NOTICE

AVIS

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Error Performance of PSK Signals Transmitted Over Correlated Rayleigh Fading Channels

by

Dominic Kwing-pang Fung

B.Sc.E.E., University of Manitoba, 1989

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE (ENGINEERING SCIENCE)

in the School

of

Engineering Science

© Dominic Kwing-pang Fung 1991

Simon Fraser University

July, 1991

Canada

APPROVAL

NAME: Dominic Kwing-pang Fung

DEGREE: Master of Applied Science (Engineering Science)

TITLE OF THESIS: **Error Performance of PSK Signals Transmitted Over Correlated Rayleigh Fading Channels**

EXAMINING COMMITTEE:

Chairman: Dr. John Jones

Dr. Paul Ho
Senior Supervisor

Dr. James Cavers
Supervisor

Dr. John Bird
Examiner

DATE APPROVED: July 16, 1991

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

"Error Performance of PSK Signals Transmitted Over Correlated Rayleigh

Fading Channels"

Author:

        (signature)

        Kwing-Pang (Dominic) FUNG

        (name)

        July 16, 1991

        (date)

ABSTRACT

Phase shift keying (PSK) is a class of constant envelope modulation which has become one of the most popular techniques for mobile communication. The mobile radio channel is characterized by Rayleigh fading in which the complex gains experienced by consecutive transmitted symbols are correlated. Various channel estimation techniques have been proposed in the literatures to combat correlated Rayleigh fading. In this thesis, a general analysis of the error performance of PSK modulations with different detection schemes in correlated Rayleigh fading channels is provided. Specifically, we derive the optimal receiver for the demodulation of N PSK symbols given M channel state estimates. Subsequently, an exact expression for the pairwise error event probability of this receiver is obtained. The results are applied to study the error performance of pilot symbol assisted modulation (PSAM), multiple symbol differential detection, and interleaved, Trellis-coded PSK modulations. The bit error rates for the above modulation schemes at various normalized Doppler frequencies have been obtained from the analytical expressions. For bit error rate above 10^{-5}, it is found that both PSAM and multiple symbol differential detection can eliminate the irreducible error floor commonly associated with conventional differential detection. In the case of interleaved, Trellis-coded differential PSK, it is observed that full interleaving does not necessarily provide the best error performance.

To God

ACKNOWLEDGMENTS

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AWGN	Additive white Gaussian noise
BPSK	Binary phase shift keying
CCITT	International Telephone and Telegraph Consultative Committee
CSI	Channel state information
dB	Decibel
DPSK	Differential phase shift keying
FEC	Forward error correction or forward error correcting codes
FIR	Finite Impulse Response
Hz	Hertz
iid	independent and identically distributed
ISI	Inter-symbol interference
kHz	kiloHertz
Ksymbol/s	kilo symbol per second
MHz	megaHertz
MPSK	M-ary phase shift keying
MLSE	Maximum likelihood sequence estimation
PSAM	Pilot symbol assisted modulation
PSK	Phase shift keying
QAM	Quadrature amplitude modulation
rms	Root mean square
RPpoles	Poles on the right-half of the complex s-plane
SNR	Signal to noise ratio
TCM	Trellis coded modulation

VARIABLES AND FUNCTIONS

$\mathbf{c}$	Transmitted codeword
$\hat{\mathbf{c}}$	Erroneous codeword
$d(\mathbf{c}, \hat{\mathbf{c}})$	Hamming distance between $\mathbf{c}$ and $\hat{\mathbf{c}}$
f_D	Maximum Doppler frequency
$f_D T$	Normalized maximum Doppler frequency or fade rate
$g(t)$	Channel fading process
g_k	A sample of the fading process
$n_w(t)$	Channel AWGN
$p(t)$	Impulse response of a pulse shaping filter
$s(t)$	Baseband transmitted signal
s_k	Transmitted symbol in the interval $kT \le t \le (k+1)T$
$r(t)$	Baseband received signal
t	Time
u_k	Fading gain corresponding to the received symbol
v_k	Fading gain corresponding to the channel state information
x_k	Channel state information in the kth interval
y_k	Received symbol in the kth interval
z_k	Differentially encoded symbol
A	Pulse amplitude
E_b	Bit energy
E_s	Symbol energy
$E\{\bullet\}$	Statistical average
$J_o(\bullet)$	Bessel function of order zero
M	Length of channel state sequence
N	Length of codeword
P_b	Bit error probability
$P(\mathbf{c} \to \hat{\mathbf{c}})$	Pairwise error event probability
$R_g(\tau)$	Autocorrelation function of the fading process
T	Symbol interval
α	Interleaving span
β	Interleaving depth
θ_k	Transmitted phase in the kth interval

λ_k	The kth eigenvalue
μ	Correlation coefficient
σ^2	variance
$\phi(i,j)$	correlation between fading gains u_i and u_j
Φ	Covariance matrix
$\Phi_D(s)$	Characteristic function

CHAPTER 1

INTRODUCTION

For the past ten years, digital transmission over mobile fading channels has become an important research area. A major concern to the mobile radio system engineers is the error performance of the communication systems. In general, a moving vehicle does not have a direct line of sight path to the transmitter. The received signal is the net resultant of many signals that reach the vehicle via multiple paths. Thus, the received signal at different locations consists of a number of waves whose amplitudes, phases, and angles of arrival are random. Under these conditions, the short-term amplitudes statistics of the resultant signal tends to be Rayleigh distributed [2]. In general, the Rayleigh fading experienced by consecutive transmitted symbols are correlated. This results in burst errors. For a coded system, a practical and effective approach for combating burst errors is to destroy the channel's memory through interleaving. Alternatively, one can exploit the correlation of the fading encountered by different transmitted symbols to improve the error performance of the system. This brings us to the central theme of the thesis:– the seeking of alternate

detection for digital modulations operating in correlated Rayleigh fading channels. Two techniques will be studied, both use the concept of block demodulation. The first technique, which is called multiple-symbol differential detection in the literature [11], requires no bandwidth expansion. The second technique, which is a refinement of the one in [6], requires a slight bandwidth expansion. For the bit error rate above 10^{-5}, both techniques are able to eliminate the irreducible error floors commonly associated with conventional detectors.

In conjunction to the study on alternate decoding strategies for correlated Rayleigh fading channels, we also considered the interleaver design issue for combined coding/interleaving systems. As mentioned earlier, this is a more common approach for combating fading. After Ungerboeck [32] proposed Trellis-coded Modulation (TCM) for the additive white Gaussian noise channel, there has been considerable interest in applying TCM to mobile fading channels. In recent years, numerous reports [8], [9], [10], [21], [5] showed that TCM, when combined with interleaving of sufficient depth, is able to provide good error performance in mobile fading channels. In many of these studies, the system was assumed to have ideal interleaving, or equivalently independent fading. This assumption is justified if the interleaving depth is "large" compared to the fade duration. However, in order to find out what a "large" interleaving depth is, one usually has to resort to time consuming computer simulation [21], [12], [17]. In this thesis, we will study analytically the error performance of TCM with finite interleaving, or equivalently correlated fading. The error probability will be given with the interleaving depth as a parameter.

1.1 Background and Literature Review

Telephone has long been accepted as an important communications tool in modern living. However, due to the fact that its use has been constrained by the connecting wire, the ultimate objective of communications— to enable anyone to communicate instantly with anyone else from anywhere, can be achieved only by mobile radio. Resulting from the success in semiconductor industry and the change in attitude of the radio regulatory authorities in making radio frequency spectrum available for commercial applications, mobile radio is no longer a luxury item. Nowadays, a mobile telephone costs less than a video recorder or even an in-car stereo unit and become an essential business tool for many people.

1.1.1 Conventional Mobile Radio Systems

During the past decade, broad applications of land mobile services have been authorized by regulatory agencies in the 800 and 900 MHz portions of the radio spectrum. Owing to the limited experience with the application of radio propagation to land mobile system in these bands, there are no procedures within the mobile engineering community that are as yet generally acknowledged or accepted as providing accurate and reliable 800/900 MHz propagation information [23]. Early in 1982, this constraint on engineering analysis of mobile radio operating in the 800/900 MHz frequency range has been identified by the IEEE as an area requiring immediate attention.

Mobile radio signals are affected by various factors such as propagation-path loss, multipath fading and Doppler effect resulting from the relatively high speed

3

of the vehicle. However, in urban areas, propagation between a mobile unit and a base station is most susceptible to the effect of multipath fading [16]. The term "multipath fading" is used to refer to the variation in the strength of a received radio carrier signal due to atmospheric changes and ground and water reflections in the propagation path as shown in Figure 1.1. Basically, there are two types of fade, *flat fades* and *frequency selective fades*. The former distorts the amplitude of the transmitted signal uniformly across the channel bandwidth while the latter attenuates the carrier signal unevenly across the band. In most cases, flat fades mainly affect the small-capacity digital radios and frequency selective fades are of major importance to the high-capacity digital radios. Throughout this thesis, only flat fades will be considered in the radio channel. Generally, with adequate path clearance and in the absence of a single specular reflection on a path, the amplitude of flat fades due to multipath propagation varies randomly with a Rayleigh distribution [16]. This kind of channel is commonly referred to as Rayleigh fading channels. When the signal falls below its statistical mean, a fade occurs that causes any digital data transmitted over the carrier to be corrupted with a noise burst [15]. For instance, a vehicle travels at 20 km/hr and sends data at 10 kbit/s with carrier frequency 850 MHz, the signal goes into a -15 dB fade at the rate of approximately six times a second. The probability that the duration of this fade is at least 8 ms is about 0.2. Thus, a block of 80 or more bits of data would be corrupted by noise once every 160 ms with a probability of 0.2.

There are several techniques to minimize the effects of multipath fading. Since the chance of having two deep fades from two uncorrelated signals at any instance is rare, the effect of the fades can be reduced by combining them. This can be achieved by space and frequency diversity. The space diversity scheme uses two (or more)
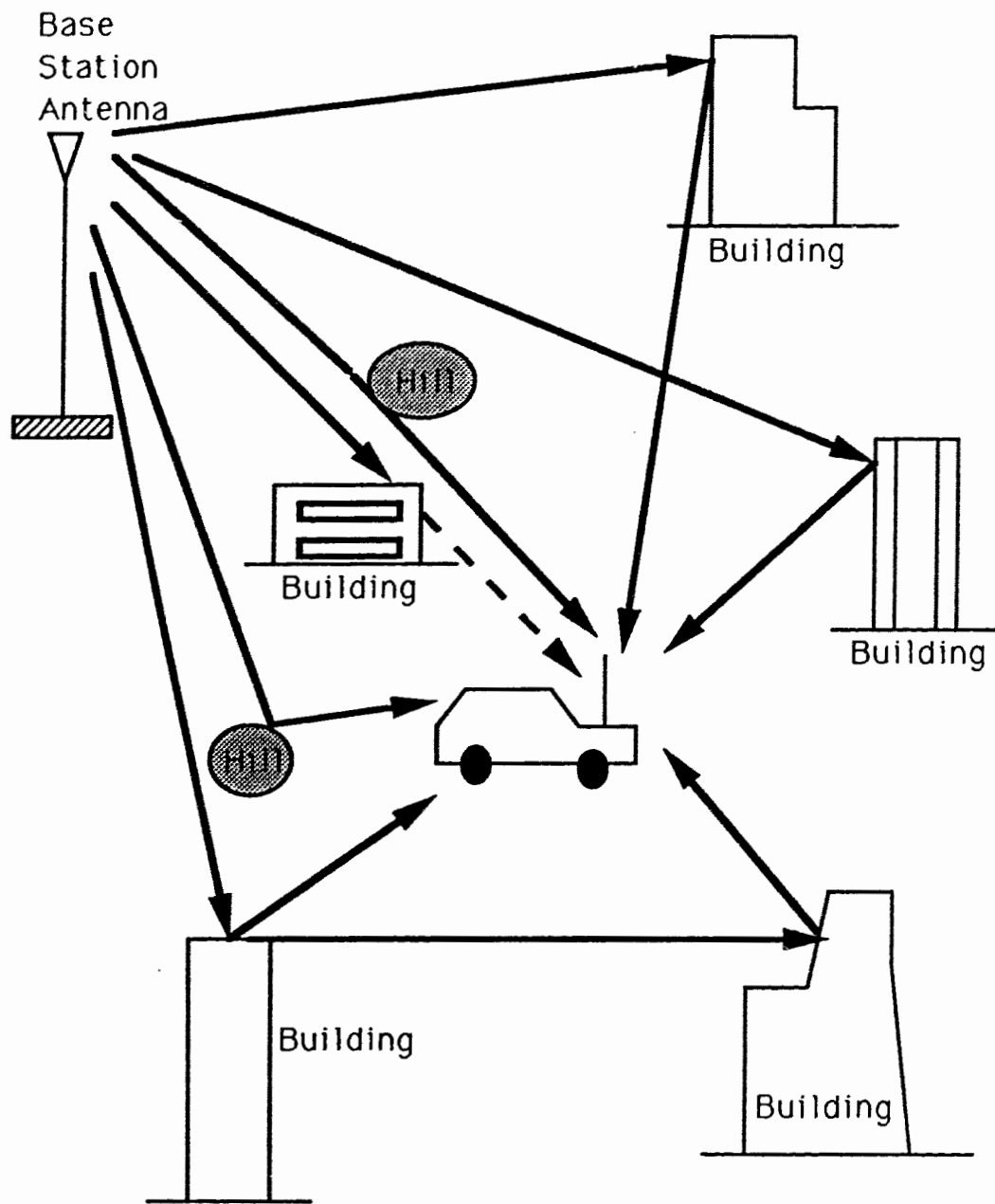
Figure 1.1: The Multipath fading phenomenon

antennas for receiving and/or transmitting the same signal simultaneously over a radio channel so that the individual signals received are uncorrelated. The frequency diversity scheme transmits the same signal simultaneously over two or more radio frequency channels which are located in the same frequency band. Although it has been shown [31] that both space diversity and frequency diversity can provide good improvement factors, their use is usually restricted due to the requirement of extra hardware or due to the inefficient use of the available frequency spectrum.

In coded systems, one simpler technique which requires minor hardware changes and is very robust to burst error, is time diversity, or *interleaving*. If all the symbols of a transmitted codeword are sent at widely spaced intervals and the intervening spaces are filled by symbols of other codewords, the statistical dependency between symbols can effectively be eliminated. An interleaver (see Figure 1.2) is a device which can rearrange the ordering of the encoded symbols so that two consecutive symbols are separated by certain spaces (interleaving depth) after interleaving. Theoretically, if the interleaver is able to break up all the burst errors caused by multipath fading, the individual symbols received will have independent fading gains and the channel is called a memoryless channel or a fully interleaved Rayleigh channel. The larger the interleaving depth is, the longer burst of error the system can handle. However, a drawback for having a large interleaving depth is the time delay which increases directly with the interleaving depth. Fortunately, there is some tolerance for this in the mobile radio environment. In particular, for the case of speech transmission, if the total coding/decoding delay is below 60 ms, it would not be noticeable to the listener [8]. Thus, decoder buffer and interleaving depth must be limited so as to produce at most a 60 ms delay. The Pan-European Digital Cellular system: Group Special Mobile (GSM), which will begin service in 1991, adopts interleaving

$$(c_1, c_{\alpha+1}, \ldots c_{(\beta-1)\alpha+1}, c_2, c_{\alpha+2}, \ldots)$$

c_1	$c_{\alpha+1}$	$\cdots$	$c_{(\beta-1)\alpha+1}$
c_2	$c_{\alpha+2}$	$\vdots$	$c_{(\beta-1)\alpha+2}$
c_3	$c_{\alpha+3}$	$\vdots$	$c_{(\beta-1)\alpha+3}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
c_α	$c_{2\alpha}$	$\vdots$	$c_{\alpha\beta}$

$$(c_1, c_2, c_3, \ldots)$$

Figure 1.2: Block interleaver

as the technique to combat the multipath fading in the system [35].

With the introduction of the interleaver[1] in the mobile communication system, burst error caused by amplitude fades of duration greater than one symbol time can be dispersed evenly on each codeword. In order to recover the original codeword, error correction technique must be applied subsequently. The ability to detect and/or correct errors in traditional forward error correction (FEC) coding is provided by the transmission of redundant bits, and thus, lowering the effective information rate per transmission bandwidth. In mobile radio applications, where both bandwidth and power limitations are imposed simultaneously, it is often not possible to just employ either bandwidth efficient modulation techniques or power efficient FEC techniques to enhance the system throughput. What is required is the integration of a bandwidth efficient modulation scheme with some form of FEC coding to exploit the best possible attributes of both.

In 1982, Ungerboeck of IBM (Zurich) [32] proposed a novel coding technique which treated channel coding and modulation as an entity and achieved a significant improvement in error performance without sacrificing data rate or requiring extra bandwidth. This technique is known as Trellis-coded Modulation. The term "Trellis" is used because this scheme can be described by a state transition (trellis) diagram similar to the trellis diagrams of binary convolutional codes [33]. Intuitively, signal waveforms representing information sequences are most impervious to noise if they are very different from each other. Mathematically, this is equivalent to the statement that the signal sequences should have large distance in the

---

[1]Interleaver must be used in conjunction with some kind of coding schemes.

Euclidean signal space. This leads to the most important new concept of TCM—to use signal-set expansion to provide redundancy for coding, and to design coding and signal-mapping functions jointly so as to maximize directly the minimum Euclidean distance (free distance) between coded signal sequences (increase the noise immunity). The resulting free distance of the coded sequences exceeds significantly the minimum distance between uncoded modulation signals at the same information rate, bandwidth, and signal power.

For example, consider the results in [32] where he compared his 8-state 8-PSK Trellis code with uncoded 4-PSK scheme. Both systems transmit two information bits per modulation interval. The free distance for uncoded 4-PSK is only 1.414 while the free distance for Trellis-coded 8-state 8-PSK is 2.141. Such an improvement in free distance resulted in code gain of 3.6 dB over uncoded 4-PSK in the additive white Gaussian noise channel (AWGN).

The most common application of TCM technique is in the new generation of modems being developed for the telephone channel. Prior to TCM, uncoded transmission at 9.6 kbit/s over voiceband channels was considered to be a practical limit for data modems. In 1984, a new generation of data modem which employed TCM along with other improvements in equalization and synchronization appeared on the market and was able to transmit data reliably over telephone lines at rates of 14.4 kbit/s and higher. In addition, TCM was also adopted by the CCITT for use in new high-speed voiceband modems [33]. Since the TCM improves the noise immunity of digital transmission systems without bandwidth expansion or reduction of data rate, it is ideal for application in the power and bandwidth limited mobile radio environment.

