variables, because $\mathrm{Re}\left\{e[n]\right\}$, $x_{01}[n]$, and $\mathrm{Re}\left\{\nu[n]\right\}$ are independent, where K is the length of filter $f_{\ell m}[n]$. Though K is not infinite, we can approximate $e_{\ell m}[n]$ as Gaussian distributed, given that $(e[k]x_{01}[k] + \nu[k])$ has finite variance [72]. Furthermore, the mean and variance of $e_{\ell m}[n]$ are

$$E\left\{e_{\ell m}[n]\right\} = 0$$

and

$$\sigma_{\ell m}^2 = E\left\{e_{\ell m}^2[n]\right\} = \frac{1}{2}E\left\{|\tilde{e}_{\ell m}[n]|^2\right\} = \frac{1}{2}(\sigma_e^2\sigma_x^2 + \sigma_\nu^2)\sum_k f_{\ell m}^2[k - 2^\ell n].$$

The accuracy of the approximation is demonstrated in Appendix C. For a wavelet packet, $\sum_k f_{\ell m}^2[k]$ is a constant for all the terminals (ℓm). Hence, $\sigma_{\ell m}^2$ is not a function of (ℓm), and all the terminals in a WPDM system experience the same performance.

The pdf of $e_{\ell m}[n]$ is

$$p_{e_{\ell m}}(z) = \frac{1}{\sqrt{2\pi\sigma_{\ell m}^2}}\exp\left(-\frac{z^2}{2\sigma_{\ell m}^2}\right).$$

The random variable $e_d[n]$ is a function of $|\hat{u}[n]|$. The pdf of a general function $e_d[n] = Y\left(\left|\hat{u}[2^\ell n]\right|\right)$ is related to the pdf of $|\hat{u}[2^\ell n]|$ [71,72] by

$$p_{e_d}(z) = \frac{p_{|\hat{u}|}\left(Y^{-1}(z)\right)}{|J|} \tag{6.32}$$

where Y^{-1} is the inverse function of Y, i.e. $|\hat{u}[2^\ell n]| = Y^{-1}(e_d[n])$, and J is the first order derivative of $e_d[n]$ with respect to $|\hat{u}[2^\ell n]|$ evaluated at $|\hat{u}[2^\ell n]| = Y^{-1}(e_d[n])$. From Eq. (6.31),

$$J = -z^2. \tag{6.33}$$

Since $\hat{u}[n]$ is a complex Gaussian distributed random variable with zero mean, $|\hat{u}[n]|$ has a Rayleigh distribution; i.e.

$$p_{|\hat{u}|}(z) = \begin{cases} \frac{2z}{\sigma_{\hat{u}}^2} \exp\left(-\frac{z^2}{\sigma_{\hat{u}}^2}\right) & z \geq 0 \\ 0 & z < 0 \end{cases} \tag{6.34}$$

Substituting Eq. (6.34) and (6.33) into (6.32), we obtain the pdf of $e_d[n]$,

$$p_{e_d}(z) = \frac{2}{z^3\sigma_{\hat{u}}^2}\exp\left(-\frac{1}{\sigma_{\hat{u}}^2 z^2}\right).$$

As a product of two independent random variables $e_{\ell m}[n]$ and $e_d[n]$, the pdf of $I_{\ell m}[n]$ is given by

$$\begin{aligned} p_{I_{\ell m}}(z) &= \int_{-\infty}^{\infty} \frac{1}{y} p_{e_{\ell m}}\left(\frac{z}{y}\right) p_{e_d}(y)dy \\ &= \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma_{\ell m}^2}\,y}\exp\left(-\frac{z^2}{2\sigma_{\ell m}^2 y^2}\right) \frac{2}{y^3\sigma_{\hat{u}}^2}\exp\left(-\frac{1}{\sigma_{\hat{u}}^2 y^2}\right) dy. \end{aligned} \tag{6.35}$$

Simplification of Eq. (6.35) results in

$$p_{I_{\ell m}}(z) = \frac{1}{2^{3/2}\sigma_{\ell m}\left(\frac{z^2}{2\sigma_{\ell m}^2} + \frac{1}{\sigma_{\hat{u}}^2}\right)^{3/2} \sigma_{\hat{u}}^2}. \tag{6.36}$$

The derivation from Eq. (6.35) to Eq. (6.36) is given in Appendix B.

When $x_{\ell m}[n]$ takes values of ± 1 with equal probability, the probability of error of $x_{\ell m}[n]$ can be evaluated

$$\begin{aligned} P_{\ell m}(e) &= \Pr(\hat{x}_{\ell m}[n] > 0 \mid x_{\ell m}[n] = -1) \\ &= \Pr(I_{\ell m}[n] > 1) \\ &= \int_1^{\infty} p_{I_{\ell m}}(z)dz \\ &= \frac{1}{2}\left(1 - \frac{\sigma_{\hat{u}}}{\sqrt{\sigma_{\hat{u}}^2 + 2\sigma_{\ell m}^2}}\right). \end{aligned} \tag{6.37}$$

As an example, suppose the Daubechies wavelet of order 14 is used in WPDM, and $\phi_{11}(t)$ and $\phi_{12}(t)$ are the two carriers for two users. After PSAM equalization for a flat fading channel and demodulation in the receiver, the overall noise to $x_{11}[n]$ has a pdf,

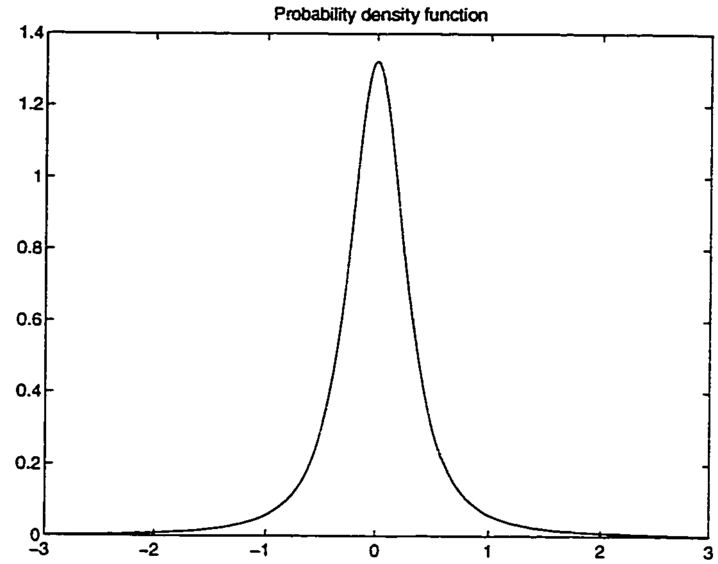
$$p_{I_{11}}(z) = \frac{1}{2^{3/2}\sigma_{11}\left(\frac{z^2}{2\sigma_{11}^2} + \frac{1}{\sigma_{\hat{u}}^2}\right)^{3/2}\sigma_{\hat{u}}^2}, \tag{6.38}$$

where

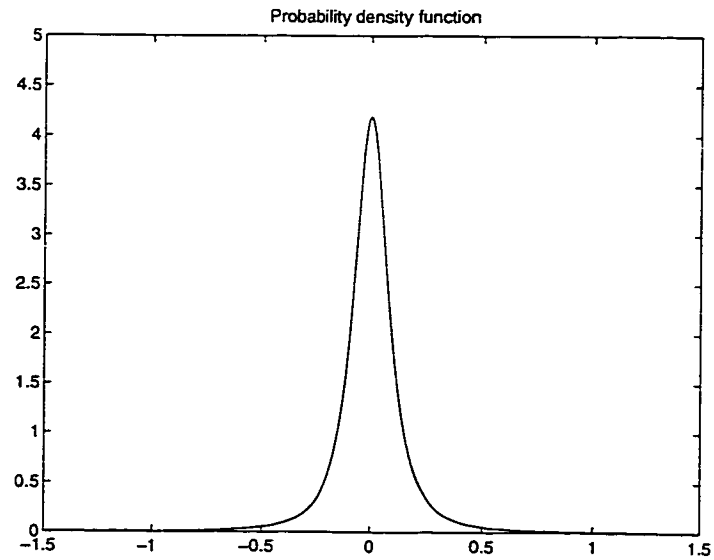
$$\sigma_{11}^2 = \frac{1}{2}(\sigma_{\nu}^2 + \sigma_x^2\sigma_e^2)\sum_k f_{11}^2[k].$$

and $f_{11}[n] = h[n]$ with $h[n]$ defined in Eq. (2.4). To get a better feeling of the overall noise due to fading, we plot the pdfs of $I_{11}[n]$ in Figs. 6.5, where $f_D T_0 = 0.025$.

In Figs. 6.5(a) and (b), the average signal power to the AWGN ratios are $\sigma_u^2/N_0 = 10$dB and $\sigma_u^2/N_0 = 20$dB, respectively. For the Daubechies wavelet, because of Eq. (2.8), i.e., $g[n] = (-1)^n h[1-n]$, and $g[n] = f_{12}[n]$, the pdf of the overall noise to $x_{12}[n]$ is the same as that of $x_{11}[n]$. Therefore, we can expect the same performance for the two users employing $\phi_{11}(t)$ and $\phi_{12}(t)$ as carriers. The expectation can be extended to a general case; i.e. different users employing $\phi_{\ell m}(t)$ would have the same performance in flat fading channels. The reason is that flat fading channels have the same distortion to different frequencies in the signal bandwidth.



(a)



(b)

Figure 6.5: The pdf of the overall interference $I_{11}[n]$ with $f_D T_0 = 0.025$, and the Daubechies wavelet of order 14. (a) $\sigma_u^2/N_0 = 10dB$. (b) $\sigma_u^2/N_0 = 20dB$.

For the purpose of comparison, we derive the probability of error for an FDM system in flat fading channels following the similar procedure as in WPDM. For each user in an FDM system, we assume BPSK is used to modulate its information sequence. Furthermore, pulse shaping is assumed so that finite bandwidths for the users are achieved. Although flat fading process broadens the spectrum of the transmitted signal, we assume that sufficiently wide guard bands are inserted to prevent the broadened spectrum from interfering with adjacent channels. The performance for each user is just the same as that for a BPSK system.

In BPSK modulation scheme, the received low-pass equivalent signal is

$$r(t) = u(t)s(t) + \nu_c(t),$$

where $u(t)$ is the fading process, $\nu_c(t)$ is the low-pass equivalent of AWGN, and

$$s(t) = \sum x[k]\phi(t - kT_0)$$

with $x[k] = \pm 1$ being the information sequence and $\phi(t)$ being the pulse shape. Employing the suboptimum receiver and PSAM for equalization of the channel distortion, we have the estimated $x[n]$

$$\hat{x}[n] = x[n] + \frac{e[n]x[n] + \nu[n]}{\hat{u}[n]},$$

where $e[n]$ is the channel distortion estimation error which is Gaussian distributed with zero mean and variance σ_e^2 and $\nu[n]$ is from the AWGN. The variance of $\nu[n]$ is σ_ν^2. Since $x[n] = \pm 1$, $e[n]x[n]$ is Gaussian distributed as proved in Appendix A, and so is $e[n]x[n]+\nu[n]$, whose variance is $\sigma_e^2 + \sigma_\nu^2$.

Let $I[n]$ denote the overall noise to $x[n]$, i.e., $I[n] = \mathrm{Re}\{e[n]x[n] + \nu[n]\}$. The pdf of $I[n]$ can then be derived as

$$p_I(z) = \frac{1}{2\sqrt{\sigma_e^2 + \sigma_\nu^2}\left(\frac{z^2}{\sigma_e^2+\sigma_\nu^2} + \frac{1}{\sigma_{\hat{u}}^2}\right)^{3/2}\sigma_{\hat{u}}^2}.$$

Hence, the probability of error for each user is

$$\begin{aligned} P(e) &= \Pr\left(\mathrm{Re}\left\{\frac{r_u[n]}{\hat{u}[n]}\right\} > 0 \mid x_{01}[n] = -1\right) \\ &= \Pr(I[n] > 1) \\ &= \frac{1}{2}\left(1 - \frac{\sigma_{\hat{u}}}{\sqrt{\sigma_{\hat{u}}^2 + (\sigma_{\tilde{\nu}}^2 + \sigma_e^2)}}\right), \end{aligned} \tag{6.39}$$

which is the same result as in [76], though different derivation approaches are followed.

To verify the validity of our derivation and the effectiveness of PSAM in combating flat fading, an example is presented, in which the performance of both WPDM and FDM are illustrated.