A couple of years after Ungerboeck had introduced the power and band-width efficient TCM technique, there were considerable research activities in studying the possibility of using TCM in mobile radio systems. McLane et al [21] studied the usefulness of rate 2/3, Trellis-coded 8-PSK, with both coherent (PSK) and differential detections (DPSK), for fast fading, shadowed mobile satellite communication channels. Their interest was mainly in speech transmission at a bit rate of 2400 bit/s and a bit error rate of 10^{-3}. A similar study on interleaved Trellis-coded 8-DPSK modulations transmitted over both Rayleigh and Rician fading channels was done by Edbauer [12]. More recently, Lee and McLane [17] repeated the study reported in [21] by replacing the block interleaver with a convolutional interleaver. The advantage of using a convolutional interleaver is that it requires only half the delay for the same bit error performance relative to the block interleaver. The common drawback in the above studies was that all the results were determined via digital computer simulation. Although simulation is capable of reflecting the actual system performance, it is a time consuming process. In addition, simulation studies cannot provide much insight into the understanding of the behavior of the system.

The first analytical result on Trellis-coded MPSK modulation transmitted over fading channel was reported by Divsalar and Simon [8] where they applied the Chernoff bound technique to obtain an upper bound on the pairwise error probability. By making use of the pairwise error probability bound and the transfer function of the pair-state transition diagram, an upper bound of the average bit error probability was obtained. Later on, Divsalar and Simon [9] used a similar technique and extended their analysis to include Trellis-coded multilevel differential phase shift keying (MDPSK). However, the upper bounds obtained by Divsalar and Simon were too loose over the normal range of signal to noise ratios (SNR) of interest.

10

Besides, the pair-state transition diagram approach may be a tedious task when the number of states in the trellis diagram becomes large. By using the characteristic function and the numerical Gauss-Konrod integration rule, McKay [20] was able to evaluate an exact pairwise error event probability for TCM in Rayleigh and Rician fading channels. With a simple modification to the standard transfer function bound, a new asymptotically tight upper bound for the bit error probability was derived. Although the results were satisfactory, the numerical evaluation of this upper bound is quite complicate. By employing the characteristic function and the residue theorem, Cavers and Ho [5] obtained an exact and easily computed expression for the pairwise error event probability of TCM operating in Rayleigh fading channels. This expression is quite general which includes not only Trellis-coded MPSK, but also Trellis-coded Quadrature Amplitude Modulation (QAM) with perfect channel state information (CSI), differential detection, or pilot tones. Accurate average bit error probabilities were obtained by considering only a small set of short error events. Among all of these analytical studies, the assumption of ideal interleaving, or equivalently independent fading were used and which is usually impractical for real systems.

In this thesis, we will study the error performance of Trellis-coded PSK modulations without the assumption of ideal interleaving. An exact expression for the pairwise error event probability of interleaved, coded PSK modulations is derived with the interleaving depth as one of the parameters. Both coherent as well as differential detection are considered. Generally, it is found that when the interleaving depth is equal to one-fifth to one-quarter the duration of a fade cycle (defined as the reciprocal of the normalized Doppler frequency), the error performance is almost as good as full interleaving.

1.1.2 Alternate Detection Strategy in Mobile Radio Systems

Since interleaving is employed only in coded systems, it is natural to ask: are there other ways of enhancing error performance for both coded and uncoded systems in correlated Rayleigh fading channels? By exploiting the statistics of the fading channels, two new detection schemes: multiple-symbol differential detection [19] and pilot symbol assisted modulations (PSAM) [26] were proposed.

In the simplest terms, a multiple-symbol differential detector is a decoder that makes a decision about a block of N consecutive PSK symbols based on $N + 1$ received samples. The first received sample is used to provide a phase reference for the entire block while the last sample is used to provide a reference for the next block. In the case when $N = 1$, a conventional differential detector is the result. The larger the value of N, the better the error performance [11].

A simulation study of the error performance of multiple-symbol differential detection of MPSK was done by Edbauer [12]. He found that relative to coherent 4-PSK, Trellis-coded 8-DPSK with a three-symbol detector achieves the same coding gain as Trellis-coded coherent 8-PSK in the AWGN channel. Recently, Divsalar and Simon [11] used the *maximum likelihood sequence estimation* technique to analyze the performance of multiple-symbol differential detection for uncoded MPSK signals in the AWGN channel. They demonstrated that the amount of improvement over conventional DPSK depends on the number of phases and the number of additional symbol intervals added to the observation. Moreover, with the addition of a few observation intervals, the error performance is approaching that of a coherent

detector.

The first investigation of multiple-symbol differential detection in the Rayleigh fading channel was shown in [19] where two or more differential detectors were jointly utilized to take advantage of the redundancy introduced by the differential encoder. Although the performance evaluation was performed via computer simulations, they found that just by using a 1-bit and a 2-bit differential detector jointly, the performance of Trellis-coded PSK schemes can be improved significantly. A more detailed study of multiple-symbol differential detection was reported in [18]. They considered two disturbances (AWGN and fading) separately and derived optimum decoders for each cases. Then, based on a linear combining approach, they combined the two decoders together and form the suboptimal decoder for the multiple-symbol detection system. Error performance was evaluated by computer simulation. For the Trellis-coded 4-PSK scheme, the proposed detection technique was able to eliminate the presence of error floors.

In this thesis, we present the true optimal multiple symbol differential detector for uncoded PSK modulation transmitted over correlated Rayleigh fading channels with AWGN. An exact expression for the pairwise error event probability of this detector is derived. As shown later, for the bit error rate above 10^{-5}, with only 2 additional observation intervals, this decoding strategy practically eliminates the irreducible error floor associated with a conventional differential detector. In addition, this decoding strategy is not very sensitive to the mismatch between the autocorrelation function of the channel fading process and the decoding metric.

Beside considering multiple symbol differential detection, we also consider block demodulation for pilot symbol assisted modulations. PSAM was first proposed by

Sampei and Sunaga in [26]. The concept of PSAM is similar to the conventional pilot tone estimator. The transmitter periodically inserts known symbols into the data sequence so that the receiver can make a good estimation of the channel state based on the pilot symbols. Unlike the conventional pilot signaling methods where a pilot signal is transmitted along with the data signal, PSAM does not require complex analog signal processing, such as frequency shifting and filtering. A simulation study of PSAM for 16QAM in the Rayleigh fading channel was reported in [26]. The hardware implementation of a PSAM modem was also outlined in [26]. From both the software and hardware simulation results, it was confirmed that PSAM can effectively suppress the error floor [26]. An analytical study of PSAM was shown in [6] where closed form results for the bit error rate of BPSK and 4-PSK, and a tight upper bound for 16QAM were obtained. The main difference between the work presented in this thesis and the one reported in [6] is that the former used block decoding strategy while the latter uses symbol by symbol detection. It is observed that block decoding of PSAM is more appropriate in a fast fading environment (e.g. 5% of the signaling rate).

1.2 Contributions of The Thesis

The major contributions of this thesis can be summarized as follows:

1. The derivation of the optimum block decoder for PSK symbols transmitted over correlated Rayleigh fading channels. The structure of this optimal decoder covers both multiple symbol-differential detection as well as PSAM.

2. The derivation of an exact expression for the pairwise error event probability for the optimal detectors in (1).

3. The derivation of an exact and easily computed expression for the pairwise error event probability of interleaved coded PSK modulations transmitted over correlated Rayleigh fading channel, with the interleaving depth and the normalized Doppler frequency as parameters.

1.3 Thesis Outline

In Chapter 2, a general description of the PSK modulation systems operating in correlated Rayleigh fading channels is given. The assumptions made in the study are clearly stated. Following that, the optimum block decoder and the corresponding pairwise error event probability is derived. In Chapter 3, the error expressions for multiple-symbol differential detection and pilot symbol assisted block demodulation are presented as examples for the general analysis in Chapter 2. Analytical results are shown subsequently. In Chapter 4, we will once again make use of the general results given in Chapter 2 to analyze the error performance of Trellis-coded PSK modulations with non-ideal interleaving. Results for both interleaved coded PSK modulations and DPSK modulations with Viterbi decoding are presented. Finally, conclusions of this study are drawn in Chapter 5.

CHAPTER 2

ANALYSIS OF PSK SIGNALS IN CORRELATED RAYLEIGH FADING CHANNEL

As stated in Chapter 1, one of our objectives is to study the error performance of PSK modulations transmitted over Rayleigh fading channel without the assumption of independent fading. In this chapter, we will firstly define the system notation and then derive the optimum block decoding metric for PSK signals operating in such an environment. Following that, an exact expression for the pairwise error event probability for the optimum decoder is derived.

2.1 Simplifying Assumptions and Definitions

In order to perform the analysis described in this thesis, a number of simplifying assumptions are made. First, the channel is assumed to be non-frequency selective fading channel (or flat fading channel) with AWGN. Second, the fading process is assumed to be slow enough that the channel complex gain is roughly constant over one symbol interval. Finally, the system is assumed to be able to perform perfect time synchronization. For coherent detection, perfect carrier frequency and phase recovery are also assumed.

Throughout this thesis, $E\{\bullet\}$ is used to represent a statistical average while $(\bullet)^*$ is used to represent complex conjugate. $(\bullet)^t$ is used to represent the transpose of a matrix or a vector while $(\bullet)^\dagger$ is used to represent the Hermitian transpose of a matrix.

2.2 System Description

The purpose of this chapter is to present a general mathematical analysis of PSK signals operating in correlated Rayleigh fading channel. Throughout this thesis, we will use the complex baseband notation. A comprehensive coverage of the relationship between baseband signal and bandpass signal can be found in Chapter 3 of [25].

Consider Figure 2.1, the input to the encoder is a sequence of binary digits and the output is a sequence of complex PSK symbols denoted by $c = (c_1, \ldots, c_k, \ldots)$. Note that the functions for the encoder, the processor G, and the processor H for

17
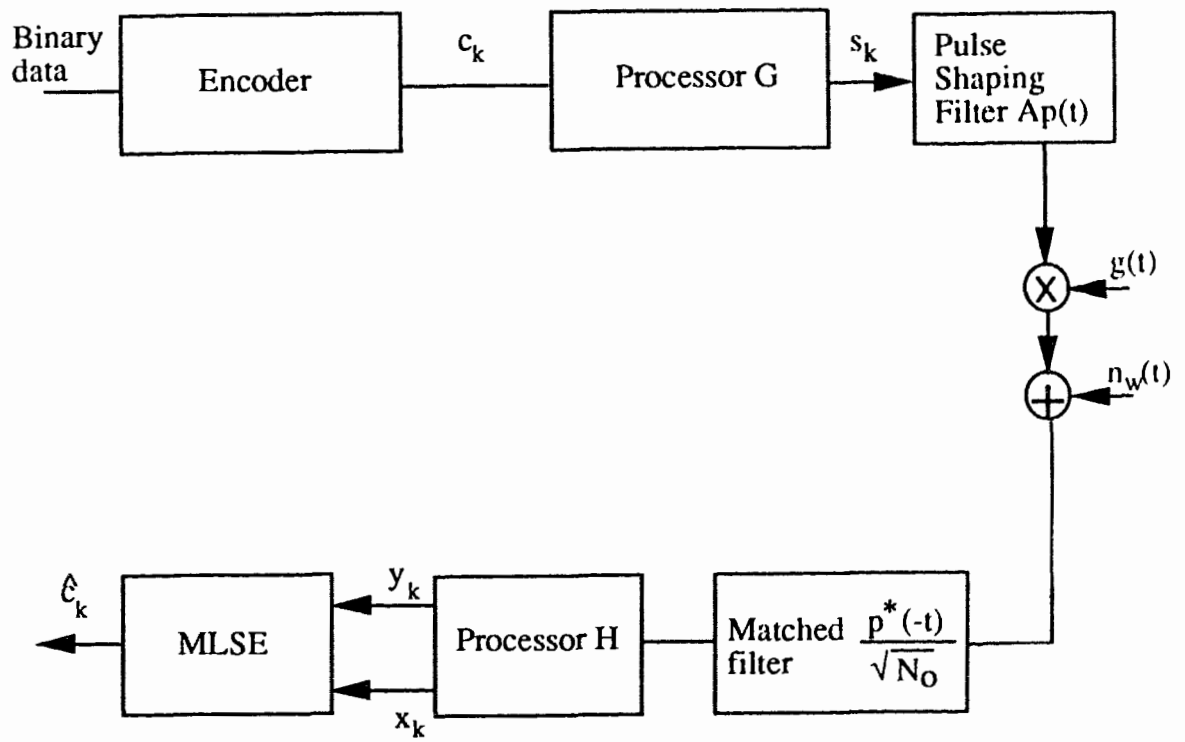
Figure 2.1: General system model.
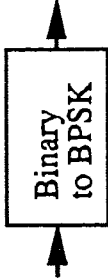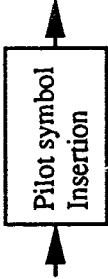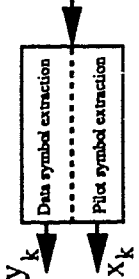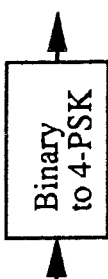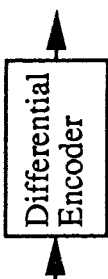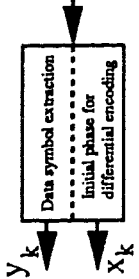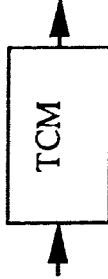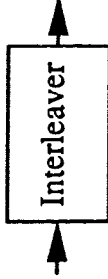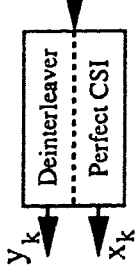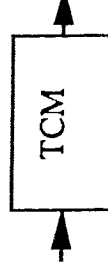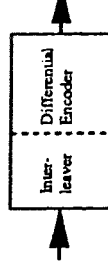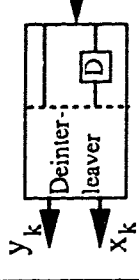
	PSAM	Multiple symbol differential detection	Interleaved coherent PSK	Interleaved DPSK
Encoder	Binary to BPSK	Binary to 4-PSK	TCM	TCM
Processor G	Pilot symbol Insertion	Differential Encoder	Interleaver	Inter-leaver / Differential Encoder
Processor H	Data symbol extraction / Pilot symbol extraction (y_k, x_k)	Data symbol extraction / initial phase for differential encoding (y_k, x_k)	Deinterleaver / Perfect CSI (y_k, x_k)	Deinter-leaver / D (y_k, x_k)

Table 2.1: The actual function of the Encoder, the processor G, and the processor H in Figure 2.1 for different detection schemes.

different detection schemes are listed in Table 2.1. For example, the encoder for PSAM is a binary to BPSK converter while for multiple symbol differential detection, the encoder represents a binary to 4-PSK converter. The output sequence c of the encoder is further processed by the processor G and produces a transmit sequence $s = (s_1, \ldots, s_k, \ldots)$. Like the c_k's, the s_k's are also complex PSK symbols. In the case of PSAM, the processor G inserts pilot symbols periodically into the sequence c while in multiple symbol differential detection, the processor G encodes the sequence c differentially. On the other hand, for coded coherent PSK, the processor G just acts as an interleaver while in coded DPSK, the processor G not only performs interleaving, but also encodes the scrambled sequence differentially; see Table 2.1.

The baseband equivalent of the transmit signal is

$$s(t) = A \sum_k s_k p(t - kT) \tag{2.1}$$

where $p(t)$ is the complex impulse response of a pulse shaping filter that satisfies the Nyquist's criterion of zero intersymbol interference [22], $1/T$ is the symbol rate, and A is a constant. The transmitted symbol s_k in the interval $kT \leq t \leq (k+1)T$ has the complex form

$$s_k = exp\{j\theta_k\} \tag{2.2}$$

where θ_k takes on one of the I uniformly distributed values $\psi_i = \frac{2\pi i}{I}; i = 0, 1, \ldots, I-1$ around the unit circle. The energy of the pulse $p(t)$ is normalized such that

$$\int_{-\infty}^{\infty} |p(t)|^2 dt = 1.$$

20

In a mobile radio environment, there are usually a large number of diffuse scatterers or reflectors which move randomly relative to one another. The received signal is therefore, a linear combination of many replicas of the original signal, each attenuated and phase shifted by a random amount [34]. By the central limit theorem [24], the distribution of the sum of many independent random variables approaches the Gaussian distribution. Hence, the baseband equivalent of the received signal can be written as

$$
\begin{aligned}
r(t) &= g(t)s(t) + n_w(t) \\
&= Ag(t)\left(\sum_k s_k p(t - kT)\right) + n_w(t)
\end{aligned}
\tag{2.3}
$$

where $g(t)$ is a zero mean, complex, Gaussian random process representing the Rayleigh fading in the channel, and $n_w(t)$ is the complex envelope of the channel's white Gaussian noise. The double-sided power spectral density (PSD) of $n_w(t)$ is N_o. For a mobile radio channel, the autocorrelation function of the fading process can be modeled as [16]

$$
\begin{aligned}
R_g(\tau) &= \frac{1}{2}E\{g^*(t)g(t + \tau)\} \\
&= \sigma_g^2 J_o(2\pi f_D \tau)
\end{aligned}
\tag{2.4}
$$

where J_o is the Bessel function of order zero, f_D is the maximum Doppler frequency (also referred as the fade rate in this study):

$$
f_D = \frac{speed\, of\, vehicle}{transmission\, wavelength},
$$

21

and $\sigma_g^2 = R_g(0)$ is the variance of the fading process $g(t)$. We assume the fading process $g(t)$ is slow enough that $g(t)$ remains roughly constant over each symbol interval. This implies the received signal in (2.3) can be simplified to

$$r(t) = A\left(\sum_k g_k s_k p(t - kT)\right) + n_w(t) \tag{2.5}$$

where g_k denotes the value of $g(t)$ during the kth interval. The average power of the signal component of $r(t)$ is thus

$$
\begin{aligned}
P_s(t) &= \frac{1}{2}E\{|g(t)s(t)|^2\} \\
&= \frac{A^2}{2}\sum_l\sum_k E\{s_l^* s_k\}E\{g_l^* g_k\}p^*(t - lT)p(t - kT) \\
&= \frac{A^2}{2}\sum_k E\{|s_k|^2\}E\{|g_k|^2\}|p(t - kT)|^2 \\
&= A^2\sigma_g^2\sum_k |p(t - kT)|^2.
\end{aligned}
\tag{2.6}
$$

The average received signal energy for each modulation symbol sent is thus

$$
\begin{aligned}
E_s &= \int_0^T P_s(t)dt \\
&= A^2\sigma_g^2\int_{-\infty}^{\infty}|p(t)|^2 dt \\
&= A^2\sigma_g^2.
\end{aligned}
\tag{2.7}
$$

If each modulation symbol carries m bits of information, then the average received energy per bit is $E_b = E_s/m$.

At the receiver side, the received signal is passed to a matched filter with an

22

impulse response equal to $p^*(-t)/\sqrt{N_o}$. The output of the filtered signal is sampled at the symbol rate and the samples are used to derived two sequences:

$$\mathbf{y} = (y_1, y_2, \ldots, y_N)^t, \tag{2.8}$$

and

$$\mathbf{x} = (x_1, x_2, \ldots, x_M)^t. \tag{2.9}$$

The sequence y contains the corrupted signals of the original data while the sequence x contains information about the fading experienced by the data symbols. When fading is slow enough that it will not distort the shape of the transmittion pulse and hence produce no appreciable intersymbol interference, then, the sample y_k can be written as [5]

$$y_k = u_k c_k + n_k. \tag{2.10}$$

The u_k's in (2.10) are complex Gaussian random variables which are called fading gains while the n_k's are independent and identically distributed (iid) complex Gaussian random variables representing samples of the filtered channel noise. Each n_k's has a zero mean and a unit variance. It should be pointed out that the u_k's and n_k's are statistically independent.

The fading gain sequence u_k's in (2.10) are correlated, zero mean Gaussian random variables. Depending on the system being studied, the u_k's can be data dependent. In any case, we let

$$\varphi(i,j) = \frac{1}{2}E\{u_i u_j^*\} \tag{2.11}$$

denote the correlation between u_i and u_j. The expression for $\phi(i,j)$ for different systems will be derived in subsequent chapters.

In addition to the sequence $\mathbf{y}$, the processor II also provides a sequence $\mathbf{x} = (x_1, \ldots, x_M)^t$ that contains the channel state information (CSI); see Table 2.1. The sequence length is M and the x_k's are dependent of the channel estimator used. For example, in coded PSK with perfect CSI, $M = N$, and $x_k = u_k$ while in multiple-symbol differential detection, $M = 1$, and x_1 is the noisy phase reference at the beginning of the data block. In any case, the x_k's are correlated, zero mean complex Gaussian random variables.

The received sequences $\mathbf{x}$ and $\mathbf{y}$ can be written in matrix form as

$$\mathbf{r} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{u} \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{n} \end{pmatrix}, \tag{2.12}$$

where

$$\mathbf{C} = \begin{pmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_N \end{pmatrix}, \tag{2.13}$$

$$\mathbf{u} = (u_1, u_2, \ldots, u_N)^t, \tag{2.14}$$

$$\mathbf{n} = (n_1, n_2, \ldots, n_N)^t, \tag{2.15}$$

24

and I is an $M \times M$ identity matrix. Since the variance of the n_k's is 1, the covariance matrix for n [24] is

$$\Phi_{nn} = \frac{1}{2}E\{nn^*\} = I, \qquad (2.16)$$

where I is an $N \times N$ identity matrix. Also for simplicity, we shall use

$$\Phi_{xx} = \frac{1}{2}E\{xx^\dagger\} \qquad (2.17)$$

and

$$\Phi_{uu} = \frac{1}{2}E\{uu^\dagger\} \qquad (2.18)$$

to represent the covariance matrices for the sequences x and u. In addition,

$$\Phi_{xu} = \frac{1}{2}E\{xu^\dagger\} = \Phi_{ux}^\dagger \qquad (2.19)$$

will be used to represent the cross-covariance matrix of the sequences x and u. The elements in the above covariance matrices will be defined in subsequent chapters. Given any data sequence $\{c_1, c_2, \ldots, c_N\}$, the covariance matrix of r is

$$\Phi_{rr} = \frac{1}{2}E\{rr^\dagger\} = \begin{pmatrix} \Phi_{xx} & \Phi_{xy} \\ \Phi_{yx} & \Phi_{yy} \end{pmatrix}, \qquad (2.20)$$

where

$$\Phi_{yy} = C\Phi_{uu}C^\dagger + I,$$

25

$$= \mathbf{C}\boldsymbol{\Phi}_{\mathbf{uu}}\mathbf{C}^\dagger + \mathbf{CIC}^\dagger$$

$$= \mathbf{C}\boldsymbol{\Phi}_{\mathbf{uu}}'\mathbf{C}^\dagger. \tag{2.21}$$

with

$$\boldsymbol{\Phi}_{\mathbf{uu}}' = \boldsymbol{\Phi}_{\mathbf{uu}} + \mathbf{I}, \tag{2.22}$$

and

$$\boldsymbol{\Phi}_{\mathbf{xy}} = \boldsymbol{\Phi}_{\mathbf{xu}}\mathbf{C}^\dagger = \boldsymbol{\Phi}_{\mathbf{yx}}^\dagger. \tag{2.23}$$

2.3 The Optimum Decoder

Let the set of all possible N-length data sequences be denoted by

$$\{\hat{\mathbf{c}} = (\hat{c}_1, \ldots, \hat{c}_N)\}$$

where $\hat{c}_k$ is the kth symbol in the sequence $\hat{\mathbf{c}}$. An optimum decoder will select the sequence for which the a-posteriori probability $P(\hat{\mathbf{c}}|\mathbf{y},\mathbf{x})$ is the largest. If all the sequences are sent with equal probability, this is equivalent to choosing the sequence with the largest conditional probability density function $p(\mathbf{y}|\hat{\mathbf{c}},\mathbf{x})$. The joint probability density function of the x_k's and the y_k's, given $\hat{\mathbf{c}}$, is

$$p(\mathbf{x},\mathbf{y}|\hat{\mathbf{c}}) = \frac{1}{(2\pi)^{N+M}||\boldsymbol{\Phi}_{\mathbf{rr}}||}exp\left\{-\frac{1}{2}\mathbf{r}^\dagger\boldsymbol{\Phi}_{\mathbf{rr}}^{-1}\mathbf{r}\right\}, \tag{2.24}$$

where $\mathbf{r}$ is defined in (2.12) and the joint probability density function of the x_k's is

26

$$p(\mathbf{x}) = \frac{1}{(2\pi)^M ||\Phi_{\mathbf{xx}}||} exp\left\{-\frac{1}{2}\mathbf{x}^\dagger \Phi_{\mathbf{xx}}^{-1}\mathbf{x}\right\}. \tag{2.25}$$

Note that in this thesis, the notation $|| \bullet ||$ is used to represent the determinant of a matrix. The conditional probability density function $p(\mathbf{y}|\hat{\mathbf{c}}, \mathbf{x})$ is, therefore,

$$
\begin{aligned}
p(\mathbf{y}|\hat{\mathbf{c}}, \mathbf{x}) &= \frac{p(\mathbf{x}, \mathbf{y}|\hat{\mathbf{c}})}{p(\mathbf{x})} \\
&= \frac{1}{(2\pi)^N ||\Phi_{\mathbf{rr}}||/||\Phi_{\mathbf{xx}}||} exp\left\{-\frac{1}{2}\left(\mathbf{r}^\dagger \Phi_{\mathbf{rr}}^{-1}\mathbf{r} - \mathbf{x}^\dagger \Phi_{\mathbf{xx}}^{-1}\mathbf{x}\right)\right\}. \tag{2.26}
\end{aligned}
$$

Although the decoding expression in (2.26) looks rather complicate, it is possible to reduce it into a simpler form; see Appendix A. The resulting optimum decoding metric is of the form:

$$M(\hat{\mathbf{C}}) = \left(\begin{array}{cc} \mathbf{x}^\dagger & \mathbf{y}^\dagger \end{array}\right) \left(\begin{array}{cc} \mathbf{0} & -\mathbf{b}^\dagger \mathbf{a}\hat{\mathbf{C}}^\dagger \\ -\hat{\mathbf{C}}\mathbf{a}\mathbf{b} & \hat{\mathbf{C}}\mathbf{a}\hat{\mathbf{C}}^\dagger \end{array}\right) \left(\begin{array}{c} \mathbf{x} \\ \mathbf{y} \end{array}\right) \tag{2.27}$$

where

$$\mathbf{a} = \left[\Phi_{\mathbf{uu}}' - \Phi_{\mathbf{ux}}\Phi_{\mathbf{xx}}^{-1}\Phi_{\mathbf{xu}}\right]^{-1}, \tag{2.28}$$

$$\mathbf{b} = \Phi_{\mathbf{ux}}\Phi_{\mathbf{xx}}^{-1}, \tag{2.29}$$

and $\Phi_{\mathbf{uu}}'$ is defined in (2.22).

2.4 The Pairwise Error Event Probability

After obtaining the optimum decoding metric, we are now in a position to derive the pairwise error event probability. Let the transmitted data sequence be c. Then, according to (2.27), the decoder will make a wrong decision if for some sequence ĉ, the random variable

$$D = M(\hat{\mathbf{C}}) - M(\mathbf{C})$$
$$= \mathbf{r}^\dagger \mathbf{F} \mathbf{r} \tag{2.30}$$

is less than zero. The probability that D is less than zero is known as the pairwise error event probability. Note that r is the column vector defined in (2.12) and F is the following matrix:

$$\mathbf{F} = \begin{pmatrix} \mathbf{0} & \mathbf{b}^\dagger \mathbf{a} \boldsymbol{\Delta}^\dagger \\ \boldsymbol{\Delta} \mathbf{a} \mathbf{b} & \hat{\mathbf{C}} \mathbf{a} \hat{\mathbf{C}}^\dagger - \mathbf{C} \mathbf{a} \mathbf{C}^\dagger \end{pmatrix}, \tag{2.31}$$

with

$$\boldsymbol{\Delta} = \mathbf{C} - \hat{\mathbf{C}}. \tag{2.32}$$

The random variable D in (2.30) is a quadratic form of dependent Gaussian variates. Following the procedures outlined in Appendix B, the random variable D can be transformed into a sum of independent quadratic forms of complex Gaussian variates; i.e.:

28

$$D = \sum_k \lambda_{b,k} |q_k|^2.$$

As showed in Appendix B, the q_k's are iid complex Gaussian variables each having a zero mean and a unit variance and $\lambda_{b,k}$'s are the eigenvalues of the matrix $\mathbf{B}$ where

$$\mathbf{B} = \Lambda_r^{\frac{1}{2}} \mathbf{U}_r^\dagger \mathbf{F} \mathbf{U}_r \Lambda_r^{\frac{1}{2}}. \tag{2.33}$$

As pointed out in Appendix B, $\Lambda_r^{\frac{1}{2}}$ is a diagonal matrix whose diagonal elements are the square root of the eigenvalues of the matrix $\boldsymbol{\Phi}_{rr}$, $\mathbf{U}_r$ is a unitary matrix whose columns are orthonormal and $(\mathbf{U}_r \Lambda_r^{\frac{1}{2}})(\Lambda_r^{\frac{1}{2}} \mathbf{U}_r^\dagger) = \boldsymbol{\Phi}_{rr}$. Using the results from [5], it is easy to calculate $P(\mathbf{c} \to \hat{\mathbf{c}})$, the pairwise error event probability. Let $p(D)$ be the probability density function of the random variable D. Then, the characteristic function, or the two-sided Laplace Transform [27] of $p(D)$ is

$$
\begin{aligned}
\Phi_D(s) &= \int_{-\infty}^{\infty} p(D) e^{-sD} dD \\
&= \prod_k \frac{-p_k}{s - p_k}
\end{aligned}
\tag{2.34}
$$

where

$$p_k = \frac{-1}{2\lambda_{b,k}} \tag{2.35}$$

and the region of convergence is the vertical strip enclosing the $j\omega$ axis bounded by the closest poles on either sides. In addition, it should be pointed out that if $\lambda_{b,k}$ is zero, the term $\frac{-p_k}{s-p_k}$ is equal to unity. From [5], the pairwise error probability

29

is equal to the sum of the residues of the function $\Phi_D(s)/s$ calculated at the poles located in the right-half of the complex s-plane, i.e.,

$$
\begin{aligned}
P(\mathbf{c} \to \hat{\mathbf{c}}) &= Prob(D \leq 0) \\
&= -\sum Residue \left[\frac{\Phi_D(s)}{s} \right]_{RPpoles}.
\end{aligned}
\tag{2.36}
$$

This expression for the pairwise error event probability can be easily computed, even in the case of higher order poles, since there exists a recursive procedure for calculating the different derivatives of the characteristic function.

The key elements in the expression (2.36) are the eigenvalues, the $\lambda_{b,k}$'s of the matrix $\mathbf{B}$ in (2.33). However, in order to find the $\lambda_{b,k}$'s, a large number of matrix calculations and decompositions have to be performed; see Appendix B. In fact, there is a much simpler way to find the $\lambda_{b,k}$'s. As found in Appendix C, the $\lambda_{b,k}$'s are also the eigenvalues of the matrix

$$
\mathbf{G'} = \mathbf{\Phi_{rr}F}.
\tag{2.37}
$$

Consequently, a lot of matrix eigenvalue calculations can be avoided in computing the poles p_k's in (2.34).

In summary, the pairwise error event probability for a PSK sequence transmitted over a correlated Rayleigh fading channel can be evaluated as follows:

1. Express the random variable D in the form $\mathbf{r^\dagger F r}$ where $\mathbf{r}$ is a column vector which contains the correlated random variables x_k's and y_k's. The probability that $D \leq 0$ is the pairwise error event probability. It should be pointed

30

out that pairwise error event probability is obtained by assuming only two codewords are present and different pairwise error events are dependent.

2. Find the covariance matrix of **r**.

3. Find the eigenvalues of the matrix $\mathbf{\Phi_{rr}F}$ in (2.37), or of the matrix **B** in (2.33).

4. Substitute those eigenvalues found in step 3 into (2.34) – (2.36) to calculate the error event probability.

In nearly all applications, we are interested in the overall bit error probability of the communication systems rather than the individual pairwise error event probability. A good approximation of the bit error probability can be obtained by summing a small set of pairwise error probabilities. This will be shown in subsequent chapters. There are two points we want the readers to take notice of. First, the general form of the decoder in (2.27) implies the receiver has prior knowledge about the autocorrelation function of the fading process. Second, although the pairwise error probability analysis has been derived for an optimum decoder, our analysis still applies to non-optimal decoders as long as the decoding metric results in a Hermitian matrix **F**, like the one in (2.31).

2.5   Summary

In this chapter, we have presented detailed procedures for the performance analysis of PSK signals transmitted over correlated Rayleigh fading channels. The optimal decoding metric is derived along with the expression for the pairwise error event

probability. The applications of the analysis developed in this chapter will be given as examples in subsequent chapters.

CHAPTER 3

PERFORMANCE OF BLOCK DECODING OF PSK SIGNALS

In this chapter, we will present two applications to illustrate the general theory that has been developed in Chapter 2. The analysis of the uncoded PSK signals with multiple-symbol differential detection will be discussed in the first part. The analysis of uncoded BPSK signals with pilot symbol assisted demodulation will be shown in the second part. The optimum decoders and the pairwise error event probabilities will be obtained for both cases. Results and discussions will be presented subsequently.

3.1 Multiple-symbol Differential Detection of PSK Signals

In a conventional differential detector, the previous received sample is used as an estimate of the channel complex gain in the current interval. This works fine as long as fading is relatively slow. As shown in [5], the irreducible error floor associated with conventional differential detection is proportional to $(\pi f_D T)^2$ where $f_D T$ is the normalized Doppler frequency. Intuitively, more accurate channel state information can be obtained by using more than one previously received samples. This leads us to the concept of multiple symbol differential detection. In simplest terms, a multiple-symbol differential detector is a decoder that makes a decision about a block of N consecutive PSK symbols based on $N + 1$ received samples. When N is equal to 1, we have a conventional differential detector. By making use of the results developed in Chapter 2, we can derive the optimum decoding metric and the pairwise error event probability for multiple-symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels. The numerical results for uncoded 4-DPSK will be presented subsequently.

3.1.1 System Description

Using the notation in section 2.2, the transmitted PSK symbol in the kth signaling interval (i.e. $kT \leq t \leq (k+1)T$, T being the symbol duration) can be expressed in complex form as:

$$s_k = exp(j\theta_k), \quad k = 1, \ldots, N \tag{3.1}$$

where θ_k is the transmitted signal phase and will take on one of the I values from the set $\{\frac{2\pi i}{I}; i = 0, 1, \ldots, I-1\}$. Due to the differential encoding process, the sequence $\mathbf{s} = (s_1, s_2, \ldots, s_N)$ is related to the message sequence $\mathbf{c} = (c_1, c_2, \ldots, c_N)$ by the following:

$$s_k = s_{k-1}c_k = s_0 z_k \tag{3.2}$$

where

$$z_k = \prod_{j=1}^{k} c_j, \tag{3.3}$$

and s_0 is the reference symbol for the differential encoding process. Like the s_k's, the c_k's and consequently the z_k's are complex phasors taken from the set $\{exp(j\frac{2\pi i}{I}); i = 0, 1, \ldots, I-1\}$.

The complex PSK symbols s_k's are transmitted over a non-frequency selective Rayleigh fading channel with additive white Gaussian noise. After filtering and sampling, the received signal in the kth interval is $y_k = u_k c_k + n_k$, $k = 1, \ldots, N$; (see (2.10)), where

$$u_k = \frac{Ag_k}{\sqrt{N_o}}s_{k-1}. \tag{3.4}$$

Note that g_k is the channel complex gain in the kth interval and u_k's are complex Gaussian random variables with variance:

$$\sigma_u^2 = \frac{A^2}{N_o}\sigma_g^2 = E_s/N_o, \tag{3.5}$$

where σ_g^2 is shown in (2.7). As before, n_k represents a filtered Gaussian noise sample and the n_k's are a set of iid variables with a variance of $\sigma_n^2 = \frac{1}{2}E\{|n_k|^2\} = 1$. The vector $\mathbf{u} = (u_1, \ldots, u_N)^t$ can be written as

$$\mathbf{u} = \frac{A}{\sqrt{N_o}}\mathbf{Sg} \tag{3.6}$$

where

$$\mathbf{S} = \begin{pmatrix} s_0 & & \\ & \ddots & \\ & & s_{N-1} \end{pmatrix}, \tag{3.7}$$

and

$$\mathbf{g} = (g_1, \ldots, g_N)^t. \tag{3.8}$$

The covariance matrix $\boldsymbol{\Phi}_{\mathbf{uu}}$ in (2.18) becomes

$$\boldsymbol{\Phi}_{\mathbf{uu}} = \mathbf{S}\boldsymbol{\Phi}_{22}\mathbf{S}^\dagger \tag{3.9}$$

where

$$\Phi_{22} = \frac{A^2}{N_o} \frac{1}{2} E\{\mathbf{g}\mathbf{g}^\dagger\} = \begin{pmatrix} \phi(1,1) & \phi(1,2) & \cdots & \phi(1,N) \\ \phi(2,1) & \phi(2,2) & \cdots & \phi(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(N,1) & \phi(N,2) & \cdots & \phi(N,N) \end{pmatrix}, \qquad (3.10)$$

and

$$\begin{aligned} \phi(i,j) &= \frac{A^2}{N_o} \frac{1}{2} E\{g_i g_j^*\} \\ &= \frac{E_s}{N_o} J_0(2\pi f_D T(i-j)), \end{aligned} \qquad (3.11)$$

see (2.4) and (2.7). Now substituting the above expression for $\Phi_{\mathbf{uu}}$ into (2.21) implies

$$\begin{aligned} \Phi_{\mathbf{yy}} &= \mathbf{C}(\mathbf{S}\Phi_{22}\mathbf{S}^\dagger + \mathbf{I})\mathbf{C}^\dagger \\ &= \mathbf{C}\mathbf{S}(\Phi_{22} + \mathbf{I})\mathbf{S}^\dagger\mathbf{C}^\dagger \\ &= \mathbf{z}(\Phi_{22} + \mathbf{I})\mathbf{z}^\dagger \end{aligned} \qquad (3.12)$$

where

$$\mathbf{z} = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_N \end{pmatrix}, \qquad (3.13)$$

and the z_k's are defined in (3.3).