Before proceeding with the example, let us clarify the definition of signal power. In fading channels, instantaneous signal power is a random variable, and therefore is not appropriate to characterize system performance. When we mention signal power, we refer to the average signal power.

Example 6.4.1 In both WPDM and FDM, we assume that there are two users sharing the channels. Each user has a transmission rate of $1/T$. In WPDM, $\phi_{11}(t)$ and $\phi_{12}(t)$ of Daubechies wavelet of order 14 are used to modulate the signal sequence $x_{11}[n]$ and $x_{12}[n]$ of two users. In FDM, no interchannel interference is considered by assuming that sufficient guardband has been added between the two users. Fig. 6.6 shows the derivation and simulation results about the performance of WPDM and FDM in flat fading channels with Doppler shift $f_D T = 0.05$ as a function of SNR, where SNR is the ratio of average signal power σ_u^2 to the spectral density N_0 of the complex envelope of AWGN. For the two user WPDM system, $T_0 = T/2$; therefore, $f_D T_0 = 0.025$. Since the two WPDM users have the same performance as we demonstrated earlier, the probability of error in Fig. 6.6 applies for both users. Furthermore, for each user in FDM, the pilot spacing M is 6. For WPDM, the pilot spacing in $x_{01}[n]$ is also 6. Later we will show that the pilot spacing used in the

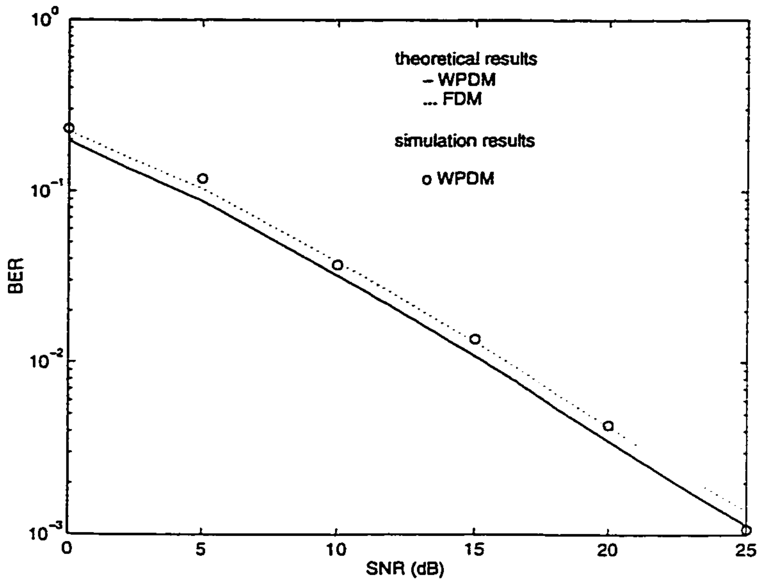


Figure 6.6: The probability of error of WPDM and FDM in flat fading channels with PSAM

example is small enough to result in an accurate channel estimate.

It can be seen that simulation results of a WPDM system closely agree with our derivation. From Fig. 6.6, the error floor is suppressed successfully. In other words, with the PSAM the probability of error can be brought to the expected level by increasing the signal to AWGN ratio.

To compare the performance of WPDM and FDM systems with different pilot spacing, we present another example in which only two users are in either the WPDM or the FDM system.

Example 6.4.2 The wavelet packet used in WPDM is generated from Daubechies scaling function of order 14. Let us assume $f_D T = 0.05$ once more, where $1/T$ is still the transmission rate of one user. For WPDM, $T_0 = T/2$; therefore, $f_D T_0 = 0.025$. Fig. 6.7 shows the theoretical and simulation results of BER vs. the pilot insertion rate M with σ_u^2/N_0 20dB.

It is clear that for the same BER the pilot spacing for WPDM can be larger than the spacing for FDM. In other words, WPDM needs less frequent pilot symbols; therefore

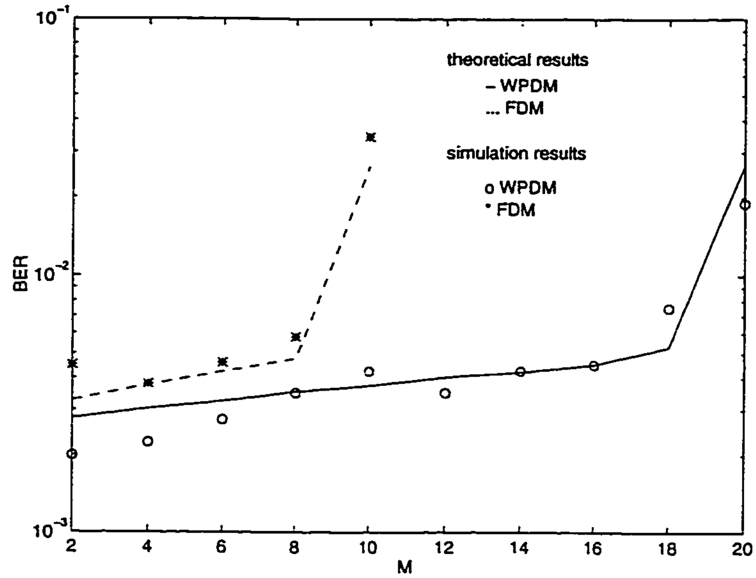


Figure 6.7: Probability of error vs. different pilot spacing M for WPDM and FDM. σ_u^2/N_0=20dB.

WPDM wastes less energy in pilot symbols and gains capacity when compared with the FDM scheme. In the given example, we assume $M = 8$ in FDM and $M = 16$ in WPDM for reliable communications. Then it is easily calculated that 12.5% capacity is used by pilot symbols in FDM, while only 6.25% capacity is used in WPDM.

With more users in WPDM and FDM systems, $f_D T_0$ is much smaller than $f_D T$. Thus, smaller percentage of capacity is spent in pilot symbols. However, T_0 has to be sufficiently large to maintain the validity of the flat fading assumption. As a consequence, $f_D T_0$ cannot be decreased indefinitely.

6.4.2 Performance of WPDM in frequency selective fading channels with PSAM

When frequency selective fading channel is a two-ray multipath channel described in Eq. (6.12), we have the low-pass equivalent of the channel output

$$r(t) = u_1(t)s(t) + u_2(t)s(t-\tau) + \nu_c(t) \tag{6.40}$$

where $s(t)$ is the multiplexed signal defined in Eq.(3.3). It is still assumed that every Mth symbol in $x_{01}[n]$ is the pilot symbol. However, how to estimate the channel distortion is not clear at the moment. First, as in the flat fading case, the received signal $r(t)$ is passed through the matched filter matching to $\phi_{01}(t)$ and a sampler. Since $u_1(t)$ and $u_2(t)$ are slow processes and can be regarded as a constant in several bit duration, the output of the matched filter is

$$\begin{aligned} r_f[n] &= \frac{1}{\sqrt{E_0}} \int r(t)\phi_{01}(t - nT_0)dt \\ &= u_1[n]x_{01}[n] + u_2[n] \sum_k x_{01}[k] R_\phi(nT_0 - kT_0 - \tau) + \nu[n], \end{aligned} \tag{6.41}$$

where $u_1[n]$ and $u_2[n]$ are the sampled values of $u_1(t)$ and $u_2(t)$ at instant nT_0, and $R_\phi(\tau)$ is the autocorrelation function of $\phi_{01}(t)$ whose definition can be found in Eq. (4.8), and $\nu[n]$ is as in Eq. (6.24).

Re-arrangement of the right hand of Eq. (6.41) results in

$$r_f[n] = \Big(u_1[n] + u_2[n]R_\phi(-\tau)\Big)\, x_{01}[n] + u_2[n] \sum_{k \neq 0} x_{01}[n-k] R_\phi(kT_0 - \tau) + \nu[n]. \tag{6.42}$$

Since $u_1(t)$ and $u_2(t)$ are two bandlimited complex Gaussian process, $u_1[n] + u_2[n]R_\phi(-\tau)$ is still a bandlimited Gaussian process.

Comparing $r_f[n]$ in Eq.(6.42) with Eq. (6.14), we can identify easily that

$$v[n] = u_1[n] + u_2[n]R_\phi(-\tau) \stackrel{\text{def}}{=} u_f[n], \tag{6.43}$$

whose power is $\sigma_{u_f}^2 = \sigma_{u_1}^2 + R_\phi^2(-\tau)\sigma_{u_2}^2$, and the total noise in frequency selective fading channels is

$$\tilde{\nu}_0[n] = \nu[n] + u_2[n] \sum_{k \neq 0} x_{01}[n-k] R_\phi(kT_0 - \tau) \stackrel{\text{def}}{=} \nu_f[n]. \tag{6.44}$$

Therefore, through pilot symbols, $u_f[kM]$ can be estimated in the same way as in a flat fading situation. Interpolating $\hat{u}_f[kM]$ gives the estimates of the channel distortion on

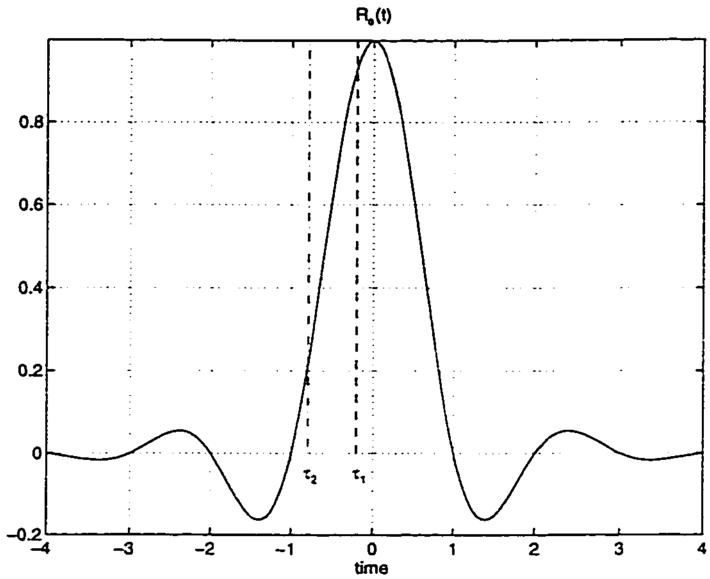


Figure 6.8: Autocorrelation function of $\phi_{01}(t)$.

data symbols. The second term $u_2[n]\sum_{k\neq 0} x_{01}[n-k]R_\phi(mT_0-\tau)$ in Eq. (6.42) cannot be estimated, therefore is an interference, and increases noise level.

The autocorrelation function $R_\phi(t)$ for $\phi_{01}(t)$ being a Daubechies wavelet function of length 14 is plotted in Fig. 6.8. From Fig. 6.8 (where T_0 is normalized to 1), we can see that when the delay τ of the second path relative to the main path is small compared with the bit duration T_0, $R_\phi(-\tau)$ is usually bigger than $R_\phi(mT_0-\tau)$ with $m\neq 0$. The second path contributes more to signal than to interference. On the other hand, if τ is big, but still within one bit duration, the second path serves mainly as an interference, since the distortion caused by $u_2[n]$ cannot be effectively estimated. When τ exceeds T_0, the second path is just an interference.

Substituting Eqs. (6.43) and (6.44) into Q_{ik} and $w_i[k]$ in Eq. (6.17) and (6.18), we have

$$\begin{aligned}
Q_{ik} &= \frac{1}{2}E\{\hat{u}_f[iM]\hat{u}_f^*[kM]\} \\
&= \tilde{R}_u\left((i-k)MT_0\right)\left(\sigma_{u_1}^2+\sigma_{u_2}^2R_\phi^2(-\tau)\right) \\
&\quad +\sigma_{u_2}^2\tilde{R}_u\left((i-k)MT_0\right)\sum_{m\neq 0}R_\phi(mT_0-\tau)R_\phi\left((kM-iM+m)T_0-\tau\right)
\end{aligned}$$

$$
\begin{aligned}
& + \frac{N_0 \delta[i-k]}{E_0 \; E\{|x_{01}[n]|^2\}} \\
w_i[k] &= \frac{1}{2} E\{u_f^*[k] \hat{u}_f[iM]\} \\
&= \left(\sigma_{u_1}^2 + \sigma_{u_2}^2 R_\phi^2(-\tau)\right) \tilde{R}_u \left((iM-k)T_0\right).
\end{aligned}
$$

Similarly the optimum Wiener filter coefficient $\mathbf{d}[k]$ can be calculated from normal equation, and the channel distortion can be estimated. The variance of the estimated $\hat{u}_f[n]$ is given by

$$\sigma_{\hat{u}_f}^2[k] = \mathbf{w}^\dagger[k] \mathbf{Q}^{-1} \mathbf{w}[k].$$

Let $e_f[n]$ be the estimation error associated with the estimate $\hat{u}_f[n]$; i.e.,

$$e_f[n] = u_f[n] - \hat{u}_f[n]. \tag{6.45}$$

Its variance is

$$\sigma_{e_f}^2[k] = \sigma_{u_f}^2[k] - \sigma_{\hat{u}_f}^2[k]. \tag{6.46}$$

As in flat fading, the channel is compensated by the channel estimate,

$$\hat{r}_f[n] = \frac{r_f[n]}{\hat{u}_f[n]}, \tag{6.47}$$

where $\hat{u}_f[n]$ is the estimate of the channel distortion, and therefore a slow complex Gaussian process.

By substituting the expression of $r_f[n]$ in Eq. (6.42) in (6.47), the compensated signal sequence $\hat{r}_f[n]$ may be further written as

$$
\begin{aligned}
\hat{r}_f[n] &= \frac{(\hat{u}_f[n] + e_f[n]) x_{01}[n] + \nu_f[n]}{\hat{u}_f[n]} \\
&= x_{01}[n] + \frac{e_f[n] x_{01}[n] + \nu_f[n]}{\hat{u}_f[n]}.
\end{aligned} \tag{6.48}
$$

Passing the compensated sequence $\hat{r}_f[n]$ through wavelet packet decomposition filters, we have

$$\hat{x}_{\ell m}[n] = \sum_k \hat{r}_f[k] f_{\ell m}[k - 2^\ell n]. \tag{6.49}$$

Insertion of $\hat{r}_f[n]$ in the above equation results in

$$\hat{x}_{\ell m}[n] = x_{\ell m}[n] + \sum_k \frac{e_f[k]x_{01}[k] + \nu_f[k]}{\hat{u}_f[k]} f_{\ell m}[k - 2^\ell n]. \tag{6.50}$$

Once again, for real $x_{01}[n]$, only the real part of the second term in Eq. (6.50) is of interest. For the notation simplicity, we denote

$$I_{f,\ell m}[n] = \mathrm{Re}\left\{\sum_k \frac{e_f[k]x_{01}[k] + \nu_f[k]}{\hat{u}_f[k]} f_{\ell m}[k - 2^\ell n]\right\}.$$

Similar to the analysis in flat fading channels, making use of the slowly varying property of $\hat{u}_f[n]$, we may write $I_{f,\ell m}[n]$ as

$$I_{f,\ell m}[n] \simeq \mathrm{Re}\left\{\frac{\sum_k \left(e_f[k]x_{01}[k] + \nu_f[k]\right) f_{\ell m}[k - 2^\ell n]}{\hat{u}_f[2^\ell n]}\right\}, \tag{6.51}$$

as long as the varying rate of $\hat{u}_f[n]$ is not fast compared with the length of $f_{\ell m}[n]$.

Using the circular symmetry property of $e_f[n]$, $\nu_f[n]$, and $\hat{u}_f[n]$, Eq. (6.51) can be written as

$$I_{f,\ell m}[n] = \frac{\mathrm{Re}\left\{\sum_k \left(e_f[k]x_{01}[k] + \nu_f[k]\right) f_{\ell m}[k - 2^\ell n]\right\}}{|\hat{u}_f[2^\ell n]|}.$$

Denote

$$e_{f,\ell m}[n] = \mathrm{Re}\left\{\sum_k (e_f[k]x_{01}[k] + \nu_f[k]) f_{\ell m}[k - 2^\ell n]\right\}.$$

Though $e_f[n]$ is a sequence of independent random variables and so is $x_{01}[n]$, $\nu_f[n] = \nu[n] + u_2[n]\sum_{k\neq 0} x_{01}[n-k]R_\phi(kT_0 - \tau)$ is a correlated sequence. Thus $e_{f,\ell m}$ is the sum of K correlated random variables, where K is the length of filter $f_{\ell m}[n]$. However, $e_{f,\ell m}[n]$ can still be approximated by a Gaussian distribution [71]. Usually for an accurate approximation, the number of correlated random variables involved in the summation should be bigger than the number of independent random variables. In Appendix D, we demonstrate the accuracy of the approximation of $e_{f,\ell m}[n]$ as a Gaussian distribution.

In the approximation, the mean and variance of $e_{f,\ell m}[n]$ are

$$E\left\{e_{f,\ell m}[n]\right\} = 0$$

and

$$\begin{aligned}
\sigma^2_{f,\ell m} &= E\{e^2_{f,\ell m}[n]\} \\
&= \frac{1}{2}\left(\sigma_e^2 E\{x_{01}^2[n]\} + \sigma_\nu^2\right) \sum_k f_{\ell m}^2[k - 2^\ell n] \\
&\quad + \frac{1}{2}\sigma_{u_2}^2 E\{x_{01}^2[n]\} \cdot \\
&\quad \sum_k \sum_j \tilde{R}_{u_2}((k-j)T_0) \sum_{i \neq 0} R_\phi(i-\tau) R_\phi(j-k+i-\tau) f_{\ell m}[k-2^\ell n] f_{\ell m}[j-2^\ell n].
\end{aligned} \tag{6.52}$$

Similar procedure to derive Eq. (6.36) can be followed to obtain the pdf of $I_{f,\ell m}[n]$ as

$$p_{I_{f,\ell m}}(z) = \frac{1}{2^{3/2}\sigma_{f,\ell m}\left(\frac{z^2}{2\sigma^2_{f,\ell m}} + \frac{1}{\sigma^2_{\hat{u}_f}}\right)^{3/2} \sigma^2_{\hat{u}_f}}. \tag{6.53}$$

If binary transmission is assumed for each user, the probability of error of $x_{\ell m}[n]$ is

$$\begin{aligned}
P_{f,\ell m}(e) &= \Pr\left(\hat{x}_{\ell m}[n] > 0 \,|\, x_{\ell m}[n] = -1\right) \\
&= \Pr(I_{f,\ell m}[n] > 1) \\
&= \int_1^\infty p_{I_{f,\ell m}}(z)\, dz \\
&= \frac{1}{2}\left(1 - \frac{\sigma_{\hat{u}_f}}{\sqrt{\sigma^2_{\hat{u}_f} + 2\sigma^2_{f,\ell m}}}\right).
\end{aligned} \tag{6.54}$$

It can be seen that the probability of error at terminal (ℓm) is a function of $\sigma^2_{f,\ell m}$, which we know from Eq. (6.52) is a function of the correlation of $f_{\ell m}[n]$. Thus, users at different terminals may experience different performance. This is the expected result, since $\phi_{\ell m}(t)$ concentrates in different frequency bands and the frequency-selective fading affects the different frequency bands in different ways. The performance selectivity will be illustrated later.

Let us first present an example to demonstrate the validity of our derivation.

Example 6.4.3 Consider a specific WPDM system consisting of two branches, using $\phi_{11}(t)$ and $\phi_{12}(t)$ as carriers for the binary message signals $x_{11}[n]$ and $x_{12}[n]$. The transmission

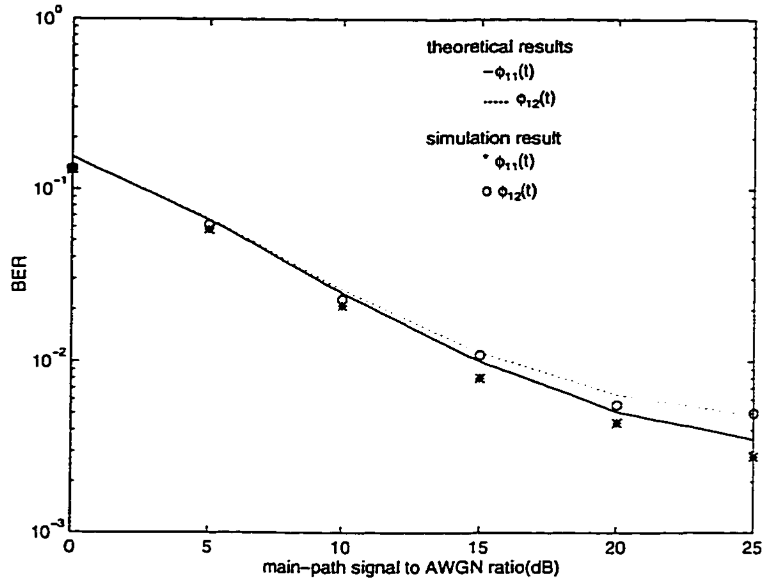


Figure 6.9: Simulation and derivation results on the probability of error of WPDM in two-ray multipath channel with PSAM.

channel is a two-ray fading channel, with each of the path having a Doppler shift of $f_D T_0 = 0.025$. The second path has a delay of $\tau = 0.1T_0$ relative to the main path, and the signal power ratio of the main path to the second path is 3dB. Moreover, the channel is corrupted by AWGN. The pilot spacing is $M = 6$. The performance of WPDM is evaluated both theoretically and by simulation, and the results are plotted in Fig. 6.9.

It can be observed that the theoretical analysis agree reasonably with the simulation results, although during the analysis, approximations were used.

We can also observe that for $x_{11}[n]$ and $x_{12}[n]$, different performance was obtained. This is the expected results, because $\phi_{11}(t)$ and $\phi_{12}(t)$ concentrate on different frequency bandwidth, and the distortion effect of a frequency selective channel varies with frequencies.

The frequency selectivity will be illustrated in next example in more detail.

Example 6.4.4 Assume the second level terminal functions $\phi_{21}(t)$, $\phi_{22}(t)$, $\phi_{23}(t)$, and $\phi_{24}(t)$ of a Daubechies wavelet packet of length 14 are used as carriers in a WPDM system. The channel of interest is still a two-ray fading channel with each path having a Doppler shift of $f_D T_0 = 0.025$, and $M = 6$ is chosen to be the pilot spacing in PSAM. Fig. 6.10 shows the performance of each user with different combination of delay τ and the signal power ratio of the main path to the second path $\sigma_{u_1}^2/\sigma_{u_2}^2$.

In Fig. 6.10(a), $\tau = 0.2T_0$ and $\sigma_{u_1}^2/\sigma_{u_2}^2 = 0$dB. In Fig. 6.10(b), $\tau = 0.1T_0$ and $\sigma_{u_1}^2/\sigma_{u_2}^2 =$ 5dB.

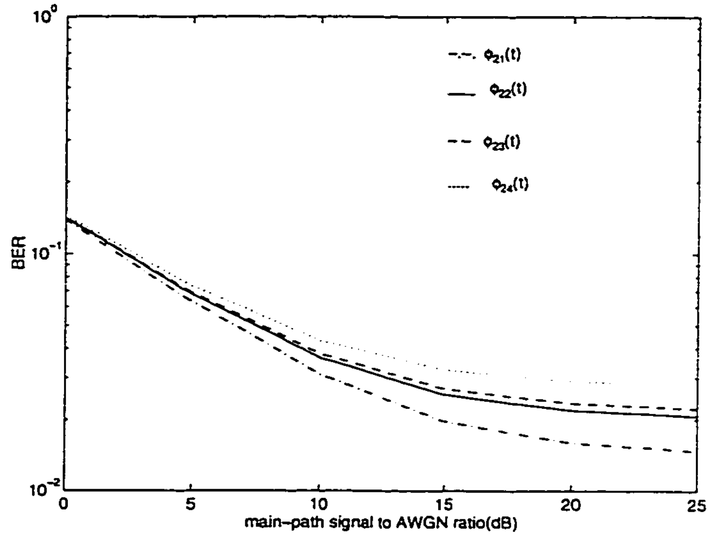
The performance selectivity on the different subchannels is obvious, and is characteristic of frequency selective channels.

We now examine the effect of the relative delay τ.