Since differential encoding/detection is used, we are using the received sample

corresponding to the initial transmitted phase s_0 to provide the CSI. As a result, the length of the sequence x in (2.9) is equal to 1 ($M = 1$) and

$$x_1 = u_0 s_0 + n_0. \tag{3.14}$$

This implies that the matrix Φ_{xu} in (2.19) becomes:

$$\Phi_{xu} = \frac{1}{2} E\{x_1 u^\dagger\} = s_0 \Phi_{12} S^\dagger \tag{3.15}$$

where

$$\Phi_{12} = (\phi(0,1), \ldots, \phi(0,N)) \tag{3.16}$$

and

$$\phi(0,j) = \frac{A^2}{N_o} \frac{1}{2} E\{g_0 g_j^*\} = \frac{E_s}{N_o} J_o(2\pi f_D T j). \tag{3.17}$$

Substituting Φ_{xu} into (2.23) implies

$$\Phi_{xy} = s_0 \Phi_{12} S^\dagger C^\dagger = \Phi_{12} z^\dagger \tag{3.18}$$

where z was defined in (3.13). As mentioned in Chapter 2, $\Phi_{xy} = \Phi_{yx}^\dagger$. Finally, the term Φ_{xx} in (2.17) becomes

$$\Phi_{xx} = \frac{1}{2} E\{|x_1|^2\} = \phi(0,0) + 1 = \frac{E_s}{N_o} + 1. \tag{3.19}$$

3.1.2 The Optimum Decoder

The received sequence $(y_1, y_2, \ldots, y_N)$ as well as x_1 are passed to a multiple-symbol differential detector implemented as a Maximum Likelihood Sequence Estimator (MLSE). The function of the MLSE is to select from all the I^N possible N-length sequences the most likely message sequence. This is accomplished by finding the sequence $\hat{\mathbf{c}} = (\hat{c}_1, \ldots, \hat{c}_N)$ for which the conditional probability density function $p(\mathbf{y}|x_1, \hat{\mathbf{c}})$ is the largest, $\mathbf{y}$ being the vector defined earlier in (2.8). Using the results from section 3.1.1 as well as those from Section 2.2, we have the following decoding metric

$$M(\hat{\mathbf{c}}) = (x_1^* \; \mathbf{y}^\dagger) \begin{pmatrix} 0 & -\tilde{\mathbf{b}}^\dagger \tilde{\mathbf{a}} \hat{\mathbf{z}}^\dagger \\ -\hat{\mathbf{z}} \tilde{\mathbf{a}} \tilde{\mathbf{b}} & \hat{\mathbf{z}} \tilde{\mathbf{a}} \hat{\mathbf{z}}^\dagger \end{pmatrix} \begin{pmatrix} x_1 \\ \mathbf{y} \end{pmatrix} \qquad (3.20)$$

where $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$ are:

$$\tilde{\mathbf{a}} = \left[\Phi'_{22} - \Phi_{21} \Phi_{xx}^{-1} \Phi_{12} \right]^{-1}, \qquad (3.21)$$

$$\tilde{\mathbf{b}} = \Phi_{21} \Phi_{xx}^{-1}, \qquad (3.22)$$

$$\Phi'_{22} = \Phi_{22} + \mathbf{I}, \qquad (3.23)$$

$$\hat{\mathbf{z}} = \begin{pmatrix} \hat{z}_1 & & \\ & \ddots & \\ & & \hat{z}_N \end{pmatrix}, \qquad (3.24)$$

and $\hat{z}_k = \prod_{j=1}^{k} \hat{c}_j$. The decoder will select the sequence $\hat{\mathbf{c}} = (\hat{c}_1, \ldots, \hat{c}_N)$ whose metric is the smallest. In using the metric in (3.20), the assumption of a prior knowledge about the autocorrelation function of the fading process must be made.

This can be done in practice by estimating the autocorrelation function from a known training sequence sent during the start-up phase of a communication session, and by continuously updating these estimates during the data transmission phase. Since estimation errors are sometimes unavoidable (especially during the acceleration or deceleration of the vehicle), we will examine in the following sections the error performance of a detector whose decoding metric is not perfectly matched to the channel statistics.

In the case of static fading, i.e. when $f_D T = 0$, the matrices $\bar{a}$ and $\bar{b}$ in (3.21) and (3.22) becomes $\mathbf{I} - f\mathbf{J}$ and $\mathbf{1}f'$ respectively, where $f = \phi(0,0)/(1 + (1 + N)\phi(0,0))$, $f' = \frac{\phi(0,0)}{\phi(0,0)+1}$, and $\mathbf{J}$ and $\mathbf{1}$ are $N \times N$ and $N \times 1$ matrices whose entries are all equal to unity; see Appendix D. Consequently, the optimum decoding metric from (3.20) becomes

$$M(\hat{\mathbf{c}}) = \sum_{k=1}^{N} |y_k|^2 - f\left[\left|\sum_{k=1}^{N} y_k \hat{z}_k^*\right|^2 + \sum_{k=1}^{N}(y_k x_1^* \hat{z}_k^* + y_k^* x_1 \hat{z}_k)\right],$$

or equivalently

$$M(\hat{\mathbf{c}}) = \left|x_1 + \sum_{k=1}^{N} y_k \hat{z}_k^*\right|^2. \tag{3.25}$$

In otherwords, for very slow fading, the function of optimum decoder is to select the sequence $(z_1, \ldots, z_N)$ which has the highest correlation with the received sequence. It should be pointed out that (3.25) is also the optimum decoding metric for the AWGN channel; see [11] and [36].

3.1.3 The Error Performance

Given the actual message sequence is $\mathbf{c} = (c_1, \ldots, c_N)$, the decoder will make a wrong decision if for some sequence $\hat{\mathbf{c}} = (\hat{c}_1, \ldots, \hat{c}_N)$, the random variable D in (2.30) is less than zero. As shown in Section 2.4, we can follow the four steps developed on page 31 to obtain the pairwise error event probability. Note that the matrix $\mathbf{F}$ in (2.31) becomes:

$$\mathbf{F} = \begin{pmatrix} 0 & \tilde{\mathbf{b}}^\dagger \tilde{\mathbf{a}} \tilde{\boldsymbol{\Delta}}^\dagger \\ \tilde{\boldsymbol{\Delta}} \tilde{\mathbf{a}} \tilde{\mathbf{b}} & \hat{\mathbf{z}} \tilde{\mathbf{a}} \hat{\mathbf{z}}^\dagger - \mathbf{z} \tilde{\mathbf{a}} \mathbf{z}^\dagger \end{pmatrix} \tag{3.26}$$

where $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$ were defined in (3.21) and (3.22) and

$$\tilde{\boldsymbol{\Delta}} = \mathbf{z} - \hat{\mathbf{z}}.$$

We will make use of the exact pairwise error event probability expression in (2.36) to evaluate the overall bit error probability. Results for uncoded 4-DPSK, with the Doppler frequency of the Rayleigh fading channel and the word length of the multiple-symbol differential detector as parameters, will be presented and compared with that of a perfect coherent detector.

In analyzing the bit error performance of an uncoded MPSK system, we note that the message sequence $\mathbf{c} = (c_1, \ldots, c_N)$ is obtained by mapping binary n-tuples, $n = log_2(I)$, to MPSK symbols using Gray coding. An example using 4-PSK is shown in Figure 3.1. We also note that the bit error probability is independent of the message sequence sent; see Appendix E. As a result, we assume $\mathbf{c}$ is the all zero-phase sequence, i.e., $\mathbf{c} = (1, \ldots, 1)$. Consequently, all the elements in the sequence

41
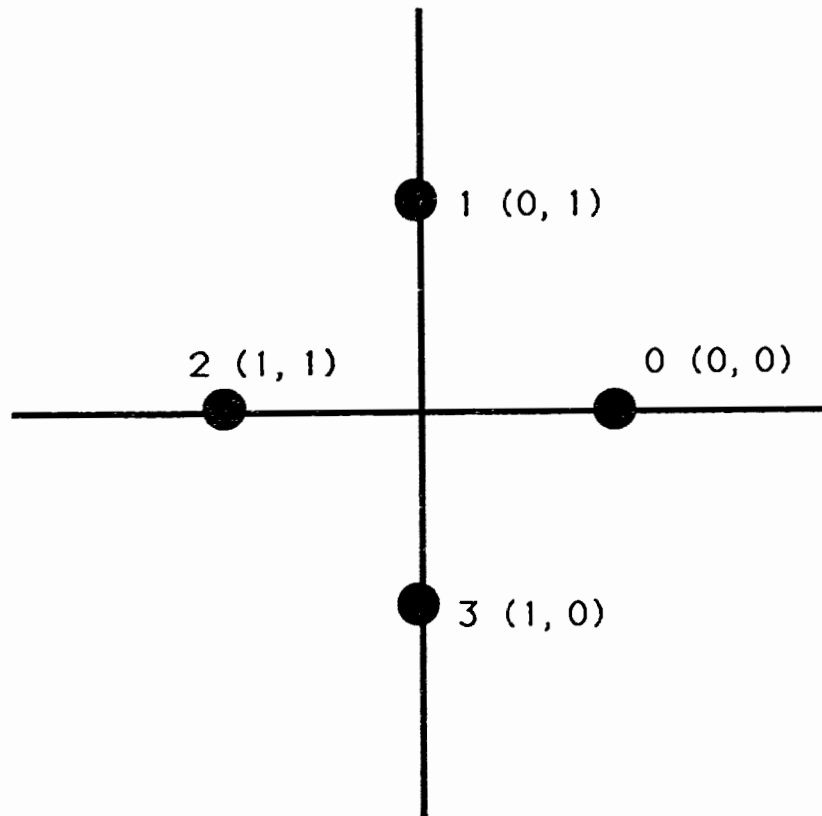
Figure 3.1: The 4-PSK constellation and Gray mapping. The magnitude of each signal vector is equal to unity.

$(z_1, \ldots, z_N)$ are equal to unity; see (3.3).

An upperbound on the bit error probability can be obtained by taking the union bound of the pairwise error event probabilities. For a multiple-symbol detector with a word length equal to N, there are $I^N - 1$ error events to consider. Clearly, the computational complexity required to obtain the union bound increases dramatically as N increases. It is shown in [11] that for an AWGN channel, an accurate estimate of the overall bit error probability can be obtained by considering only the set of most likely (or dominant) error events. These most likely error events are determined by those sequences $(\hat{z}_1, \ldots, \hat{z}_N)$ which have the highest correlation with the sequence $(z_1, \ldots, z_N)$, where the correlation is defined as

$$\mu = \left| 1 + \sum_{k=1}^{N} z_k^* \hat{z}_k \right|^2. \tag{3.27}$$

As mentioned above, we assumed all the z_k's are equal to unity. Subsequently, (3.27) becomes

$$\mu = \left| 1 + \sum_{k=1}^{N} \hat{z}_k \right|^2.$$

This expression indicates that there are $2(N+1)$ sequences that give the same largest μ and they are denoted by:

$$\hat{\mathbf{z}}^{(i)} = (\hat{z}_1^{(i)}, \ldots, \hat{z}_N^{(i)}); \qquad i = 1, \ldots, 2(N+1).$$

Note that $\hat{\mathbf{z}}^{(1)} = (e^{j2\pi/I}, \ldots, e^{j2\pi/I})$, $\hat{\mathbf{z}}^{(2)} = (e^{-j2\pi/I}, \ldots, e^{-j2\pi/I})$, and each of the remaining $\hat{\mathbf{z}}^{(i)}$ contains one and only one symbol of the form $e^{j\pm2\pi/I}$ and the remaining $N - 1$ symbols are all equal to unity. Since a static fading channel $(f_D T = 0)$ is

equivalent to an AWGN channel throughout the duration of each block of N symbols, consequently, the most likely error events for these two channels are identical. As a matter of fact, we have shown earlier that the optimum decoding metric is the same for both the AWGN and the static fading channel; see section 3.1.2. In general, we believe that as long as $f_D T$ is relatively small, the most likely error events in a fading channel are identical to those in the AWGN channel. Consequently, only those error events mentioned above will be included in the bit error probability analysis.

Let the erroneous message sequence corresponding to $\hat{\mathbf{z}}^{(i)} = (\hat{z_1}^{(i)}, \ldots, \hat{z_N}^{(i)})$ be denoted by $\hat{\mathbf{c}}^{(i)} = (\hat{c_1}^{(i)}, \ldots, \hat{c_N}^{(i)})$ where $\hat{z_k}^{(i)} = \prod_{j=1}^{k} \hat{c_j}^{(i)}$. Because of the use of differential encoding, $2N - 2$ of these erroneous message sequences have the symbol $e^{j2\pi/I}$ (or $e^{-j2\pi/I}$) followed immediately by $e^{-j2\pi/I}$ (or $e^{j2\pi/I}$) and the rest of the symbols in these sequences are all equal to unity. The remaining 4 sequences are denoted by: $\hat{\mathbf{c}}^{(1)} = (e^{j2\pi/I}, 1, \ldots, 1)$, $\hat{\mathbf{c}}^{(2)} = (e^{-j2\pi/I}, 1, \ldots, 1)$, $\hat{\mathbf{c}}^{(3)} = (1, \ldots, 1, e^{j2\pi/I})$, and $\hat{\mathbf{c}}^{(4)} = (1, \ldots, 1, e^{-j2\pi/I})$. Since Gray coding is used to map binary n-tuples to MPSK symbols, this implies the Hamming distance between the binary equivalent of $\mathbf{c}$ and $\hat{\mathbf{c}}^{(i)}$ is

$$d(\mathbf{c}, \hat{\mathbf{c}}^{(i)}) = \begin{cases} 1 & 1 \leq i \leq 4 \\ 2 & 5 \leq i \leq 2(N+1) \end{cases} \tag{3.28}$$

Consequently, an approximation to the bit error probability of a multiple-symbol differential detector in a Rayleigh fading channel is:

$$\mathbf{P_b} \approx \frac{1}{N log_2(I)} \sum_{i=1}^{2N+2} d(\mathbf{c}, \hat{\mathbf{c}}^{(i)}) \mathbf{P}(\mathbf{c} \to \hat{\mathbf{c}}^{(i)}). \qquad 2 \leq N \tag{3.29}$$

Since the enumeration of the erroneous messages is different for a conventional differential detector, the above equation is only used when $N > 1$. However, it can be shown that for a conventional differential detector ($N = 1$), the bit error probability is upperbound by [5]:

$$\mathbf{P_b} \leq \frac{3}{4} \frac{\left(\Gamma sin^2(\frac{\pi}{I})\right)^{-1}}{\left(1 + \sqrt{1 + \left(\Gamma sin^2(\frac{\pi}{I})\right)^{-1}}\right)\left(\sqrt{1 + \left(\Gamma sin^2(\frac{\pi}{I})\right)^{-1}}\right)} \qquad N = 1, \qquad (3.30)$$

where

$$\Gamma = \frac{\phi(1,0)^2}{(\phi(0,0) + 1)^2 - \phi(1,0)^2}, \qquad (3.31)$$

and $\phi(i,j)$ is the autocorrelation function defined in (3.11). Note that (3.30) can also be used to determine the bit error probability of uncoded MPSK with perfect coherent detection. In that case, simply replace the term Γ in (3.30) by the signal-to-noise ration E_s/N_o; see [5].

We show in Figure 3.2 the analytical results for the error performance of uncoded 4-DPSK in a static Rayleigh fading channel ($f_D T = 0$). Also showed is the result for 4-PSK with perfect coherent detection. It is observed that the original 3 db gap between coherent and conventional differential detection can be narrowed to 2 dB by using a multiple-symbol differential detector with a word length, N, equal to 5. Another 1 dB gain can be achieved if N is increased to 10. The performance improvement of these two detectors, however, are achieved at the expense of higher decoding complexities. Note that in general, a multiple-symbol differential detector requires to perform I^N metric computations to decode $N log_2(I)$ bits of information.
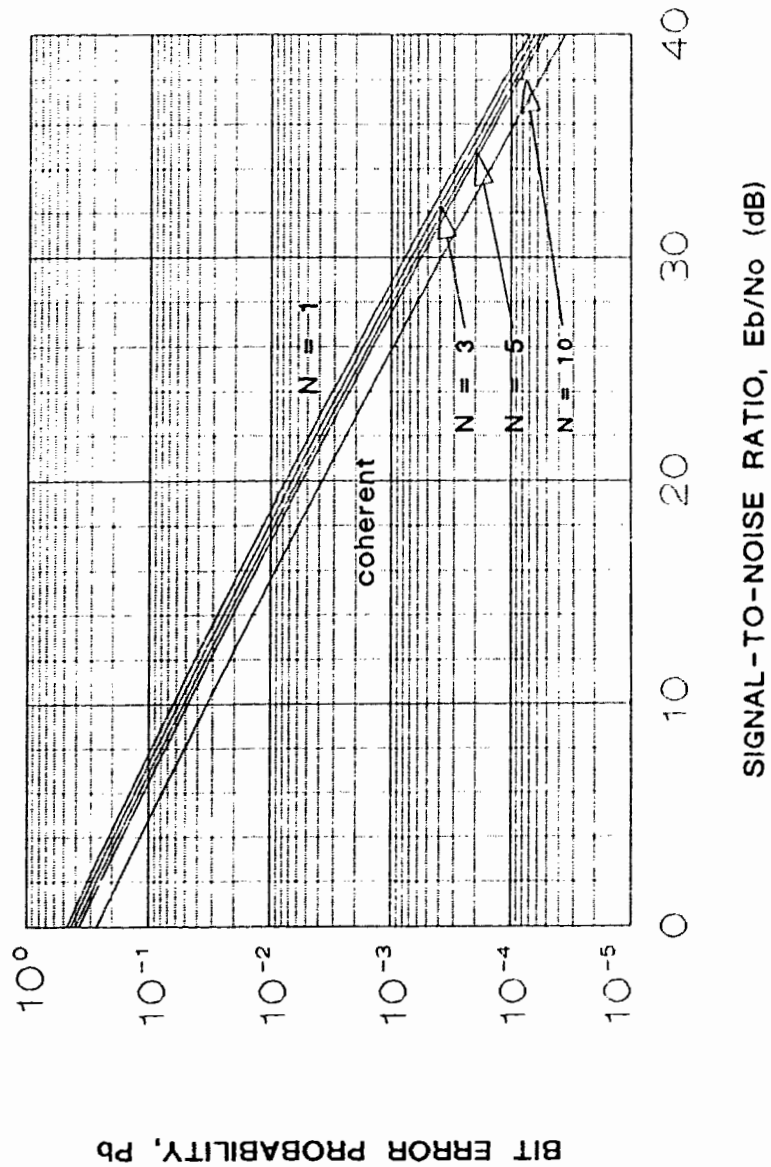
Figure 3.2: Error performance of 4-PSK in a static Rayleigh fading channel. The parameter N is word length used by the multiple-symbol differential detector. $N = 1$ corresponds to conventional differential detection.

46

The scenario for which multiple-symbol differential detection is most useful is when the mobile unit is traveling at very high speed. As mentioned earlier, the maximum Doppler frequency is linearly proportional to the speed of the vehicle. We show in Figure 3.3 the error performance of 4-PSK when the normalized Doppler frequency[1] is equal to 0.01 (corresponds roughly to the case in which the carrier frequency is 900 MHz, the vehicle speed is 60 miles/hour, and the signaling rate is 8 Ksymbol/s). It is observed from Figure 3.3 that for the bit error rate above 10^{-5}, the irreducible error floor associated with the conventional differential detector can be eliminated by using a multiple-symbol detector with N as small as 2. When $N = 3$, the gap between differential and coherent detection is roughly 4 dB at a bit error rate of 10^{-4}. If complexity is not an issue, then a differential detector with $N = 10$ can be used to narrow the gap further down to 2 dB. Similar observations are made for a Doppler frequency of 0.03. It is shown in Figure 3.4 that for this Doppler frequency, a detector with $N = 2$ can practically eliminate the irreducible error floor associated with a conventional differential detector at the bit error rate above 10^{-5}. When the word length is $N = 5$, the gap between coherent detection and differential detection is about 4 dB at an error probability of 10^{-4}. Another dB gain can be achieved by using a detector with $N = 10$.

Recall that the analytical results showed in Figure 3.2 - 3.4 are obtained by assuming the bit error probability of a multiple-symbol detector is dominated by those sequences $(\hat{z}_1, \ldots, \hat{z}_N)$, which have the highest correlation with the sequence $(z_1, \ldots, z_N)$. To check the validity of this assumption, we show also in Figure 3.4 the simulation results for the $N = 3$ detector. It is observed that the simulation results

---

[1]Normalized by signaling rate, i.e., $\frac{Maximum\ Doppler\ frequency}{Signaling\ rate}$.

47
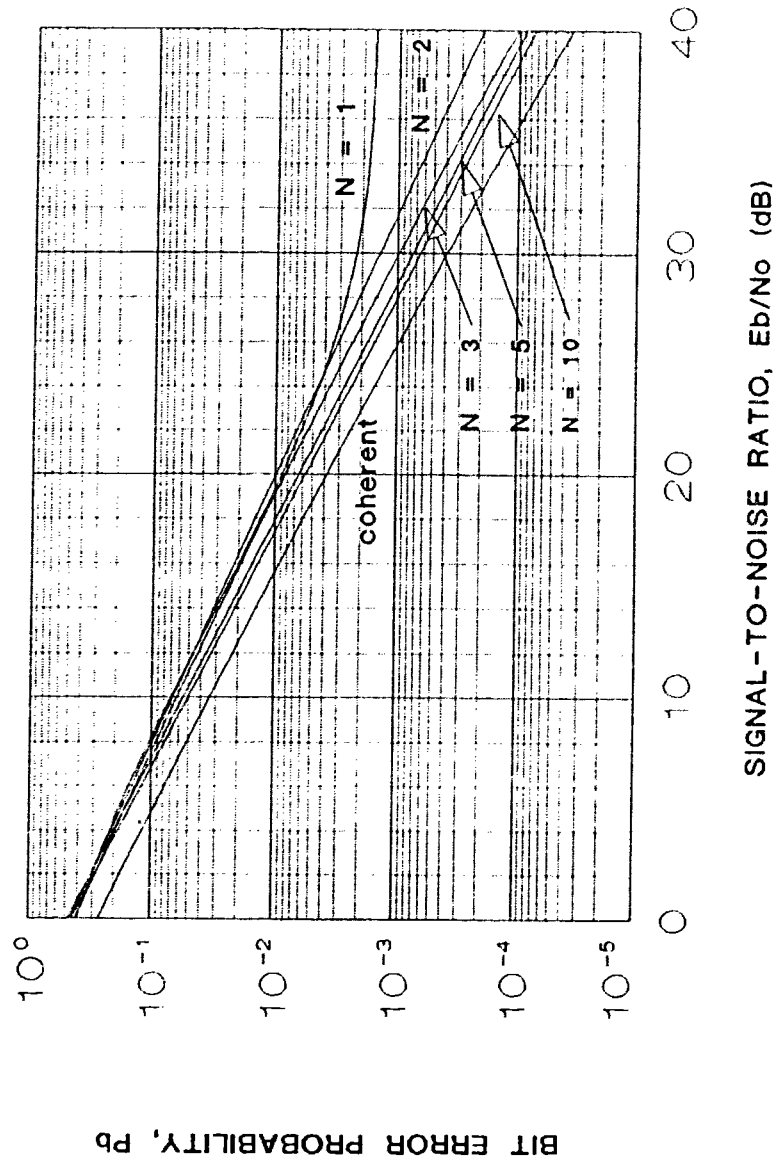
Figure 3.3: Error performance of 4-PSK in a Rayleigh fading channel with a Doppler frequency of 0.01.

48
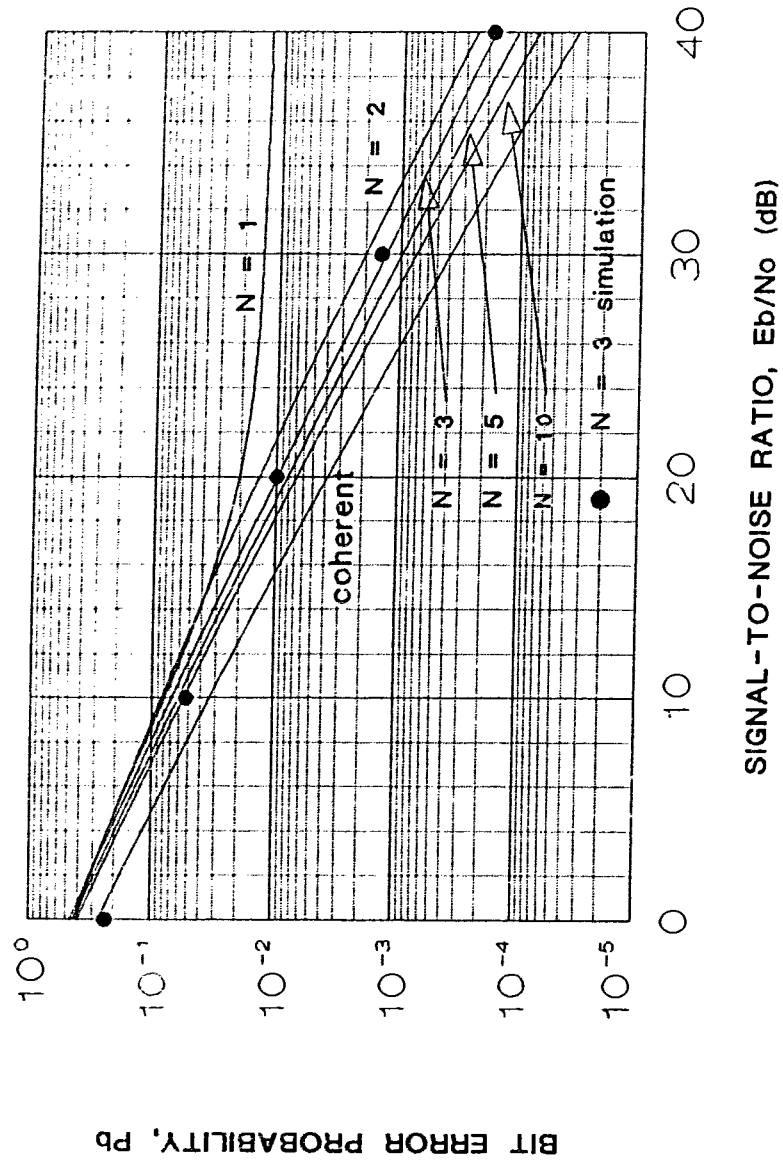
Figure 3.4: Error performance of 4-PSK in a Rayleigh fading channel with a Doppler frequency of 0.03. The circles are the simulation results for the $N = 3$ detector.

agree very well with the analytical results at large signal-to-noise ratios (equal to or above 20 dB). At lower SNR, the assumption appears to be slightly pessimistic.

As mentioned in Section 3.1.2, a practical implementation of the multiple-symbol differential detector requires the receiver to estimate from the received samples the autocorrelation function of the channel fading process. Clearly, estimation errors are unavoidable and this may result in a decoding metric that is mismatched to the channel statistics. We show in Figure 3.5 the error performance of a $N = 3$ detector that uses a decoding metric which is optimum for $f_D T = 0.01$ while the actual Doppler frequency has a different value. Two cases are considered:

1. the actual Doppler frequency is 0 (static fading),

2. the actual Doppler frequency is 0.02.

It is observed that the maximum degradation caused by a mismatch between the decoding metric and the channel statistics is about 2 dB at an error probability of 10^{-4} and this is for the case when the actual Doppler frequency is zero while the estimated Doppler frequency is 0.01. The effect of the mismatch is rather minimal if the actual Doppler frequency is 0.02. It should be pointed out that the results in Figure 3.5 are obtained by replacing the matrix $\mathbf{F}$ in step 3 of page 31 by the one derived from the estimated Doppler frequency. On the other hand, the matrix $\mathbf{\Phi_{rr}}$ in step 3 of page 31 is left unchanged and it still represents the covariance matrix of the received samples associated with the actual Doppler frequency. Moreover, Figure 3.5 is optimistic in the sense that the shape of the autocorrelation function of the fading process (in this case is the Bessel function of order zero, see (3.11)) is known to the receiver. If there is also a mismatch in the shape of the autocorrelation

50
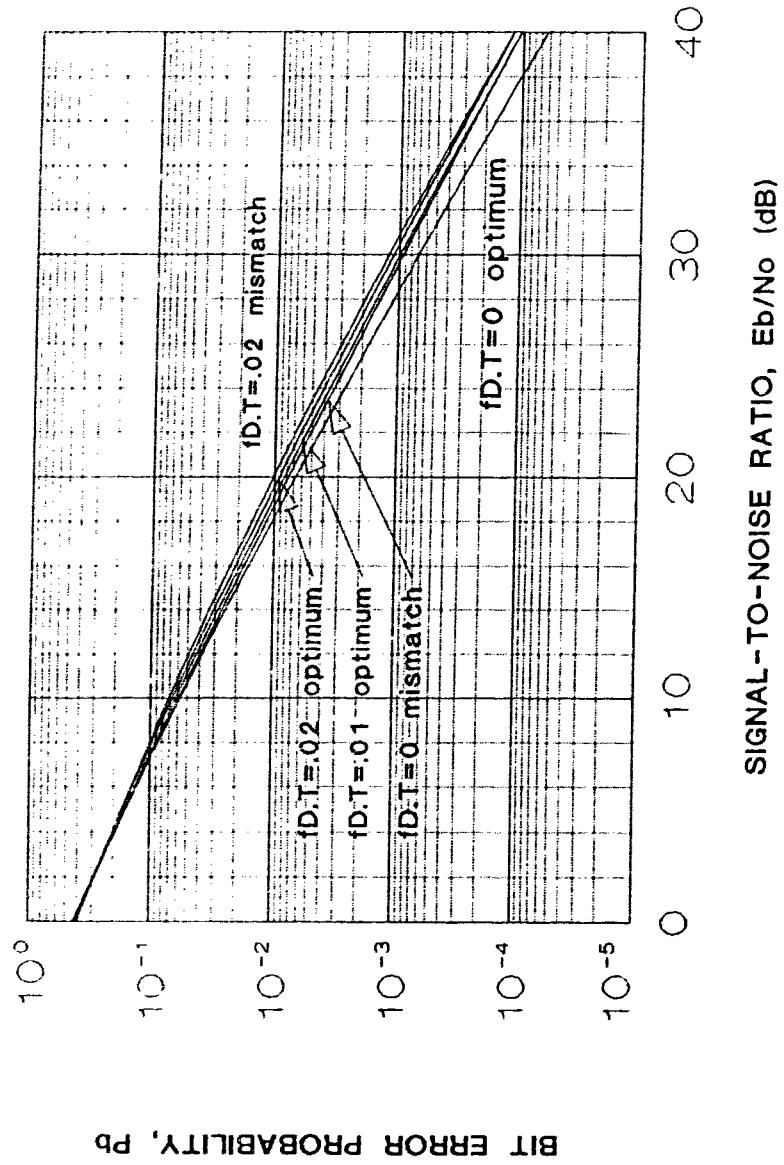
Figure 3.5: Error performance of a multiple-symbol differential detector whose decoding metric is mismatched to the channel statistics.

function, more degradation in the error performance will be resulted.

3.2 Pilot Symbol Assisted Demodulation of uncoded BPSK

In this section, we will use the general theory outlined in Chapter 2 to carry out the analysis for uncoded BPSK signals with pilot symbol assisted demodulation (PSAM) in correlated Rayleigh fading channel. Unlike [6], the analysis shown here is based on a block decoding strategy rather than symbol by symbol detection. The concept of PSAM is to insert a known symbol periodically into the transmitted sequence so that the fading distortion can be estimated at the receiver. Since there are additional pilot symbols being transmitted, the effective bandwidth will be reduced. Despite this throughput drawback, PSAM does not change the transmitted pulse shape [6] and the processing at transmitter and receiver is also simpler than with conventional pilot tone.

3.2.1 System Description

The frame structure for a PSAM system is shown in Figure 3.6 where 1 pilot symbol is added to every N data symbols. Since block demodulation is used in this analysis. The decision on any block of N data symbols will be based on the N received samples for that block, as well as on the

52
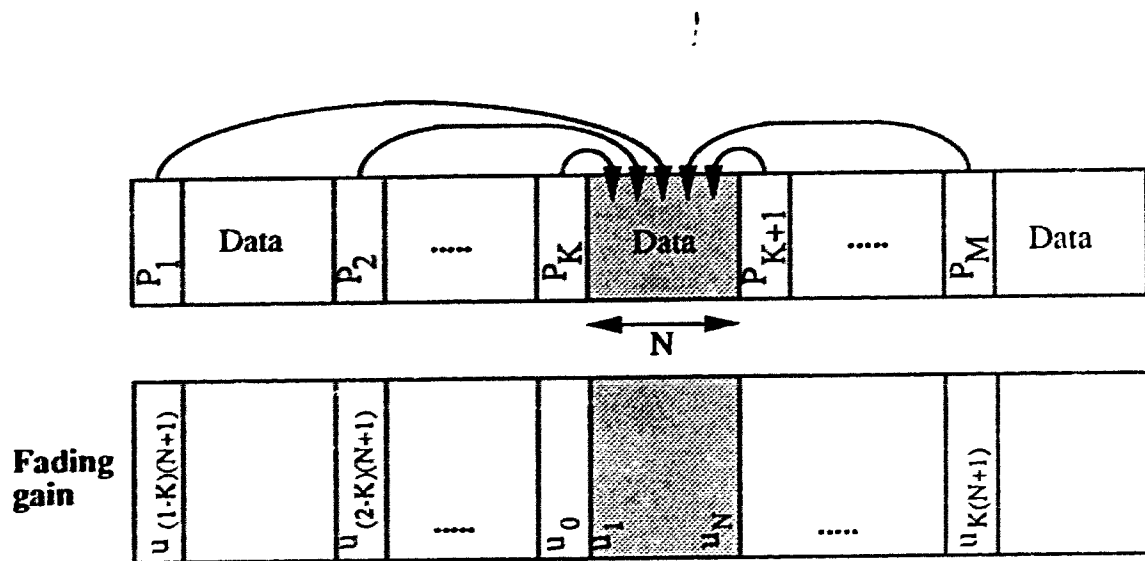
Figure 3.6: Transmitted frame structure of PSAM.

$$M = 2K \tag{3.32}$$

received pilot symbols surrounding that block. This decoding strategy is very similar to the multiple-symbol differential detection in previous section, except that in the multiple-symbol differential detection system, the effective "pilot" symbol is the symbol at the end of the previous block. As a result, multiple-symbol differential detection can be viewed as a pilot symbol assisted modulation technique with the "pilot" symbol added without bandwidth expansion.

To start our analysis, we use $c = (c_1, \ldots, c_N)$ and $p = (p_1, \ldots, p_M)$ to denote the data symbols and the pilot symbols respectively. Adopting the notation in Section 2.2, the complex received baseband data signal at $t = kT$ is denoted by $y_k = u_k c_k + n_k$ in (2.10) and the covariance matrix Φ_{yy} for the random vector y is defined in (2.21). For PSAM, the u_k's are:

$$u_k = \frac{A}{\sqrt{N_o}} g_k \tag{3.33}$$

where g_k once again represents the channel complex gain in the kth interval. By comparing (3.33) with (3.4), it is not difficult to see that the covariance matrix of the u_k's is

$$\Phi_{uu} = \Phi_{22} \tag{3.34}$$

where Φ_{22} was defined in (3.10).

On the other hand, the received pilot symbols are:

$$x_k = p_k u_{(k-K)(N+1)} + e_k, \qquad k = 1, 2, \ldots, M \qquad (3.35)$$

where K is defined in (3.32) and the e_k's are iid zero mean complex Gaussian random variables with a unit variance. Now, the random vector $\mathbf{x} = (x_1, \ldots, x_M)^t$ in (2.9) can be written in matrix form as

$$\mathbf{x} = \mathbf{Pv} + \mathbf{e} \qquad (3.36)$$

where

$$\mathbf{P} = \begin{pmatrix} p_1 & & \\ & \ddots & \\ & & p_M \end{pmatrix}, \qquad (3.37)$$

$$\mathbf{v} = \left(u_{(1-K)(N+1)}, u_{(2-K)(N+1)}, \ldots, u_{K(N+1)} \right)^t, \qquad (3.38)$$

and

$$\mathbf{e} = (e_1, \ldots, e_M)^t. \qquad (3.39)$$

This implies the matrix $\mathbf{\Phi_{xx}}$ in (2.17) becomes

$$\mathbf{\Phi_{xx}} = \mathbf{P}\mathbf{\Phi_{vv}}\mathbf{P}^\dagger + \mathbf{I} \qquad (3.40)$$

where

$$\Phi_{vv} = \begin{pmatrix} \phi((1-K)N',(1-K)N') & \phi((1-K)N',(2-K)N') & \ldots & \phi((1-K)N',KN')) \\ \phi((2-K)N',(1-K)N') & \phi((2-K)N',(2-K)N') & \ldots & \phi((2-K)N',KN') \\ \vdots & \vdots & & \vdots \\ \phi(KN',(1-K)N') & \phi(KN',(2-K)N') & \ldots & \phi(KN',KN') \end{pmatrix};$$

$$(3.41)$$

with $N' = N + 1$ and $\phi(i,j)$ was defined in (3.11). As for the matrix Φ_{xu} in (2.19), we have

$$\Phi_{xu} = \frac{1}{2}E\{(Pv + e)u\} = P\Phi_{vu} \qquad (3.42)$$

where Φ_{vu} is a matrix whose $(i,j)^{th}$ element, denoted by $\phi_{vu}(i,j)$ is

$$\phi_{vu}(i,j) = \phi((i-K)(N+1),j). \qquad \begin{cases} i = 1,2,\ldots,2K \\ j = 1,2,\ldots,N \end{cases} \qquad (3.43)$$

Once again, all the $\phi(i,j)$'s in the above covariance matrices are defined in (3.11).