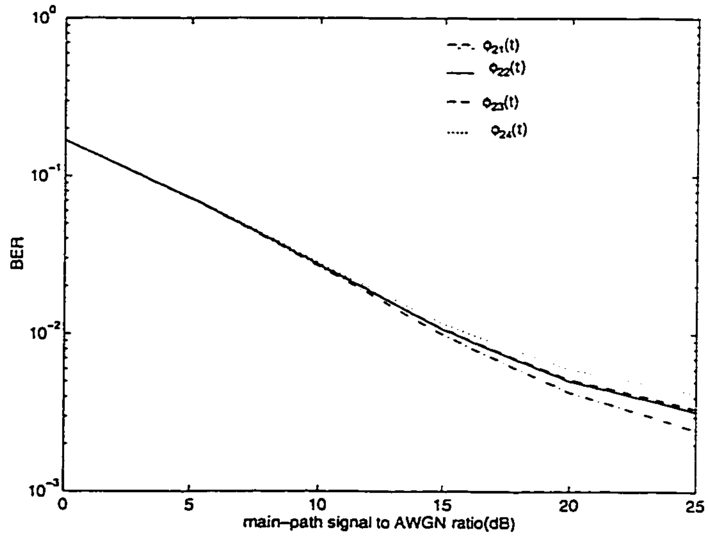
Example 6.4.5 Once again, we consider a two user WPDM system using $\phi_{11}(t)$ and $\phi_{12}(t)$ modulating $x_{11}[n]$ and $x_{12}[n]$. Fig. 6.11 shows the performance of WPDM in two-ray fading channel with $f_D T_0 = 0.025$ and various τ and $\sigma_{u_1}^2/\sigma_{u_2}^2$.

In each of Fig. 6.11(a), (b), and (c), the probability of error for $\tau = 0.1T_0$, $\tau = 0.2T_0$, and $\tau = 0.3T_0$ are plotted. Fig. 6.11(a), (b), and (c) correspond to $\sigma_{u_1}^2/\sigma_{u_2}^2 = 0$dB, $\sigma_{u_1}^2/\sigma_{u_2}^2 =$ 5dB, and $\sigma_{u_1}^2/\sigma_{u_2}^2 = 10$dB, respectively.

Fig. 6.11 shows that when the relative delay τ of the second path is small, for example $\tau = 0.1T_0$ or $\tau = 0.2T_0$, WPDM performs well. In this case, the second path contributes more to signal than to interference. When τ is big, say $\tau = 0.3T_0$, PSAM does not help improve the performance. It can also be seen that PSAM does not suppress the error floor in frequency selective fading as successfully as in flat fading. The reason for this is that PSAM is intended for the error floor removal in flat fading channels, and the performance improvement with PSAM is limited in the frequency selective fading channels. When main-path signal to AWGN ratio is large, the dominant noise is the interference from the second path, which is determined by the relative delay τ and the power of the second path. Therefore, increasing the main-path signal to AWGN ratio cannot give better performance indefinitely; i.e., the error-floor still exists as shown in Fig. 6.11. Indeed, when the main path to the

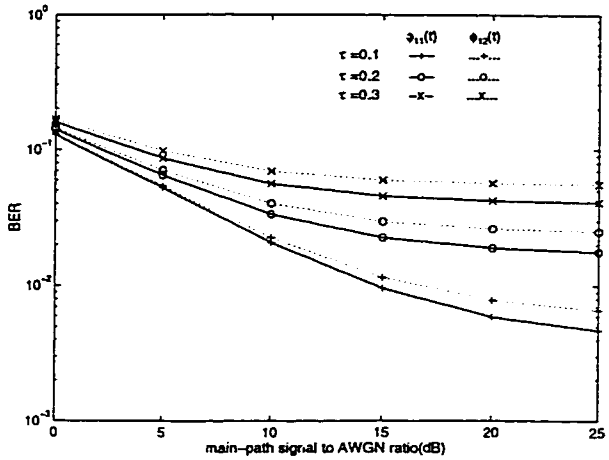


(a)

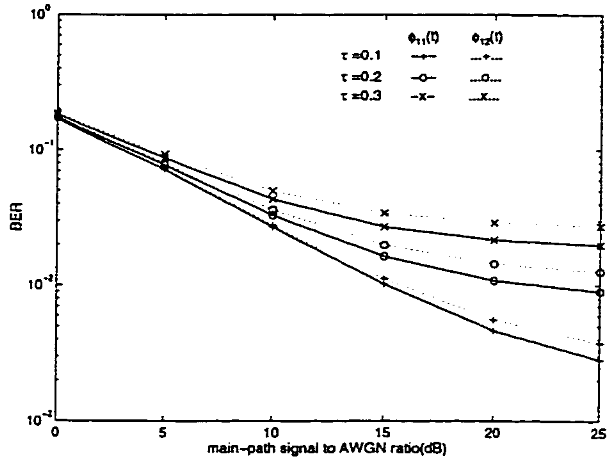


(b)

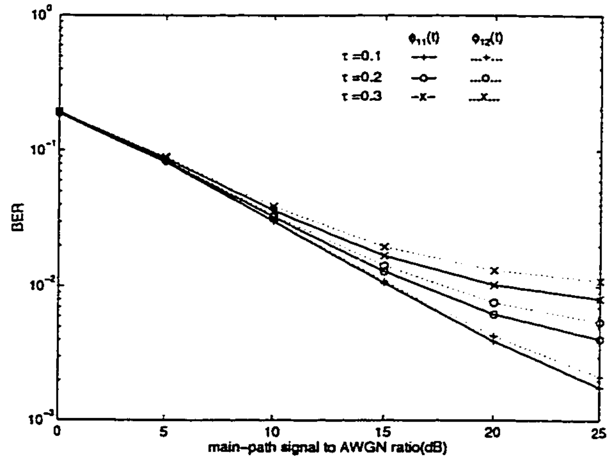
Figure 6.10: The selectivity of two-ray multipath channel on different sub-channels of WDPM. (a) $\sigma_{u_1}^2/\sigma_{u_2}^2 = 0dB$ and $\tau = 0.2T_0$; (b) $\sigma_{u_1}^2/\sigma_{u_2}^2 = 5dB$ and $\tau = 0.1T_0$.



(a)



(b)



(c)

Figure 6.11: The probability of error of WPDM in two-ray multipath channel with PSAM. The main-path signal to the second path ratio (a) 0dB (b) 5dB (c) 10dB

second path signal power ratio is bigger, the error floor suppression is more striking.

6.5 Summary

Since fading characterizes the mobile channels, the performance of WPDM in fading channels determines the success of WPDM in mobile communication systems. In this chapter, the performance of WPDM system in both flat fading and two-ray frequency selective fading channels was derived. As in many other communications systems, to combat the irreducible error due to fading, a time domain equalization PSAM was applied to WPDM systems.

It was shown that in flat fading channels, the error floor was suppressed successfully with PSAM, although WPDM has a non-constant envelope. For the purpose of comparison, the performance of an FDM system in the same channel environment was also evaluated. The calculation and corresponding simulation results illustrated that the pilot spacing for WPDM can be larger than the spacing for FDM. In other words, WPDM needs less frequent pilot symbols and therefore gains capacity in flat fading channels compared to the FDM scheme.

For a special frequency selective fading channel, the two-ray multipath channel, the performance analysis of WPDM revealed that when the relative delay τ of the second path to the main path is small compared with the bit duration T_0, the irreducible error is successfully suppressed with PSAM. While τ is big in comparison to T_0, the error floor cannot be completely suppressed by using PSAM, especially when the main path to the second path signal power ratio is small. The reason for this is that the second path could not be effectively estimated and the signal power in it could not be constructively used. Hence, the second path contributes more significantly to the interference than to the signal. This is in fact the limitation of the time-domain equalization.

Chapter 7

Conclusion

A new scheme for transmitting multiple signals using wavelet packets (WPDM) has been proposed. The concepts of the scheme have been outlined and the transmission and reception systems, with their equivalent realizations, have been developed. Wavelet packet basis functions are employed as coding waveforms because they provide self orthogonality based on time translation and mutual orthogonality based on occupancy of different orthogonal subspaces. As a general form of FDM and TDM, when WPDM is used in the extreme case as a kind of FDM, it has the advantages of a parallel transmission system. However, the orthogonalities in a wavelet packet allow the waveforms to be overlapped in both time and frequency, eliminating the need for guard bands in FDM, and thereby increasing the bandwidth efficiency of the transmission channel. In addition, the transmission system also offers communication security, since knowledge of the choice of wavelet packet is required for decoding of the message signals. While WPDM offers these advantages, we note that in common with several other communication schemes (e.g., rectangular 16-QAM), the transmitted signal does not have a constant envelope. In order to control the effects of envelope variations, we may have to limit the transmission power in order to avoid operating the transmission amplifiers near their non-linear saturation regions [51]. If an attempt is made to restore the constant envelope characteristic by passing the composite signal though a hard-limiter, the orthogonality of the constituent waveforms may be destroyed, resulting in potentially unacceptable performance.

We have examined the problem of timing error in WPDM and observed that the interference due to a timing discrepancy between the transmitter and receiver is made up of two components, ISI and crosstalk, which can be respectively modeled as the message bits from the given channel and the adjacent channels passed through corresponding filters. Though the interference filters take different forms under constant timing error and timing jitters, the simple filter models led us to propose an optimum design of the coding waveforms (i.e., the scaling function and its constituents) in which the total timing-error interference is minimized. We also obtained expressions for the probability of error under the effects of timing error and additive white Gaussian channel noise. Simulations confirmed these expressions and showed that employing the optimum scaling function as the fundamental waveform results in a lower probability of error in transmission than other commonly used scaling functions.

Throughout the thesis, we have focused our attention on orthogonal multiple signal transmission. We can easily obtain extra degrees of freedom by expanding our formulation to biorthogonal signaling using biorthogonal wavelet packets and filter banks [62, 79]. However, we note that Proposition 4.3.1 no longer holds and the minimization of the total interference energy becomes structure dependent. We must therefore reformulate the optimization in terms of the interference energies at each level. Furthermore, care must be taken in order to avoid excessive noise gain and numerical instability [80] in the wavelet packet decomposer.

As a parallel transmission system, WPDM is intuitively more robust to impulsive noise than a serial transmission system, such as TDM. The performance of a WPDM system in the presence of both Gaussian and impulse noise has been analyzed, and has been compared with TDM system. It has been found that WPDM indeed provides greater immunity to impulsive noise than TDM.

With bandwidth efficiency and flexibility in assignment of subchannels and limited signal power to maximize channel capacity, WPDM is promising for mobile communications. Since fading characterizes the mobile and radio channels, the performance of a WPDM in fading channels determine the success of WPDM in mobile communication systems. We have derived the performance of WPDM system in both flat fading and two-ray frequency selective fading channels. As in many other communications systems, to combat the irreducible error due to fading, a time domain equalization PSAM has been applied in WPDM system.

In flat fading channels, the error floor has been suppressed successfully with PSAM, although WPDM has a non-constant envelope. For the purpose of comparison, the performance of an FDM system in the same channel environment has also been evaluated. The calculation and corresponding simulation results illustrate that the pilot spacing for WPDM can be larger than the spacing for FDM, and thus WPDM gains capacity and wastes less signal power on pilot symbols in flat fading channels compared to the FDM scheme.

For a special frequency selective fading channel, the two-ray multipath channel, the performance analysis of WPDM has revealed that when the relative delay τ of the second path to the main path is small compared with the bit duration T_0, the irreducible error is successfully suppressed with PSAM. When τ is big in comparison to T_0, the error floor cannot be completely suppressed by using PSAM, especially when the main path to the second path signal power ratio is small. The reason for this is that the second path cannot be effectively estimated and the signal power in it cannot be constructively used. Hence, the second path contributes more significantly to the interference than to the signal. This is in fact the limitation of time-domain equalization. To achieve the significant performance improvement in frequency selective fading channels, frequency domain equalization techniques or diversity techniques should be used.

In summary, given the ever increasing demand on channel capacity in communication systems, the bandwidth efficient advantages of WPDM, together with the simplicity of its implementation, its emerging robustness properties, and the flexibility in the assignment of the subchannels and the limited signal power, show that WPDM holds considerable promise as a multiple signal transmission technique in both wireline and wireless communication channels.

7.1 Future work

Although the concepts of WPDM have been developed and the WPDM has been studied in several environments, there are some interesting issues worth of investigations.

First, by analogy with frequency-hopped communiction schemes, incorporation of time variation into the WPDM scheme may offer the potential for further performance improvements, especially in frequency selective fading channels, and further security improvements.

Second, since orthogonality in wavelet packet basis functions is not a necessary condition for a WPDM system, non-orthogonal wavelet packets (e.g. biorthogonal wavelet packets) may be used to provide us with a more bandwidth efficient WPDM system. However, in this case, some properties in orthogonal wavelet packets multiplexing may no longer hold.

Third, in WPDM, there may be more properties to be exploited to further improve the bandwidth efficiency. One of the properties may be cyclostationary property in wavelet packet multiplexed signals.

These further investigations may lead to a better understanding and a fuller exploitation of wavelet packets in the applications to multi-user communication systems.

Appendix A

The pdf of $e[n]x_{01}[n]$

In this appendix, we present a proof for that $z[n] = e[n]x_{01}[n]$ has a Gaussian distribution if $e[n]$ is Gaussian distributed with zero mean, $x_{01}[n] = \pm 1$, and $e[n]$ is independent of $x_{01}[n]$.

Let us denote $p_e(.)$ and $p_z(.)$ as the pdf of $e[n]$ and $z[n]$, respectively.

The distribution $F_z(z)$ of $z[n]$ is given by

$$F_z(z) = P(e[n]x_{01}[n] \leq z).$$

Assuming $x_{01}[n]$ takes on $+1$ and -1 with probability of P_1 and $1 - P_1$, respectively, and making use of the assumption that $e[n]$ and $x_{01}[n]$ are independent, we have

$$F_z(z) = P_1 P(e[n] \leq z) + (1 - P_1)P(e[n] \geq -z).$$

Because $e[n]$ is assumed to have a zero mean Gaussian distribution, $p_e(z)$ is an even function and

$$P(e[n] \geq -z) = P(e[n] \leq z).$$

Thus we have

$$\begin{aligned} F_z(z) &= P_1 P(e[n] \leq z) + (1 - P_1)P(e[n] \leq -z) \\ &= P(e[n] \leq z) \end{aligned}$$

$$= F_e(z),$$

where $F_e(z)$ is the distribution of $e[n]$. The same distributions correspond to the same pdf; i.e.,

$$p_z(z) = p_e(z),$$

and $z[n]$ has a Gaussian distribution.

Appendix B

Simplifying $p_{I_{\ell m}}(z)$

We re-write Eq. (6.35) for convenience,

$$
\begin{aligned}
p_{I_{\ell m}}(z) &= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_{\ell m}^2 y}} \exp\left(-\frac{z^2}{\sigma_{\ell m}^2 y^2}\right) \frac{2}{y^3\sigma_{\hat{u}}^2} \exp\left(-\frac{1}{\sigma_{\hat{u}}^2 y^2}\right) dy \\
&= \frac{2}{\sqrt{2\pi\sigma_{\ell m}^2}\sigma_{\hat{u}}^2} \int_0^\infty \frac{1}{y^4} \exp\left(-\frac{\frac{z^2}{2\sigma_{\ell m}^2} + \frac{1}{\sigma_{\hat{u}}^2}}{y^2}\right) dy. \qquad \text{(B.1)}
\end{aligned}
$$

Let us denote

$$
a = \frac{z^2}{2\sigma_{\ell m}^2} + \frac{1}{\sigma_{\hat{u}}^2}
$$

and change variable $t = 1/y$. Eq. (B.1) becomes

$$
p_b(z) = \frac{2}{\sqrt{2\pi\sigma^2}\sigma_{\hat{u}}^2} \int_0^\infty t^2 \exp(-at^2) dt. \qquad \text{(B.2)}
$$

The integration part in Eq. (B.2) equals to $\frac{\sqrt{\pi}}{4a^{3/2}}$. Therefore we have

$$
p_b(z) = \frac{1}{2^{3/2}\sigma_{\ell m}\left(\frac{z^2}{2\sigma_{\ell m}^2} + \frac{1}{\sigma_{\hat{u}}^2}\right)^{3/2}\sigma_{\hat{u}}^2}. \qquad \text{(B.3)}
$$

Appendix C

Approximation of $e_{\ell m}[n]$ by a Gaussian distribution

Re-write the definition of $e_{\ell m}[n]$ here for convenience,

$$e_{\ell m}[n] = \mathrm{Re}\left\{\sum_k \left(e[k]x_{01}[k] + \nu[k]\right) f_{\ell m}[k - 2^{\ell} n]\right\},$$

where $e[n]$ is the channel distortion estimation error, which has a complex Gaussian distribution, $x_{01}[n]$ is reconstructed from $x_{\ell m}[n] = \pm 1$ by wavelet packet reconstruction filter, and $\nu[n]$ is from the AWGN. The random variables $e[n]$, $x_{01}[n]$, and $\nu[n]$ are all independent.

As the sum of independent random variables, $e_{\ell m}[n]$ can be approximated by Gaussian distribution with

$$E\left\{e_{\ell m}[n]\right\} = 0$$

and

$$\sigma_{\ell m}^2 = E\left\{e_{\ell m}^2[n]\right\} = \frac{1}{2}(\sigma_e^2\sigma_x^2 + \sigma_\nu^2)\sum_k f_{\ell m}^2[k - 2^{\ell} n].$$

The pdf of $e_{\ell m}[n]$ is therefore,

$$p_{e_{\ell m}}(z) = \frac{1}{\sqrt{2\pi\sigma_{\ell m}^2}} \exp\left(-\frac{z^2}{2\sigma_{\ell m}^2}\right).$$

Now we demonstrate the accuracy of the approximation. For a wavelet packet, $\sum_k f_{\ell m}^2[k]$ is a constant for all (ℓm), $\sigma_{\ell m}^2$ is not a function of (ℓm). Thus in terms of pdf property, $e_{\ell m}[n]$ from any (ℓm) would be representative of every (ℓm). In this appendix, the data from all $e_{\ell m}[n]$ in a wavelet packet are used to obtain a histogram.

We generate a long sequence of $e_{\ell m}[n]$ in the following procedures.

- Generate a long sequence of Gaussian distributed $\mathrm{Re}\{e[n]\}$ with zero mean and variance $\frac{1}{2}\sigma_e^2$.

- Generate a long sequence of Gaussian distributed $\mathrm{Re}\{\nu[n]\}$ with zero mean and variance $\frac{1}{2}\sigma_\nu^2$.

- Generate $x_{01}[n]$ from $x_{\ell m} = \pm 1$ by wavelet packet reconstruction filter. The sequence $x_{01}[n]$ depends on the wavelet packet chosen in WPDM. Here we choose the Daubechies wavelet with $N = 14$.

- To get $e_{\ell m}[n]$, compute $\sum_k (\mathrm{Re}\{e[k]\}\, x_{01}[k] + \mathrm{Re}\{\nu[k]\})\, f_{\ell m}[k - 2^\ell n]$.

For the experimentally generated $e_{\ell m}[n]$, a histogram $Hist(b_i, y_i)$ can be obtained, where b_i is the center of ith bin, the bin size b_s is given by $b_{i+1} - b_i$, and y_i is the accumulation number. For the purpose of comparison, we normalize a histogram by

$$\overline{y}_i = \frac{y_i}{b_s \sum_i y_i},$$

such that

$$\sum_i \overline{y}_i b_s = 1.$$

Comparing the histogram with the assumption of the pdf of $e_{\ell m}[n]$, we can see heuristically how accurate the approximation is. Moreover we use the square error

$$d^2 = \sum_i \left(\overline{y}_i - p_{e_{\ell m}}(b_i)\right)^2 b_s$$

to be a measure of the approximation accuracy. To have a more accurate quantitative comparison, we estimate the up to 4th moments of the experimentally generated $e_{\ell m}[n]$ [71, 81] by

$$\begin{aligned}
\hat{m}_1 &= \frac{1}{L_0}\sum_{i=0}^{L_0} e_{\ell m}[i] \\
\hat{m}_2 &= \frac{1}{L_0}\sum_{i=0}^{L_0} (e_{\ell m}[i]-\hat{m}_1)^2 \\
\hat{m}_3 &= \frac{1}{L_0}\sum_{i=0}^{L_0} (e_{\ell m}[i]-\hat{m}_1)^3 \\
\hat{m}_4 &= \frac{1}{L_0}\sum_{i=0}^{L_0} (e_{\ell m}[i]-\hat{m}_1)^4 .
\end{aligned}$$

The difference between $\hat{m}_i (i = 1, 2, 3, 4)$ and the corresponding moments m_i of Gaussian distribution characterizes the accuracy of the approximation. Smaller difference indicates more accurate approximation. The moments of Gaussian distribution can be calculated [72]as

$$\begin{aligned}
m_{2n} &= \frac{1}{\sqrt{2\pi\sigma^2_{e_{\ell m}}}}\int_{-\infty}^{\infty} x^{2n}\exp\left(-\frac{x^2}{2\sigma^2_{\ell m}}\right) \\
&= \frac{(2n)!}{2^n n!}\sigma^2_{\ell m},
\end{aligned}$$

and

$$m_{2n+1} = 0.$$

In the following examples, the number of bins of the histogram are arbitrarily chosen to be 31.

First we assume that $\phi_{11}(t)$ and $\phi_{12}(t)$ are used for carrying the the information sequence in WPDM. In Fig. C.1, histograms and the Gaussian distribution are shown for (a) $\sigma_u^2/N_0 =$ 25dB, $f_D T_0 = 0.025$ and (b) $\sigma_u^2/N_0 = 10dB$, $f_D T_0 = 0.0125$.

Since $e_{11}[n]$ and $e_{12}[n]$ have the same distribution property, the histogram is obtained from $(e_{11}[n], e_{12}[n])$. The variance of $e[n]$ are computed from Eq. (6.25). In Tables C.1, and C.2, we list the estimated moments $\hat{m}_i$, the moments of Gaussian distribution m_i, and their

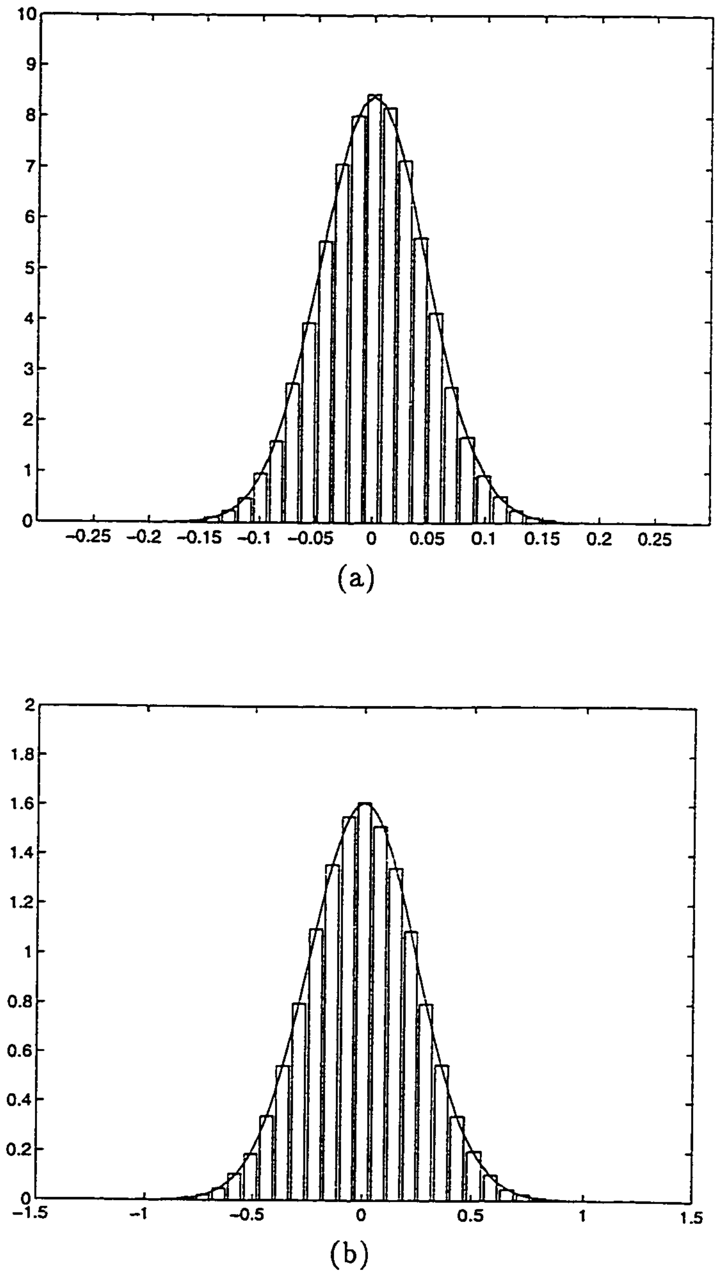


Figure C.1: Histogram of $e_{11}[n]$ and ideal Gaussian distribution. (a) $f_D T_0 = 0.025$, and $\sigma_u^2/N_0 = 25$dB. (b) $f_D T_0 = 0.0125$, and $\sigma_u^2/N_0 = 10$dB.