Recall the term E_s is the received signal energy per symbol of a non PSAM detection system. Since a pilot symbol is inserted into every N data symbols. If the same amount of energy is used in a PSAM system, the effective data signal energy is only $(\frac{N}{N+1})E_s$. Therefore, for PSAM system, the effective signal to noise ratio should be

$$SNR = \frac{N}{N+1}\frac{E_s}{N_c}. \qquad (3.44)$$

3.2.2 The Optimum Decoder

For PSAM, the decoding metric can be obtained by putting the Φ_{xu} in (3.42) and

$$
\begin{aligned}
\Phi_{xx}^{-1} &= (P\Phi_{vv}P^\dagger + I)^{-1} \\
&= (P\Phi_{vv}P^\dagger + PIP^\dagger)^{-1} \\
&= P(\Phi'_{vv})^{-1}P^\dagger
\end{aligned}
\tag{3.45}
$$

where

$$
\Phi'_{vv} = \Phi_{vv} + I
$$

into (2.28) and (2.29) and then, substituting

$$
a = [\Phi_{22} + I - \Phi_{uv}\Phi_{vv}'^{-1}\Phi_{vu}]^{-1},
\tag{3.46}
$$

and

$$
b = \Phi_{uv}\Phi_{vv}'^{-1}P^\dagger,
\tag{3.47}
$$

into (2.27). The same expressions, when substituted into (2.31) gives us the F matrix. Once we have Φ_{rr} and F, we can then proceed to calculate the pairwise error event probability as outlined in page 31.

3.2.3 The Error Performance

In this sub-section, we will consider the error performance of uncoded BPSK signals with PSAM. To find an approximation on the average bit error probability of the

PSAM system, we must first find the pairwise error probability. As before, we can follow the four steps developed in Section 2.4 (page 31) to calculate the pairwise error probability. The bit error rate for uncoded BPSK, with the Doppler frequency of the Rayleigh fading channel, the length of data block N, and the length of pilot symbols sequence M as parameters, will be presented and compared with that of a perfect coherent detector.

As in the multiple-symbol differential detection analysis, an approximation of the bit error probability can be obtained by summing a set of dominant pairwise error event probabilities. In otherwords, we are only interested in those error events which the data sequence $c = (c_1, \ldots, c_N)$ and the erroneous sequence $\hat{c} = (\hat{c}_1, \ldots, \hat{c}_N)$ have the highest correlation. For a PSAM system, the correlation coefficient is defined as (see (3.27) as well):

$$\mu = \left| M + \sum_{i=1}^{N} c_i^* \hat{c}_i \right|^2 .$$

(3.48)

Without loss of generality, the data symbols c_i's are assumed to be 1. This implies that the correlation coefficient can then be written as

$$\mu = \left| M + \sum_{i=1}^{N} \hat{c}_i \right|^2 .$$

(3.49)

Therefore, for BPSK signaling, there are N erroneous sequences $\hat{c} = (\hat{c}_1, \ldots, \hat{c}_N)$ such that the correlation coefficient in (3.49) is maximized. Those are the sequences with only one symbol $\hat{c}_i$ equals to -1 and the rest of the symbols in the sequences are all equal to 1.

For comparison purpose, we include the bit error probability of coherent BPSK

[25]:

$$P_b \approx \frac{1}{4 \times SNR} \qquad (3.50)$$

in all Figures. We show in Figure 3.7 the error performance of uncoded BPSK with the length of data block equal to 6, and the length of pilot sequence equal to 11. The performance for the block PSAM decoder agrees very well with the symbol by symbol detector in [6]. However, the block decoder performs slightly better than the symbol by symbol detector in fast fading channels. For example, when the fade rate equals to 0.05, the gap between block decoding PSAM and coherent detection is only 1.5 dB at a bit error rate of 10^{-4} while the gap for the symbol by symbol decoding PSAM shown in [6] is 3 dB. This improvement implies that block decoding is more appropriate in fast fading environment. Figure 3.8 and Figure 3.9 show the error performance at a fade rate equal to 0.05 with different data length N and different lengths of pilot sequence M as parameters. We observe that as the length of the data block increases, the bit error rate decreases subsequently. This is in agreement with the report in [6] and [26]. The reason being that as the length of the data block increases, the effective signal energy increases; see (3.44). However, we cannot keep increasing the size of the data block in order to get better performance. Since the pilot symbol spacing depends on the data block size, the larger block size means the pilot symbols separate further apart and hence, their correlation is less significant. Following this line of reasoning and observation from Figure 3.8 and 3.9, the number of pilot symbols used for decoding is less important for larger data block. In other words, the difference between using 4 pilot symbols and 20 pilot symbols for decoding a 6 symbols data block is very small; see Figure 3.9. Fewer pilot symbols used for decoding imply less delay is required for filling up the pilot symbols buffer.
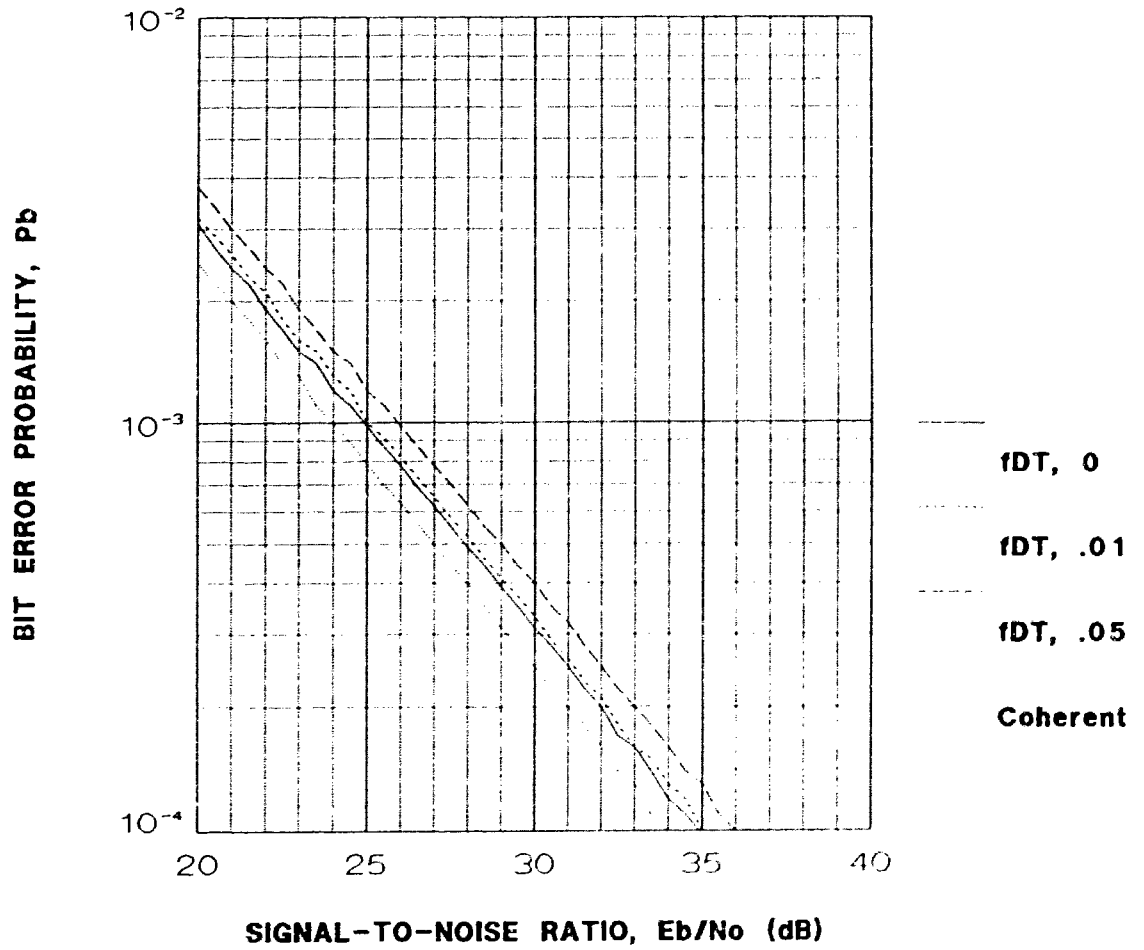
Figure 3.7: Error performance of uncoded BPSK with PSAM and $M = 11$, $N = 6$ $f_D T = 0, 0.01, 0.05$.
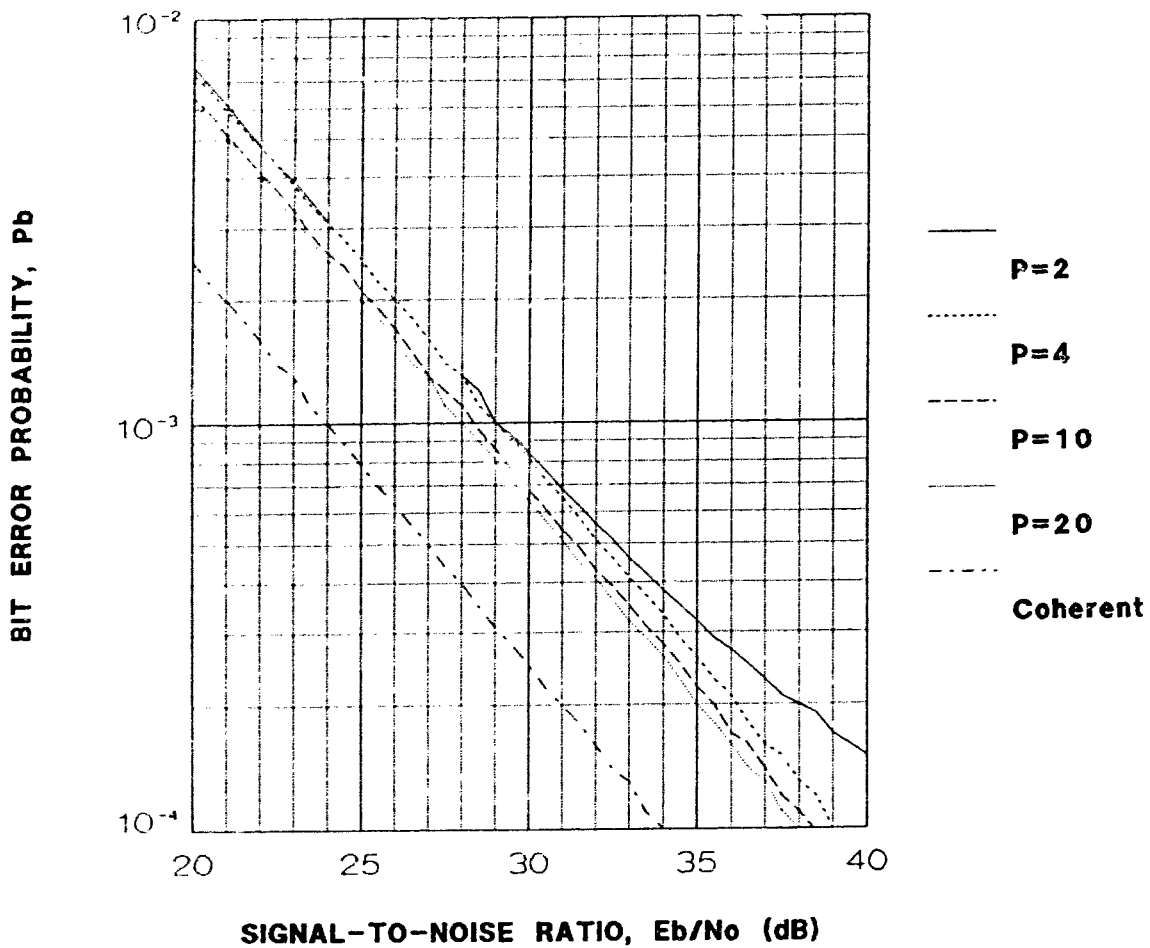
Figure 3.8: Error performance of uncoded BPSK with PSAM and $N = 1$, $f_D T = 0.05$
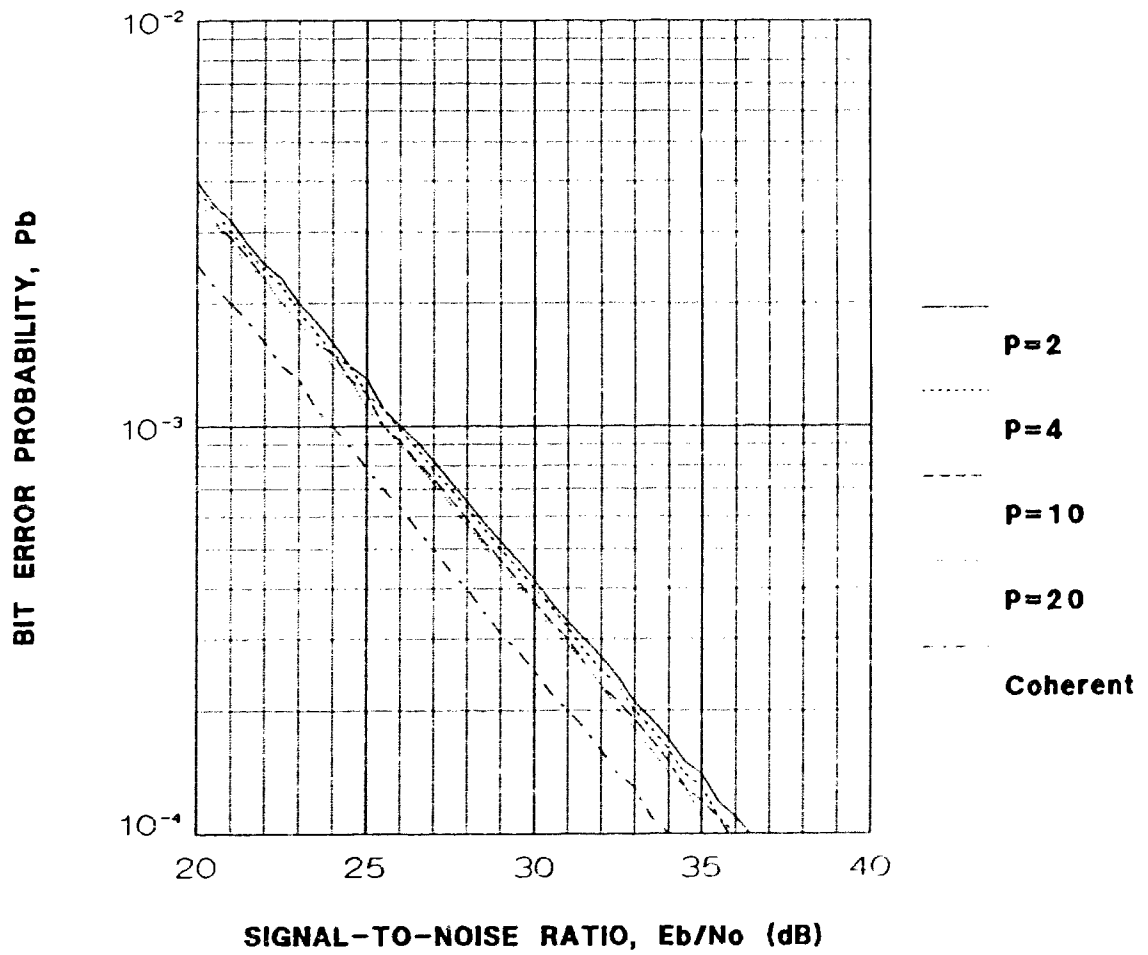
Figure 3.9: Error performance of uncoded BPSK with PSAM and $N = 6$, $f_D T = 0.05$

3.3 Summary

In this chapter, two examples have been used to demonstrate the general analysis presented in Chapter 2. In the first part, the pairwise error event probability of uncoded PSK signals with multiple-symbol differential detection has been derived from the optimum decoder. Error performance at different fade rates have been presented. For the bit error rate above 10^{-5}, it is found that multiple-symbol decoding strategy can remove the error floor associated with conventional symbol by symbol differential decoding. In the second part, the pairwise error event probability of uncoded BPSK signals with pilot symbol assisted modulation (PSAM) has been developed. Results for different fade rate, data block size N and length of pilot symbol sequence M are presented. We observed that in fast fading channels, the block decoder used in this thesis for PSAM performs slightly better that the symbol by symbol detector reported in [6].

CHAPTER 4

PERFORMANCE OF INTERLEAVED TCM IN CORRELATED RAYLEIGH FADING CHANNELS

We consider in this chapter the performance of interleaved, Trellis coded PSK signals transmitted over flat Rayleigh fading channels. Unlike [5], [8], [9], and [12], the systems consider here have finite interleaving depth. This implies after de-interleaving, the fading gains experienced by the different transmitted symbols are correlated. We will study the error performance of various TCM schemes with the interleaving depth as a parameter. First, we will make use of the general theory developed in Chapter 2 to analyze the interleaved, coded PSK signals with perfect CSI. An exact expression for the pairwise error event probability of interleaved, coded co-

herent PSK signals is obtained. Second, we will study the error performance of interleaved, coded DPSK signals. However, the optimum decoding metric in (2.27) cannot be directly adopted. The reason being that the metric in (2.27) is used for block decoding and is not directly applicable in the Viterbi decoding of TCM. As a result, we use a suboptimal decoding metric which was employed in [9]. Although the decoding metric is changed, the four steps listed on page 31 can still be used to determine the pairwise error event probability of interleaved, coded DPSK signals. Finally, by making use of the pairwise error event expressions, we can obtain an approximation of the bit error probability. Bit error curves of various interleaving depths at different maximum Doppler frequency will be shown subsequently. Since TCM will be of our interest, we will start this chapter by introducing some of the basic concept and error performance of TCM.

4.1 Ungerboeck's Trellis Codes

Trellis-coded modulation (TCM) is a coding scheme with which channel coding and modulation are combined as one entity. It is very suitable for those communication systems where both bandwidth and power limitations are imposed simultaneously. TCM achieves coding gain without any bandwidth expansion. At first it may seem that this statement violates some basic power-bandwidth trade-off principle. However, there is still a trade-off at work; namely TCM achieves coding gain at the expense of decoder complexity.