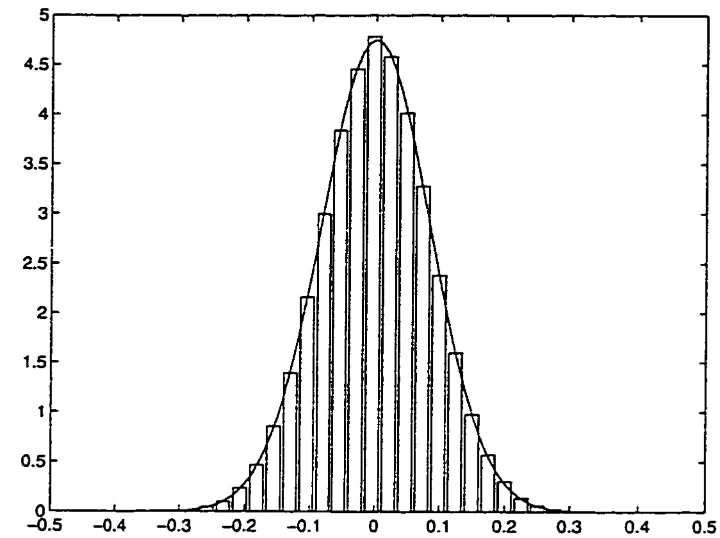


Figure C.2: Histogram of e_{2i} and ideal Gaussian distribution. $f_D T_0 = 0.025$, and $\sigma_u^2/N_0 =$ 20dB.

difference. As a measure of the approximation accuracy, d^2 for each approximation is also shown in Tables C.1, and C.2.

For the WPDM employing $\phi_{2i}(t), (i = 1, 2, 3, 4)$ as the multiplexing waveforms, the histogram of $e_{2i}[n]$ and the ideal Gaussian distribution for $\sigma_u^2/N_0 = 20\text{dB}$ and $f_D T_0 = 0.025$ are plotted in Fig. C.2.

The moments comparison and the square error d^2 are listed in Table C.3.

It can be observed from Figs. C.2, C.2 and Tables C.1,C.2, C.3 that the approximation of $e_{\ell m}[n]$ as a Gaussian distribution with zero mean and variance $\sigma_{\ell m}^2$ is accurate.

	i=1	i=2	i=3	i=4	d^2
$\hat{m}_i$	5.858e-005	2.262e-003	-1.165e-006	1.567e-005	4.785e-004
m_i	0.000e+000	2.256e-003	0.000e+000	1.527e-005	
$\|\hat{m}_i - m_i\|$	5.858e-005	5.394e-006	1.165e-006	3.932e-007	

Table C.1: Estimated and ideal moments of e_{11} in correspondence to Fig. C.1(a).

	i=1	i=2	i=3	i=4	d^2
$\hat{m}_i$	7.029e-004	6.162e-002	1.786e-004	1.151e-002	3.781e-005
m_i	0.000e+000	6.162e-002	0.000e+000	1.139e-002	
$\|\hat{m}_i - m_i\|$	7.029e-004	3.061e-007	1.786e-004	1.197e-004	

Table C.2: Estimated and ideal moments of e_{11} in correspondence to Fig. C.1(b).

	i=1	i=2	i=3	i=4	d^2
$\hat{m}_i$	1.973e-004	7.066e-003	5.916e-006	1.503e-004	1.624e-004
m_i	0.000e+000	7.051e-003	0.000e+000	1.491e-004	
$\|\hat{m}_i - m_i\|$	1.973e-004	1.567e-005	5.916e-006	1.213e-006	

Table C.3: Estimated and ideal moments of e_{21} in correspondence to Fig. C.2

Appendix D

Approximation of $e_{f,\ell m}[n]$ by a Gaussian distribution

Re-write the definition of $e_{f,\ell m}[n]$ here for convenience,

$$e_{f,\ell m}[n] = \mathrm{Re}\left\{\sum_k \left(e_f[k]x_{01}[k] + \nu_f[k]\right) f_{\ell m}[k - 2^\ell n]\right\},$$

where $e_f[n]$ is the channel distortion estimation error, which is Gaussian distributed, $x_{01}[n]$ is reconstructed from $x_{\ell m}[n] = \pm 1$ by wavelet packet reconstruction filter, and $\nu_f[n]$ is

$$\nu_f[n] = \nu[n] + u_2[n] \sum_{k \neq 0} x_{01}[n-k] R_\phi(kT_0 - \tau) \tag{D.1}$$

The random variables $e[n]$, $x_{01}[n]$, and $\nu_f[n]$ are all independent of each other, but $\nu_f[n]$ is a correlated sequence.

Though $e_{f,\ell m}[n]$ is the sum of correlated random variables, it can still be approximated by Gaussian distribution [71] with

$$E\left\{e_{f,\ell m}[n]\right\} = 0$$

and

$$\sigma^2_{f,\ell m} = E\{e^2_{f,\ell m}[n]\}$$

$$= \frac{1}{2}\left(\sigma_e^2 E\{x_{01}^2[n]\} + \sigma_\nu^2\right) \sum_k f_{\ell m}^2[k - 2^\ell n]$$
$$+\frac{1}{2}\sigma_{u_2}^2 E\{x_{01}^2[n]\} \cdot$$
$$\sum_k \sum_j \tilde{R}_{u_2}((k-j)T_0) \sum_{i \neq 0} R_\phi(i-\tau) R_\phi(j-k+i-\tau) f_{\ell m}[k-2^\ell n] f_{\ell m}[j-2^\ell n].$$

Since $\sigma_{f,\ell m}^2$ is a function of $f_{\ell m}$, different terminal (ℓm) would have different Gaussian distributions. In the appendix, the histograms are obtained for each terminal of WPDM.

Now we demonstrate the accuracy of the approximation.

We generate a long sequence of $e_{f,\ell m}[n]$ in the following procedures.

- Generate a long sequence of Gaussian distributed $\mathrm{Re}\{e_f[n]\}$ with zero mean and variance $\frac{1}{2}\sigma_{e_f}^2$.
- Generate a long sequence of Gaussian distributed $\mathrm{Re}\{\nu[n]\}$ with zero mean and variance $\frac{1}{2}\sigma_\nu^2$.
- Generate $x_{01}[n]$ from $x_{\ell m} = \pm 1$ by wavelet packet reconstruction filter. The sequence $x_{01}[n]$ depends on the wavelet packet chosen in WPDM.
- Compute $R_\phi(kT_0 - \tau)$.
- Generate Gaussian distributed slowly varying process $\mathrm{Re}\{u_2[n]\}$ with zero mean and variance $\frac{1}{2}\sigma_{u_2}^2$.
- Generate $\mathrm{Re}\{\nu_f[n]\}$ by Eq. (D.1).
- To get $e_{f,\ell m}[n]$, compute $\sum_k \left(\mathrm{Re}\{e_f[k]\} x_{01}[k] + \mathrm{Re}\{\nu_f[k]\}\right) f_{\ell m}[k - 2^\ell n]$.

For the experimentally generated $e_{f,\ell m}[n]$, a histogram can be obtained. Comparing the histogram with the assumption of the pdf of $e_{f,\ell m}[n]$, we can see heuristically how accurate the approximation is. To have a more accurate quantitative comparison, we estimate the up to 4th moments of the experimentally generated $e_{f,\ell m}[n]$ by

$$\begin{aligned}
\hat{m}_1 &= \frac{1}{L_0}\sum_{i=0}^{L_0} e_{f,\ell m}[i] \\
\hat{m}_2 &= \frac{1}{L_0}\sum_{i=0}^{L_0} \left(e_{f,\ell m}[i] - \hat{m}_1\right)^2 \\
\hat{m}_3 &= \frac{1}{L_0}\sum_{i=0}^{L_0} \left(e_{f,\ell m}[i] - \hat{m}_1\right)^3 \\
\hat{m}_4 &= \frac{1}{L_0}\sum_{i=0}^{L_0} \left(e_{f,\ell m}[i] - \hat{m}_1\right)^4 .
\end{aligned}$$

The difference between $\hat{m}_i (i = 1,2,3,4)$ and the corresponding moments $m_i$of Gaussian distribution characterizes the accuracy of the approximation. Smaller difference indicates more accurate approximation. The calculation of m_i can be found in Appendix C.

Again, we use the square error

$$d^2 = \sum_i \left(\bar{y}_i - p_{e_{f,\ell m}}(b_i)\right)^2 b_s$$

to be a measure of the approximation accuracy.

First we assume $\phi_{11}(t)$ and $\phi_{12}(t)$ of Daubechies wavelet of order 14 are used for carrying the the information sequence in WPDM. In Figs. D.1 and D.2, histograms of $e_{f,11}[n]$ and $e_{f,12}[n]$ and the respective Gaussian distribution are shown for various combination of $N_0, f_D T_0, \tau$, and main-to-second path signal power ratio $\sigma_{u_1}^2/\sigma_{u_2}^2$.

The variance of $e_f[n]$ are computed from Eq. (6.46). In Tables D.1 and D.2, we list the estimated moments $\hat{m}_i$, the moments m_i of Gaussian distribution, and their difference.

For the WPDM employing $\phi_{2i}(t), (i = 1,2,3,4)$ as the multiplexing waveforms, the histogram of $e_{f,2i}[n]$ and the ideal Gaussian distribution for $\sigma_{u_1}^2/N_0 = 15$dB, $f_D T_0 = 0.1$, $\tau = 0.1T_0$, and $\sigma_{u_1}^2/\sigma_{u_2}^2 = 0$dB are plotted in Fig. D.3.

The moments comparison are listed in Table D.3.

		i=1	i=2	i=3	i=4	d^2
$e_{f,11}$	$\hat{m}_i$	1.165e-003	9.986e-002	3.648e-004	3.032e-002	9.379e-005
	m_i	0.000e+000	1.006e-001	0.000e+000	3.038e-002	
	$\lvert\hat{m}_i - m_i\rvert$	1.165e-003	7.778e-004	3.648e-004	6.680e-005	
$e_{f,12}$	$\hat{m}_i$	1.736e-019	1.312e-001	-2.548e-004	4.979e-002	5.222e-004
	m_i	0.000e+000	1.312e-001	0.000e+000	5.168e-002	
	$\lvert\hat{m}_i - m_i\rvert$	1.736e-019	2.838e-005	2.548e-004	1.887e-003	

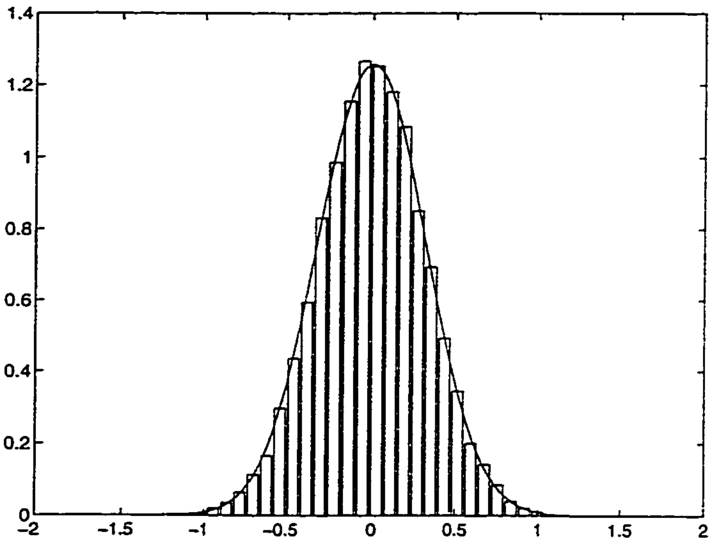
Table D.1: Estimated and ideal moments of e_{1i} in correspondence to Fig. D.1

		i=1	i=2	i=3	i=4	d^2
$e_{f,11}$	$\hat{m}_i$	3.973e-004	1.094e-002	-6.025e-006	3.657e-004	5.746e-004
	m_i	0.000e+000	1.102e-002	0.000e+000	3.643e-004	
	$\lvert\hat{m}_i - m_i\rvert$	3.973e-004	8.230e-005	6.025e-006	1.369e-006	
$e_{f,12}$	$\hat{m}_i$	6.980e-020	1.492e-002	3.079e-006	6.458e-004	1.671e-003
	m_i	0.000e+000	1.495e-002	0.000e+000	6.708e-004	
	$\lvert\hat{m}_i - m_i\rvert$	6.980e-020	3.698e-005	3.079e-006	2.492e-005	

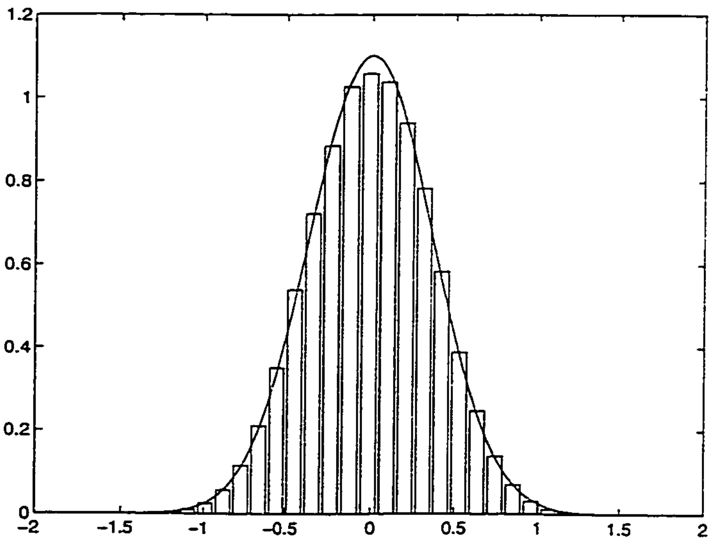
Table D.2: Estimated and ideal moments of e_{1i} in correspondence to Fig. D.2

		i=1	i=2	i=3	i=4	d^2
$e_{f,21}[n]$	$\hat{m}_i$	2.194e-003	2.658e-002	2.364e-004	2.169e-003	5.722e-004
	m_i	0.000e+000	2.670e-002	0.000e+000	2.139e-003	
	$\lvert\hat{m}_i - m_i\rvert$	2.194e-003	1.168e-004	2.364e-004	3.006e-005	
$e_{f,22}[n]$	$\hat{m}_i$	3.579e-019	3.282e-002	5.853e-005	3.214e-003	1.369e-004
	m_i	0.000e+000	3.318e-002	0.000e+000	3.303e-003	
	$\lvert\hat{m}_i - m_i\rvert$	3.579e-019	3.586e-004	5.853e-005	8.947e-005	
$e_{f,23}[n]$	$\hat{m}_i$	1.357e-019	3.414e-002	1.599e-006	3.492e-003	2.515e-004
	m_i	0.000e+000	3.427e-002	0.000e+000	3.524e-003	
	$\lvert\hat{m}_i - m_i\rvert$	1.357e-019	1.298e-004	1.599e-006	3.198e-005	
$e_{f,24}[n]$	$\hat{m}_i$	6.983e-020	4.101e-002	-2.938e-005	4.871e-003	3.954e-004
	m_i	0.000e+000	4.151e-002	0.000e+000	5.169e-003	
	$\lvert\hat{m}_i - m_i\rvert$	6.983e-020	4.982e-004	2.938e-005	2.975e-004	

Table D.3: Estimated and ideal moments of e_{2i} in correspondence to Fig. D.3

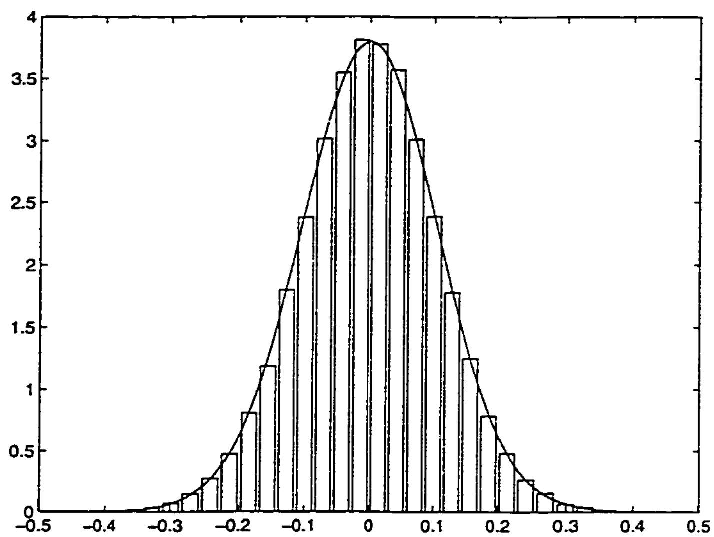


(a) $e_{f,11}[n]$

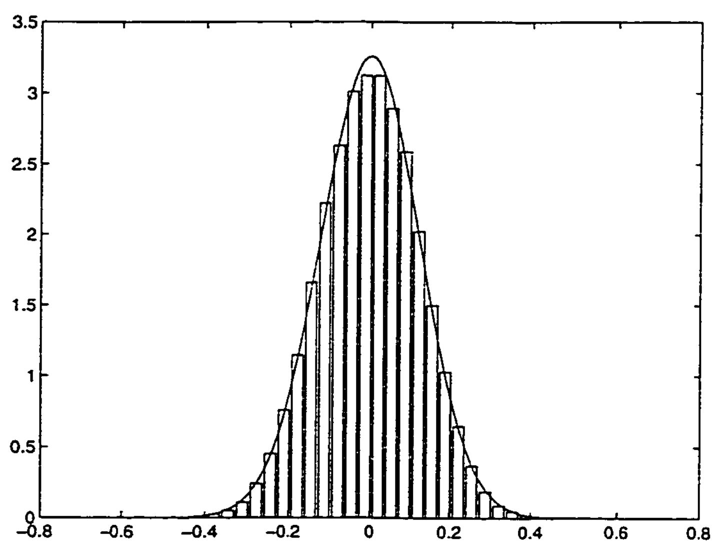


(b) $e_{f,12}[n]$

Figure D.1: Histogram of $e_{f,1i}[n]$ and ideal Gaussian distribution. $f_D T_0 = 0.025$, $\sigma_{u_1}^2/N_0 =$ 10dB, the relative delay of the second path $\tau = 0.2$, and $\sigma_{u_1}^2/\sigma_{u_2}^2 = 0$dB.

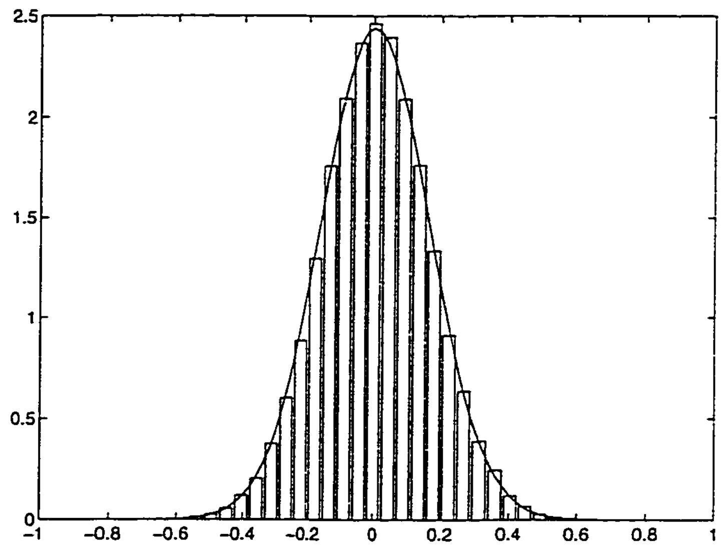


(a) $e_{f,11}[n]$

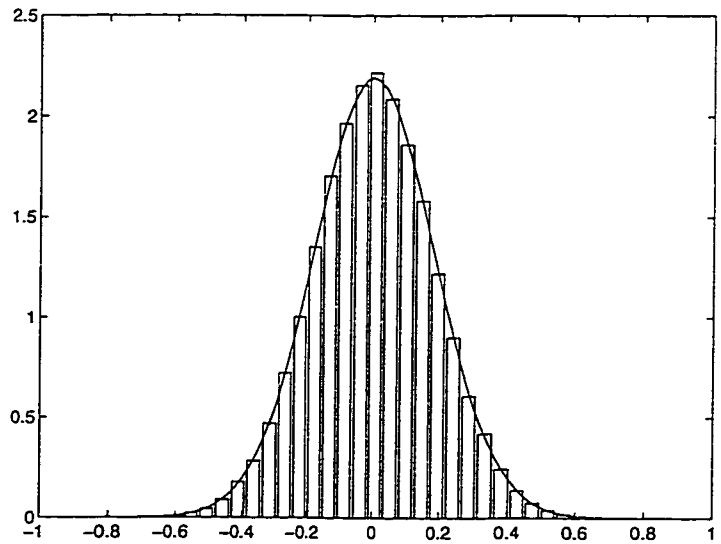


(b) $e_{f,12}[n]$

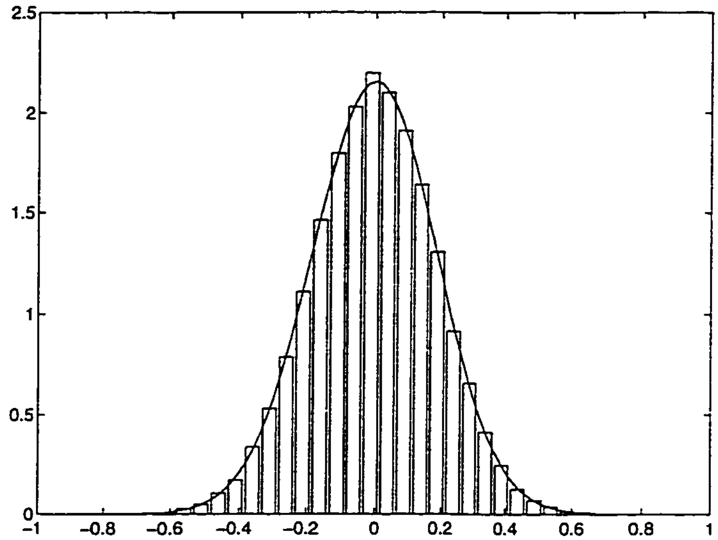
Figure D.2: Histogram of $e_{f,1i}[n]$ and ideal Gaussian distribution. $f_D T_0 = 0.025$, $\sigma_{u_1}^2 / N_0 =$ 20dB, the relative delay of the second path $\tau = 0.1$, and $\sigma_{u_1}^2 / \sigma_{u_2}^2 = 3$dB.



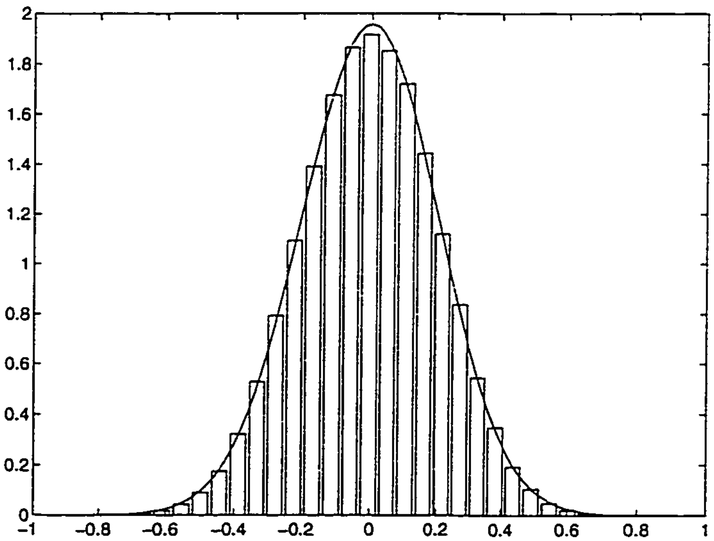
(a) $e_{f,21}[n]$



(b) $e_{f,22}[n]$



(c) $e_{f,23}[n]$



(d) $e_{f,24}[n]$

Figure D.3: Histogram of $e_{f,2i}[n]$ and ideal Gaussian distribution. $f_D T_0 = 0.025$, $\sigma_{u_1}^2/N_0 =$ 15dB, the relative delay of the second path $\tau = 0.1T_0$, and $\sigma_{u_1}^2/\sigma_{u_2}^2 = 0$dB.

It can be observed from Figs. D.1, D.2, D.3, and Tables D.1, D.2, D.3 that the approximation of $e_{f,\ell m}[n]$ as a Gaussian distribution with zero mean and variance $\sigma^2_{f,\ell m}$ is not as accurate as the approximation of $e_{\ell m}[n]$ in Appendix C, especially when $\phi_{11}(t)$ and $\phi_{12}(t)$ are used. This is an expected result since $e_{f,\ell m}[n]$ is the sum of a correlated sequence, and on the other hand $e_{\ell m}[n]$ is the sum of an independent sequence. However, the approximation is reasonably accurate to be adopted.

Bibliography

[1] B. Sklar, *Digital Communications. Fundamentals and Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1988.