TCM combines a multilevel/phase modulation signaling set with a state-oriented trellis coding scheme. Multilevel/phase signal sets are signal constellations having

multiple amplitudes, multiple phases, or a combinations of multiple amplitudes and multiple phases. A trellis code is one that can be characterized with a trellis diagram, like the one showed in Figure 4.1. The dots in Figure 4.1 represent the states and the branches represent transitions between states. During a transition, the encoder will emit a signal alphabet by considering its present state and the input bits. For example, if the encoder is in state S0 and the input bits are 11, the encoder will send out the symbol labeled 6 or the complex symbol $exp(j\frac{6\pi}{4})$ to the channel. TCM is based on the fact that redundancy is introduced by increasing the number of signal alphabet through multilevel/phase signaling, so that no redundant symbols are transmitted. For example, if the number of symbols required in the alphabet is 2^m for an uncoded system, TCM uses 2^{m+1} for the coded system. This implies there are $2^{m+1} - 2^m$ symbols used for redundancy. However, for a constant average power, the minimum distance between adjacent signal alphabet decreases as the number of signal alphabet increases. In uncoded modulation, assuming a constant average power, the reduction in the minimum distance between the signal alphabet degrades the error performance. As a result, trellis coding must be able to increase the minimum distance between the signals that are most likely to be confused, without increasing the average power. Consider the trellis in Figure 4.1, since there are 8 PSK symbols available while there are only 4 possible transitions from each state, coding here means the proper assignment of PSK symbols to the encoder transitions so as to maximize the free Euclidean distance. The concept of mapping by set partitioning [32] is used to achieve that. Figure 4.2 shows the set partitioning of 8-PSK constellation. This technique divides a signal set successively into smaller subsets with maximally increasing smallest intra-set distance Δ_i^2, $i = 0, 1, 2$. Each partition is two-way. After the first partition, the signal subset A and B obtained
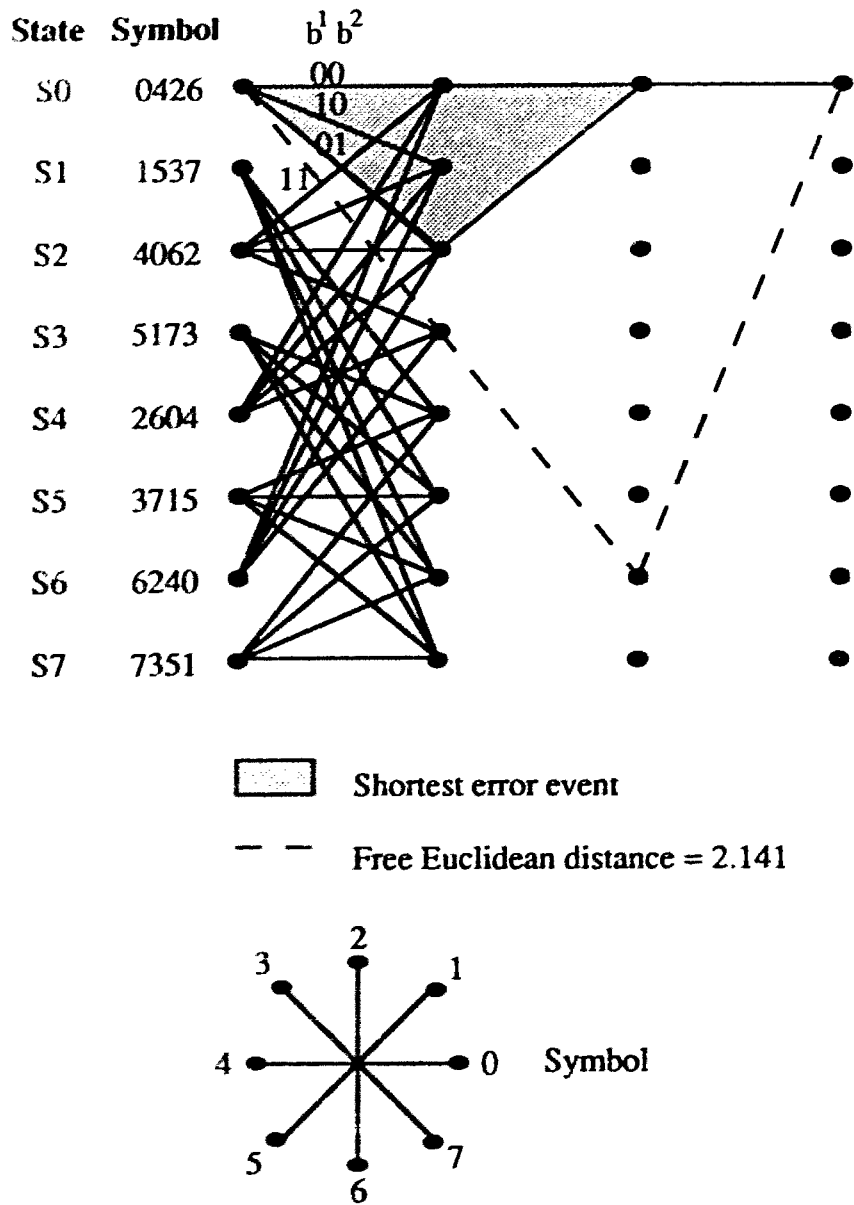
Figure 4.1: Trellis diagram of Ungerboeck's 8 state code. Branch labels indicate the binary information vector b_k associated with transitions from any state.

$\Delta_0^2 = 0.586$

Subset A c^1 0   1   Subset B

$\Delta_1^2 = 2$

c^2 0   1   0   1

A0   A1   B0   B1

$\Delta_2^2 = 4$

c^3 0   1   0   1   0   1   0   1

$c^1 c^2 c^3$ 000   001   010   011   100   101   110   111
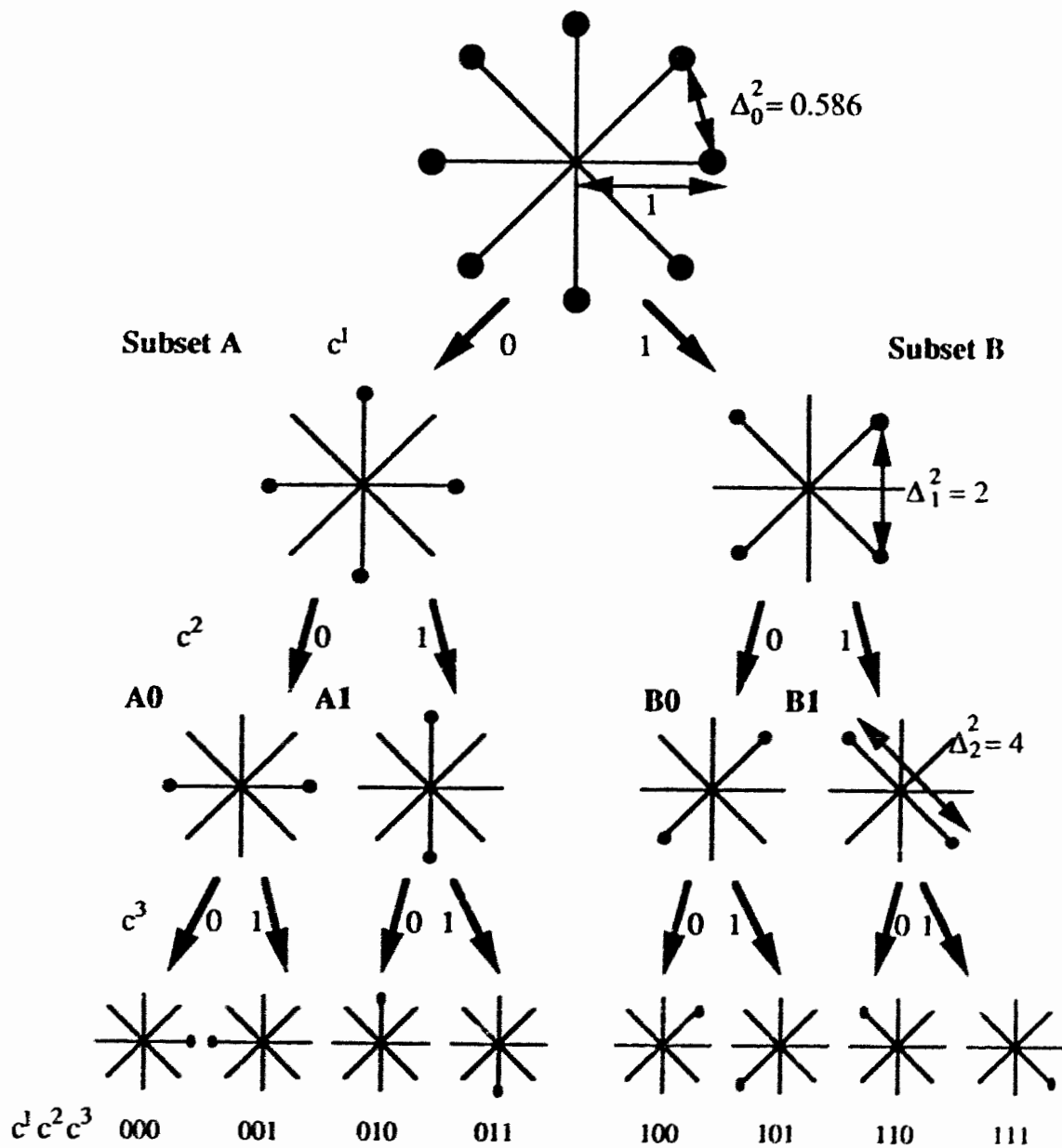
Figure 4.2: Set partitioning of the 8-PSK constellation.

by taking the even signal points and the odd signal points respectively, increase the intra-set distance Δ_i^2 from 0.586 to 2. The partitioning is repeated once more so that the smallest intra-set distance further increases to 4. If we assign symbols from the same subset (A or B) to those transitions originating from the same state and those terminating at the same state in Figure 4.1, we will have at least a squared free Euclidean distance of

$$\Delta_i^2 + \Delta_i^2 = 4.$$

This is the lower bound for the squared free Euclidean distance since it only represents the squared distance accrued during the initial split and the final remerge of two paths in the trellis and there may be additional distances accrued during other intervals. Compared with uncoded 4-PSK where the minimum squared distance is 2, the simple 8-state code in Figure 4.1 provides a coding gain (without bandwidth expansion).

The following is a summary of Ungerboeck's code design rules for AWGN channels:

1. All channel (8 PSK) symbols should occur with equal frequency and with a fair amount of regularity and symmetry.

2. Transitions originating from or merging in the same state receive signals either from subset A or B.

3. Parallel transitions receive signals either from subset A0 or A1, or B0 or B1.

These rules guarantee reasonably good codes for the AWGN channels. However, to

find out the best code for a given encoder structure, one has to rely on computer search.

The trellis encoder of Figure 4.1 can be implemented in the form of a rate $2/3$ convolutional encoder followed by an 8-PSK signal constellation and bit-to-symbol mapper, see Figure 4.3.

In summary, trellis coding for bandlimited channels employs larger signal alphabet achieved through multilevel/phase signaling (e.g. MPSK), such that channel bandwidth is not increased. Although the increase in signal set size reduces the minimum distance between symbols, the free Euclidean distance between trellis code sequence more than compensates for the signal points being crowded together [29]. The result is a net error performance gain of at least 3 dB without any bandwidth expansion [32].

4.1.1   The Performance of TCM

It is well known that the appropriate criterion for optimum TCM design on the AWGN channel is the maximization of the free Euclidean distance [32]. However, reports from [10] and [37] suggested that the performance of TCM in (severe) fading environment may no longer depend solely upon the maximization of the free Euclidean distance. The primary concern in such environment is to maximize the code diversity or the "length" of the shortest error event path. In fact, the "length" of an error event path is just the Hamming distance between the transmitted sequence and the selected sequence, counted by the channel symbols. For example, Figure 4.4 illustrates two error event paths. The sequence marked by $c = \ldots, c_1, c_2, c_3, \ldots$
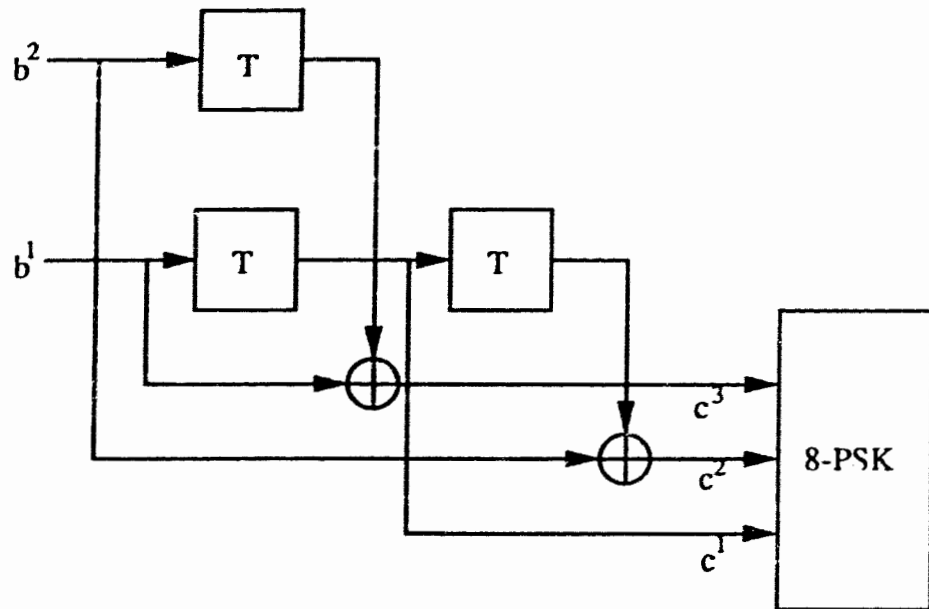
Symbol	c^1	c^2	c^3
0	0	0	0
4	0	0	1
2	0	1	0
6	0	1	1
1	1	0	0
5	1	0	1
3	1	1	0
7	1	1	1
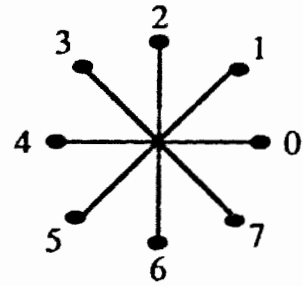
Figure 4.3: Realization of Ungerboeck's 8-state code using a rate 2/3, 8-state Convolutional code.
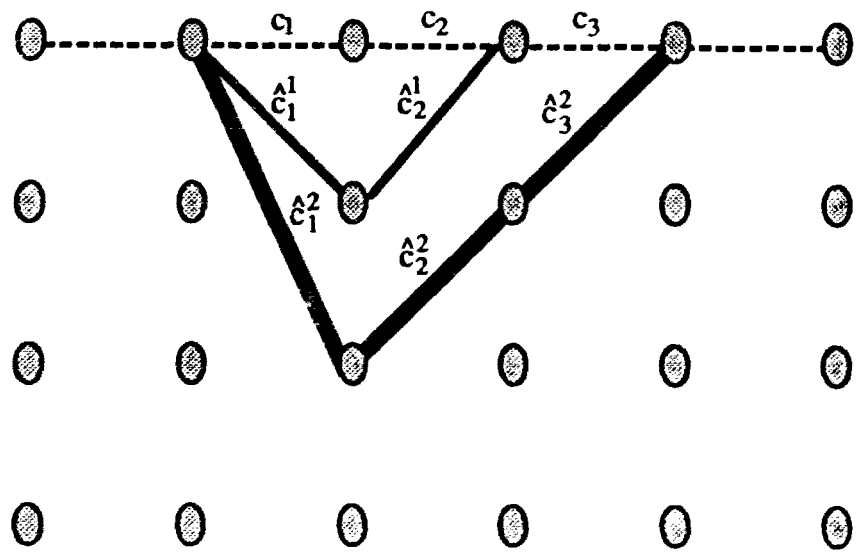
Figure 4.4: Examples of error events.

is the transmitted sequence where c_j's are the channel symbols. There are two sequence $\hat{c}^1$ and $\hat{c}^2$ which diverge and then remerge with the transmitted sequence through different paths. Assume that $\hat{c}_j^i \neq c_j$. Then, the length of error events associated with the sequence $\hat{c}^1$ and $\hat{c}^2$ are 2 and 3 respectively. Since there is no parallel transition and $\hat{c}^1$ has the smallest length, it is the shortest error event and the code diversity of this trellis code is 2. On the other hand, if fading effects are minimal, the optimum performance is once again achieved by a trellis code designed to maximize the free Euclidean distance.

Now, let's consider some examples of Ungerboeck's Trellis codes. The rate 2/3, 8-state, 8-PSK trellis code and the rate 2/3, 16-state, 8-PSK trellis code are the first TCM schemes found by Ungerboeck using the rules of set partitioning outlined in page 69. The trellis diagram of the 8-state code is shown in Figure 4.1 and the trellis diagram of the 16-state code and the structure of the trellis encoder is shown in Figure 4.5. In both diagrams, the shaded area depicts the shortest error event while the dash line denotes the free Euclidean distance. In addition, we assume the transmitted codeword is represented by the path at the top of the trellis. As written on the diagrams, the free Euclidean distance for the 8-state code is only 2.141 while the free Euclidean distance for the 16-state code is 2.274. The increment in free Euclidean distance causes the 16-state code achieved a 4.1 dB coding gain over uncoded 4-PSK while the coding gain for the 8-state code is only 3.6 dB. Thus, in terms of the coding gain, the 16-state code is more powerful than the 8-state code in the AWGN channel. In addition, the 16-state code is also more suitable for fading channels since it has a code diversity order of 3 while the 8-state code has a code diversity order of 2. However, since more states are involved, finding the set of dominant error events for the 16-state code is also more time consuming than

State	Symbol
S0	0426
S1	1537
S2	40(2
S3	5137
S4	2604
S5	3715
S6	6240
S7	7351
S8	4062
S9	5137
S10	0426
S11	1537
S12	6240
S13	7351
S14	2604
S15	3715
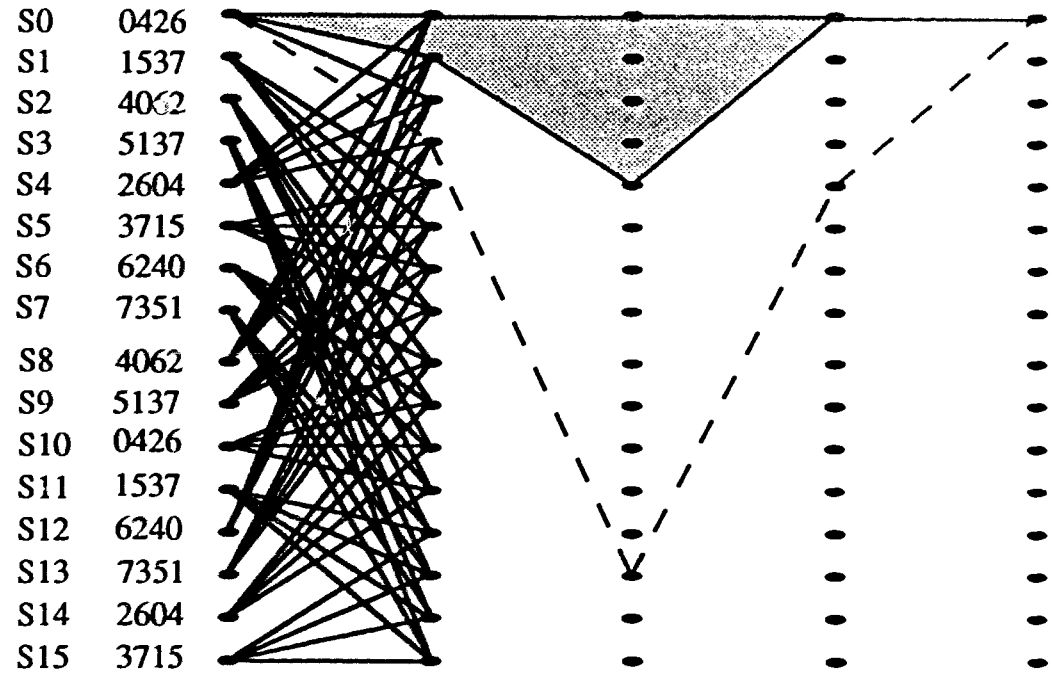
Shortest error event

Free Euclidean distance = 2.274
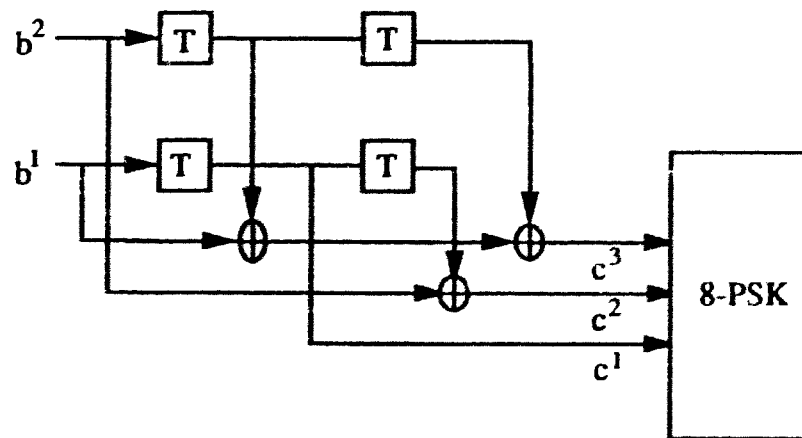
Figure 4.5: Realization of Ungerboeck's 16-state code using a rate 2/3, 16-state convolutional code.