[2] L. E. Franks, *Signal Theory.* Englewood Cliffs, NJ: Prentice-Hall, revised ed., 1969.

[3] S. Haykin, *Communication Systems.* New York: John Wiley & Sons, third ed., 1994.

[4] B. Hirosaki, "An orthogonally multiplexed QAM system using the discrete Fourier transform," *IEEE Transactions on Communications*, vol. 29, pp. 982–989, July 1981.

[5] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *The Bell System Technical Journal*, vol. 45, pp. 1775–1796, Dec. 1966.

[6] R. W. Chang and R. A. Gibbey, "A theoretical study of performance of an orthogonal multiplexing data transmission scheme," *IEEE Transactions on Communication Technology*, vol. 16, pp. 529–540, Aug. 1968.

[7] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Transactions on Communication Technology*, vol. 19, pp. 628–634, Oct. 1971.

[8] J. Chow, J. Tu, and J. Coiffi, "A discrete multitone transceiver system for HDSL applications," *IEEE Journal on Selected Areas in Communications*, vol. 9, pp. 895–908, Aug. 1991.

[9] J. P. Linnartz, "Special issue on multi-carrier modulation." *http://diva.eecs.berkeley.edu/ linnartz/MCCDMA.html.*

[10] A. Vahlin and N. Holte, "OFDM for broadcasting in presence of analogue co-channel interference," *IEEE Transactions on Broadcasting*, vol. 41, pp. 89–93, Sept. 1995.

[11] K. M. Wong, J. Wu, Q. Jin, and T. N. Davidson, "Wavelet packet division multiplexing and wavelet packet design under timing error effects," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2877–2890, Dec. 1997.

[12] J. Wu, K. M. Wong, and Q. Jin, "Wavelet packet division multiplexing," CRL Research Report 310, Communications Research Laboratory, McMaster University, Hamilton. Ontario, Canada, Sept. 1995.

[13] J. Wu, Q. Jin, and K. M. Wong, "Multiplexing based on wavelet packets," in Szu [82], pp. 315–326.

[14] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*. vol. 11. pp. 674–693, July 1989.

[15] A. N. Akansu and M. J. T. Smith, eds., *Subband and Wavelet Transforms. Design and Applications.* Boston: Kluwer Academic Publishers, 1996.

[16] M. A. Tzannes and M. C. Tzannes, "Bit-by-bit channel coding using wavelets," in *Proceedings of the IEEE Global Telecommunications Conference*, pp. 684–688, 1992.

[17] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Transactions on Information Theory (Special Issue on Wavelet Transforms and Multiresolution Signal Analysis)*, vol. 38, pp. 713–718, Feb. 1992.

[18] P. P. Ghandi, S. S. Rao, and R. S. Pappu, "On waveform coding using wavelets," in *Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA), pp. 901–905, 1993.

[19] R. E. Learned, H. Krim, B. Claus, A. S. Willsky, and W. C. Karl, "Wavelet-packet-based multiple access communication," in *Wavelet Applications in Signal and Image Processing II* (A. F. Laine and M. A. Unser, eds.), vol. 2303 of *Proceedings of the SPIE*, pp. 246–259, Oct. 1994.

[20] S. D. Sandberg and M. A. Tzannes, "Overlapped discrete multitone modulation for high speed copper wire communications," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1571–1585, Dec. 1995.

[21] K. Hetling, G. J. Saulnier, and P. K. Das, "Optimized perfect reconstruction quadrature mirror filter (PR-QMF) based codes for multi-user communications," in Szu [82], pp. 248–259.

[22] R. S. Orr, C. Pike, and M. J. Lyall, "Wavelet transform domain communication systems," in Szu [82], pp. 271–282.

[23] A. R. Lindsey, "Multi-dimensional signalling via wavelet packets," in Szu [82], pp. 303–314.

[24] G. W. Wornell, "Emerging applications of multirate signal processing and wavelets in digital communications," *Proceedings of the IEEE (Special Issue on Wavelets)*, vol. 84, pp. 586–603, Apr. 1996.

[25] J. Leonard J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency-division multiplexing," *IEEE Transactions on Communications*, vol. 33, pp. 665–675, July 1985.

[26] K. M. Wong, J. Wu, , T. N. Davidson, and Q. Jin, "Performance of wavelet packet division multiplexing in impulsive and gaussian noise," 1997. Submitted to the *IEEE Transactions on Communications*.

[27] J. Wu, K. M. Wong, Q. Jin, and T. N. Davidson, "Performance of wavelet packet division multiplexing in impulsive and Gaussian noise channels," in *Wavelet Applications in Signal and Image Processing IV* (M. A. Unser, A. Aldroubi, and A. F. Laine, eds.), vol. 2825 of *Proceedings of the SPIE*, pp. 516–527, Oct. 1996.

[28] J. Wu, K. M. Wong, and Q. Jin, "Performance of wavelet packet division multiplexing in timing errors and flat fading channels," in *Proceedings of the 3rd International Workshop on Image and Signal Processing* (B. G. Mertzios and P. Liatsis, eds.), (Manchester, England), pp. 381–384, Elsevier Science, Amsterdam, Nov. 1996.

[29] P. A. Bello and R. Esposito, "A new method for calculating probabilities of errors due to impulsive noise," *IEEE Transactions on Communication Technology*, vol. COM–17, pp. 368–379, June 1969.

[30] P. A. Bello and R. Esposito, "Error probabilities due to inpulsive noise in linear and hard-limited DPSK systems," *IEEE Transactions on Communication Technology*, vol. COM–19, pp. 14–20, Feb. 1971.

[31] S. Oshita and K. Feher, "Performance of coherent M-ary PSK sytems in an impulsive and gaussian noise environment," *IEEE Transactions on Communications*, vol. COM–30, pp. 2458–2464, Nov. 1982.

[32] S. Oshita and K. Feher, "Performance of coherent PSK and DPSK sytems in an impulsive and gaussian noise environment," *IEEE Transactions on Communications*, vol. COM–30, pp. 2540–2546, Dec. 1982.

[33] Y. Meyer, *Ondelettes*, vol. 1 of *Ondelettes et opérateurs*. Paris: Hermann, 1990. (In French. An English translation, by D. H. Salinger, appears as: Y. Meyer. *Wavelets and Operators*. Cambridge University Press, Cambridge, England, 1992).

[34] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: SIAM, 1992.

[35] C. K. Chui, *An Introduction to Wavelets*. Boston: Academic Press, 1992.

[36] A. N. Akansu and R. A. Haddad, *Multiresolution Signal Decomposition. Transforms, Subbands, and Wavelets*. Boston: Academic Press, 1992.

[37] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[38] G. G. Walter, *Wavelets and Other Orthogonal Systems with Applications*. Boca Raton, FL: CRC Press, 1994.

[39] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[40] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge Press, 1996.

[41] G. Strang, "Wavelets and dilation equations: A brief introduction," *SIAM Review*, vol. 31, pp. 614–627, Dec. 1989.

[42] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Magazine*, vol. 8, pp. 14–38, Oct. 1991.

[43] B. Jawerth and W. Sweldens, "An overview of wavelet based multiresolution analysis," *SIAM Review*, vol. 36, pp. 377–412, Sept. 1994.

[44] A. Cohen and J. Kovačević, "Wavelets: The mathematical background," *Proceedings of the IEEE (Special Issue on Wavelets)*, vol. 84, pp. 514–522, Apr. 1996.

[45] N. Hess-Nielsen and M. V. Wickerhauser, "Wavelets and time-frequency analysis," *Proceedings of the IEEE (Special Issue on Wavelets)*, vol. 84, pp. 523–540, Apr. 1996.

[46] K. Ramchandran, M. Vetterli, and C. Herley, "Wavelets, subband coding, and best bases," *Proceedings of the IEEE (Special Issue on Wavelets)*, vol. 84, pp. 541–560, Apr. 1996.

[47] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[48] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. XLI, pp. 909–996, Oct. 1988.

[49] A. V. Oppenheim, A. S. Willsky, and I. T. Young, *Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[50] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, third ed., 1995.

[51] E. A. Lee and D. G. Messerschmitt, *Digital Communications*. Boston: Kluwer Academic Publishers, second ed., 1993.

[52] J. A. C. Bingham, "Multicarrier modulation for data transsmission: An idea whose time has come," *IEEE Communications Magazine*, vol. 28, pp. 5–14, May 1990.

[53] C. Sundberg, "Continous phase modulation: a class of jointly power and bandwidth efficient digital modulation schemes with constant amplitude," *IEEE Communications Magazine*, vol. 24, pp. 25–38, Apr. 1986.

[54] F. Xiong, "Modem techniques in satellite communications," *IEEE Communications Magazine*, pp. 84–98, Aug. 1994.

[55] L. E. Franks, "Carrier and bit synchronization in data communications — A tutorial review," *IEEE Transactions on Communications*, vol. COM–28, pp. 1107–1121, Aug. 1980.

[56] M. H. Meyers and L. E. Franks, "Joint carrier phase and symbol timing recovery for PAM systems," *IEEE Transactions on Communications*, vol. COM–28, pp. 1121–1129, Aug. 1980.

[57] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE Signal Processing Magazine*, vol. 8, pp. 14–36, Apr. 1991.

[58] W. A. Gardner, "The role of spectral correlation in design and performance analysis of synchronizer," *IEEE Transactions on Communications*, vol. 34, pp. 1089–1095, Nov. 1986.

[59] W. Kaplan, *Advanced Calculus*. Reading, MA: Addison-Wesley, 1952.

[60] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*. Reading, MA: Addison-Wesley, second ed., 1994.

[61] N. Saito and G. Beylkin, "Multiresolution representations using the autocorrelation functions of compactly supported wavelets," *IEEE Transactions on Signal Processing (Special Issue on Wavelets and Signal Processing)*, vol. 41, pp. 3584–3590, Dec. 1993.

[62] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Transactions on Signal Processing*, vol. 40, pp. 2207–2232, Sept. 1992.

[63] I. Daubechies, "Orthonormal bases of compactly supported wavelets II. Variations on a theme," *SIAM Journal on Mathematical Analysis*, vol. 24, pp. 499–519, Mar. 1993.

[64] "Matlab version 4.2 (software)," *The MathWorks Corp.*, 1994.

[65] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization.* London: Academic Press, 1981.

[66] A. Grace, "Optimization toolbox user's guide," *The MathWorks Corp.*, 1994.

[67] P. G. Lemarié, "Ondelettes à localisation exponentielle," *Journal de Mathématiques Pures et Appliquées*, vol. 67, no. 3, pp. 227–236, 1988.

[68] G. Battle, "A block spin construction of ondelettes. Part I: Lemarié functions," *Communications in Mathematical Physics*, vol. 110, no. 4, pp. 601–615, 1987.

[69] R. L. Freeman, *Telecommunication System Engineering.* John Wiley and Sons, 1989.

[70] R. L. Freeman, *Practical Data Communications.* John Wiley and Sons, 1995.

[71] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems.* Plenum Press, 1994.

[72] A. Papoulis, *Probability, Random Variable, and Stochastic Processes.* New York: McGraw-Hill, third ed., 1991.

[73] J. W. C. Jakes, ed., *Microwave Mobile Communications.* New York: John Wiley & Sons, 1974.

[74] H. V. Tree, *Detection, Estimation, and Modulation. Part I.* New York: Wiley, 1968.

[75] J. K. Cavers, "Pilot symbol assisted modulation and differential detection in fading and delay spread," *IEEE Transactions on Communications*, vol. COM–43, pp. 2206–2212, July 1995.

[76] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. VT–40, pp. 686–693, Nov. 1991.

[77] C.-L. Liu and K. Feher, "Pilot-symbol aided coherent M-ary PSK in frequency-selective fast Rayleigh fading channels," *IEEE Transactions on Communications*, vol. COM–42, pp. 54–62, Jan. 1994.

[78] T. S. Rappaport, *Wireless Communications Principle and Practice*. Upper Saddle River, NJ: Prentice Hall, 1996.

[79] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. XLV, pp. 485–560, June 1992.

[80] A. Cohen and I. Daubechies, "On the instability of arbitrary biorthogonal wavelet packets," *SIAM Journal on Mathematical Analysis*, vol. 24, pp. 1340–1354, Sept. 1993.

[81] K. M. Wong and S. Chen, "Detection of narrow-band sonar signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 597–613, May 1987.

[82] H. H. Szu, ed., *Wavelet Applications II*, vol. 2491 of *Proceedings of the SPIE*, Apr. 1995.