74

the 8-state code. For the sake of simplicity, we will mainly study the Ungerboeck's 8-state, 8-PSK trellis code in this thesis.

4.2 System Model and Notation for Interleaved PSK Systems

The discrete-time system model for the interleaved, coded PSK symbols transmitted over correlated Rayleigh fading channel will be discussed in this section. Although we have presented similar materials in Chapter 2, a more detailed description for the TCM system with interleaver will be shown in this section. Both systems with perfect CSI and differential detection will be presented.

The block diagram of the system is shown in Figure 4.6. The input to the trellis encoder is a sequence of binary digits denoted by $\mathbf{m} = (m_1, \ldots, m_k, \ldots)$, and the output is a sequence of coded PSK symbols denoted by $\mathbf{c} = (c_1, \ldots, c_k, \ldots, c_N)$. In order to disperse possible deep fades in the channel, the sequence of modulation symbols is passed to a block interleaver with a buffer of size α rows by β columns, where β is the interleaving depth, and $\beta\alpha = N$ is the length of each TCM codeword; see Figure 1.2. The modulation symbols in the codeword $\mathbf{c} = (c_1, \ldots, c_k, \ldots, c_N)$ will fill the interleaver buffer column by column, and the output sequence of the interleaver is denoted by $\mathbf{c}' = (c_1', \ldots, c_k', \ldots, c_N')$. Let the time index k in c_k be written as $k = b\alpha + a + 1$, where $0 \le b \le \beta - 1$ and $0 \le a \le \alpha - 1$. This implies that after interleaving, the symbol c_k will occupy the $(a + 1, b + 1)^{th}$ position in the interleaver buffer, where $a + 1$ is the row number and $b + 1$ is the column number. According to Figure 1.2, c_k will be sent out in the $(a\beta + b + 1)^{th}$ time
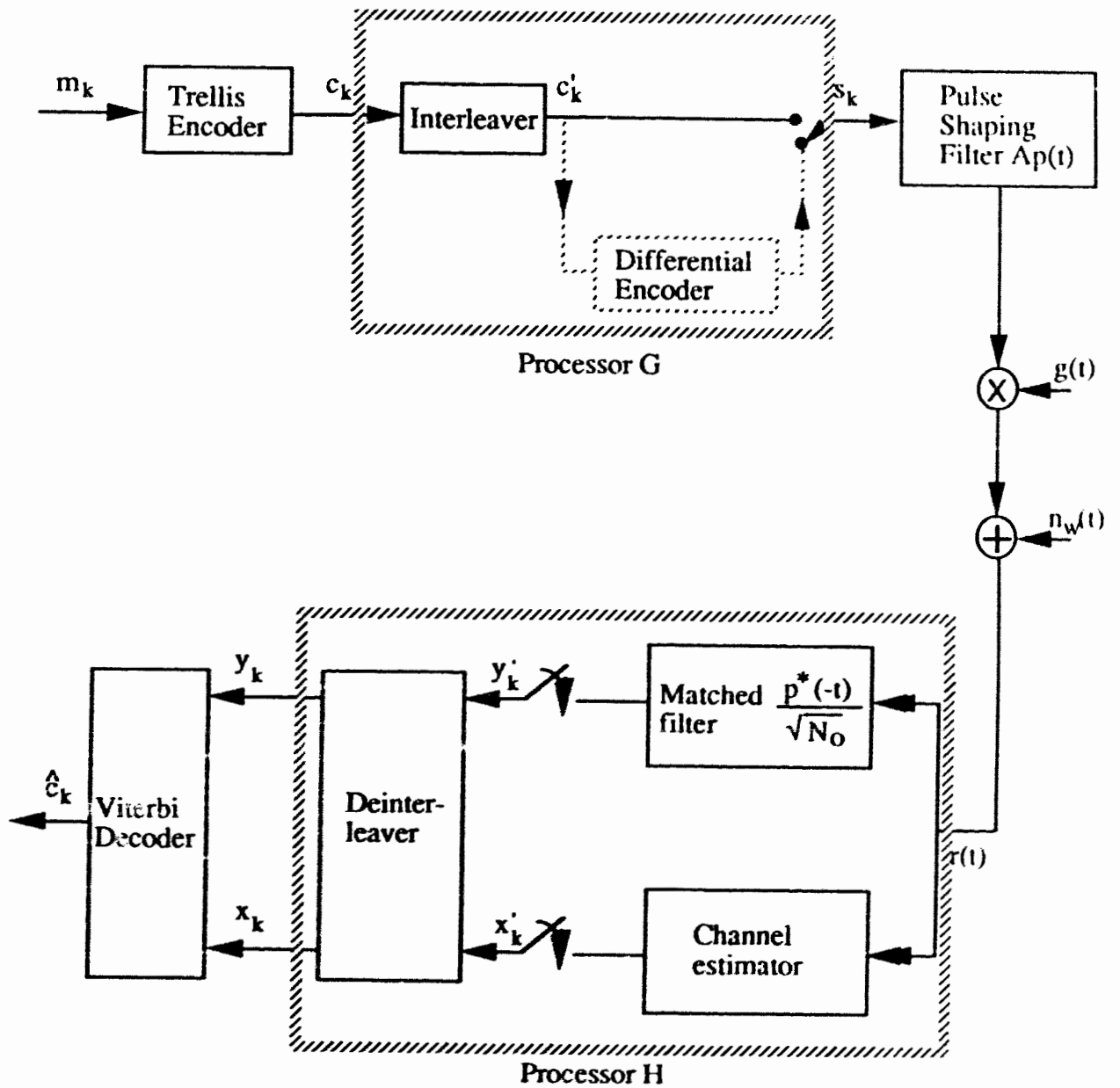
Figure 4.6: Block diagram of the Trellis-coded system.

slot. The relationship between the two sequences $\mathbf{c} = (c_1, \ldots, c_k, \ldots, c_N)$ and $\mathbf{c}' = (c'_1, \ldots, c'_k, \ldots, c'_N)$ is thus:

$$c_{b\alpha+a+1} = c'_{a\beta+b+1}.$$

(4.1)

Since in practical systems, the interleaving depth β is finite. This implies that, after deinterleaving on the receiver side, the complex gains, or fading, experienced by different modulation symbols in the codeword are correlated.

For DPSK, the sequence $\mathbf{c}' = (c'_1, \ldots, c'_k, \ldots, c'_N)$ at the interleaver output will be differentially encoded into another sequence before transmission. In either case, the transmitted sequence is denoted by $\mathbf{s} = (s_1, \ldots, s_k, \ldots, s_N)$ where

$$s_k = \begin{cases} c'_k & \text{Coherent PSK} \\ c'_k s_{k-1} & \text{Differential PSK} \end{cases}$$

(4.2)

Without loss of generality, we assume $s_0 = 1$ for DPSK.

Consider the transmission of MPSK signal over a Rayleigh fading channel with AWGN, as showed in Figure 4.6, the received signal is passed to a matched filter and the output is then sampled every T seconds to produce the sequence $(y'_1, \ldots, y'_k, \ldots, y'_N)$. The sample y'_k can then be written as (see Section 3.1.1 and 3.2.1)

$$y'_k = \frac{A}{\sqrt{N_o}} g_k s_k + n'_k.$$

(4.3)

As before, the n'_k's in the above equation are independent and identically distributed (iid) complex Gaussian random variables representing samples of the filtered channel noise. Each n'_k's has a zero mean and a unit variance. The samples in the sequence

$(y_1', \ldots, y_k', \ldots, y_N')$ are deinterleaved to produce the sequence $(y_1, \ldots, y_k, \ldots, y_N)$. Using (4.1), it is easy to see that for any $0 \leq a \leq \alpha - 1$ and $0 \leq b \leq \beta - 1$, the samples in the sequences $(y_1', \ldots, y_k', \ldots, y_N')$ and $(y_1, \ldots, y_k, \ldots, y_N)$ have the relationship that $y_{b\alpha+a+1} = y_{a\beta+b+1}'$. Using the notation in (2.10), the deinterleaved sample y_k can be written as

$$y_k = u_k c_k + n_k$$

where

$$u_k = \begin{cases} \frac{A}{\sqrt{N_o}} g_{a\beta+b+1} & \text{Coherent PSK} \\ \frac{A}{\sqrt{N_o}} g_{a\beta+b+1} s_{a\beta+b} & \text{DPSK} \end{cases} \qquad k = b\alpha + a + 1 \qquad (4.4)$$

represents the fading experienced by the kth data symbol, and

$$n_k = n_{a\beta+b+1}' \qquad k = b\alpha + a + 1 \qquad (4.5)$$

is a sample of the filtered channel's additive white Gaussian noise.

Due to a finite interleaving depth, the u_k's in (4.4) are a set of correlated, zero mean, complex Gaussian random variables. Let two of the symbols in the TCM codeword, c_{k_1} and c_{k_2}, occupy respectively the $(a_1+1, b_1+1)^{th}$ and the $(a_2+1, b_2+1)^{th}$ positions in the interleaver buffer, where $0 \leq a_1, a_2 \leq \alpha - 1$ and $0 \leq b_1, b_2 \leq \beta - 1$. This implies that if $k_1 = b_1\alpha + a_1 + 1$, $k_2 = b_2\alpha + a_2 + 1$, and the complex gains u_{k_1} and u_{k_2} experienced by these two symbols have a correlation equal to:

$$\phi(k_1, k_2) = \frac{1}{2} E\{u_{k_1} u_{k_2}^*\}$$

$$
= \begin{cases} \rho(\beta(a_1 - a_2) + b_1 - b_2) & \text{Coherent PSK} \\ \rho(\beta(a_1 - a_2) + b_1 - b_2)s_{a_1\beta+b_1}s^*_{a_2\beta+b_2} & \text{DPSK} \end{cases} \qquad (4.6)
$$

where

$$
\begin{aligned}
\rho(k) &= \frac{A^2}{N_o}\frac{1}{2}E\{g^*_m g_{m+k}\} \\
&= \frac{A^2}{N_o}R_g(kT) \\
&= \frac{E_s}{N_o}J_0(2\pi k f_D T), \qquad (4.7)
\end{aligned}
$$

and $R_g(\tau)$ is the autocorrelation function defined in (2.4). Note that if c_{k_1} and c_{k_2} are placed in the same column in interleaver buffer, then b_1 must equal to b_2, and $\phi(k_1, k_2)$ becomes $\rho(\beta(k_1 - k_2))$ for coherent PSK.

The channel estimator in Figure 4.6 will extract from the received signal information about the channel's complex gains. Let the sequence that appears at the estimator output be denoted by $(x'_1, \ldots, x'_k, \ldots, x'_N)$. In the most ideal situation, the estimator provides perfect CSI. This implies

$$
x'_k = \frac{A}{\sqrt{N_o}}g_k \qquad \text{Perfect CSI.} \qquad (4.8)
$$

Although techniques like pilot tone and embedded pilot symbols can be used to provide close to perfect CSI, for applications like mobile radio, the simpler differential detection is sometimes preferred. As pointed out in [5], differential detection can be viewed as a channel estimation technique for which the symbol received in the previous signaling interval is used as an estimate of the channel's complex gain in the current interval. Therefore, in DPSK

79

$$x'_k = y'_{k-1} = \frac{A}{\sqrt{N_o}} g_{k-1} s_{k-1} + n'_{k-1}. \qquad DPSK \qquad (4.9)$$

In either coherent PSK or DPSK, the samples in the sequence $(x'_1, \ldots, x'_k, \ldots, x'_N)$ are deinterleaved to produced the sequence $(x_1, \ldots, x_k, \ldots, x_N)$. Once again, if $k = b\alpha + a + 1$ (with $0 \le a \le \alpha - 1$ and $0 \le b \le \beta - 1$), then

$$x_k = \begin{cases} u_k & \text{Perfect CSI} \\ \frac{A}{\sqrt{N_o}} g_{a\beta+b} s_{a\beta+b} + n'_{a\beta+b} & \text{DPSK} \end{cases}, \qquad k = b\alpha + a + 1. \qquad (4.10)$$

The $n'_{a\beta+b}$'s in (4.10) are iid complex Gaussian random variables each having a zero mean and unit variance.

4.3 Coded PSK with Perfect CSI

We will provide in this section two derivations of the pairwise error event probability of interleaved TCM with perfect CSI. The first derivation is the one given in Chapter 2. The second derivation is a simpler one and it allows us to express the error probability in terms of the branch squared distances between the coded sequences.

4.3.1 The Optimum Decoder

For the system with perfect CSI, the vector $x = (x_1, \ldots, x_N)^t$ and the gain vector $u = (u_1, \ldots, u_N)^t$ are identical. Consequently, the covariance matrices Φ_{xx}, Φ_{xu} in (2.17) and (2.19) are identical. This implies the matrices a and b in (2.28) and

(2.29) are both an identity matrix of size N. As a result, the optimal decoder in (2.27) becomes

$$M(\hat{c}) = \begin{pmatrix} x^t & y^t \end{pmatrix} \begin{pmatrix} 0 & -\hat{C}^\dagger \\ -\hat{C} & I \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{4.11}$$

Clearly this is equivalent to selecting the codeword whose metric

$$M(\hat{c}) = \sum_{k=1}^{N} (y_k^* \hat{c}_k x_k + x_k^* \hat{c}_k^* y_k) \tag{4.12}$$

is the largest.

4.3.2 The Pairwise Error Event Probability

Let the transmitted codeword $c = (c_1, \ldots, c_N)$ and the erroneous codeword $\hat{c} = (\hat{c}_1, \ldots, \hat{c}_N)$ be different in the $k_1, k_2, \ldots, k_L$ positions. For convenience, we assume $k_i < k_j$ if $i < j$. Moreover, we let each k_i be written as $k_i = b_i\alpha + a_i + 1$ where $0 \le a_i \le \alpha - 1$, $0 \le b_i \le \beta - 1$, and α and β are the number of rows and columns in the interleaver buffer; see Figure 1.2. From the optimal decoding metric in (4.12), we see that if the random variable

$$\begin{aligned} D &= M(\hat{c}) - M(c) \\ &= \sum_{i=1}^{L} \{ y_{k_i} x_{k_i}^* (\hat{c}_{k_i}^* - \hat{c}_{k_i}^*) + y_{k_i}^* x_{k_i} (c_{k_i} - \hat{c}_{k_i}) \} \end{aligned} \tag{4.13}$$

is less than zero, then a decoding error will occur. The random variable D can be written in matrix form as:

$$D = \begin{pmatrix} \tilde{\mathbf{x}}^\dagger & \tilde{\mathbf{y}}^\dagger \end{pmatrix} \begin{pmatrix} \mathbf{0} & \tilde{\boldsymbol{\Delta}}^\dagger \\ \tilde{\boldsymbol{\Delta}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{y}} \end{pmatrix} \tag{4.14}$$

where

$$\begin{align}
\tilde{\mathbf{y}} &= (y_{k_1}, y_{k_2}, \dots, y_{k_L})^t, \tag{4.15} \\
\tilde{\mathbf{x}} &= (x_{k_1}, x_{k_2}, \dots, x_{k_L})^t, \tag{4.16} \\
\tilde{\boldsymbol{\Delta}} &= \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_L \end{pmatrix}, \tag{4.17}
\end{align}$$

with

$$d_i = c_{k_i} - \hat{c}_{k_i}. \tag{4.18}$$

If we let

$$\tilde{\mathbf{r}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{y}} \end{pmatrix}, \tag{4.19}$$

then using the results from Chapter2, the poles of the characteristics function of the random variable D are the reciprocals of the eigenvalues of

$$\tilde{\mathbf{G}}' = \tilde{\boldsymbol{\Phi}}_{\mathbf{rr}} \tilde{\mathbf{F}} \tag{4.20}$$

where

82

$$\tilde{\mathbf{F}} = \begin{pmatrix} \mathbf{0} & \tilde{\mathbf{\Delta}}^{\dagger} \\ \tilde{\mathbf{\Delta}} & \mathbf{0} \end{pmatrix},$$

(4.21)

and

$$\tilde{\mathbf{\Phi}}_{\mathbf{rr}} = \frac{1}{2}E\{\tilde{\mathbf{r}}\tilde{\mathbf{r}}^{\dagger}\}.$$

(4.22)

The covariance matrix $\tilde{\mathbf{\Phi}}_{\mathbf{rr}}$ can be obtained according to (2.17) - (2.23). Simply replace every matrix $\mathbf{M}$ in these equations by the corresponding matrix $\bar{\mathbf{M}}$.

Consider the matrix

$$\tilde{\mathbf{\Phi}}_{\mathbf{uu}} = \frac{1}{2}E\{\tilde{\mathbf{u}}\tilde{\mathbf{u}}^{\dagger}\}$$

where $\tilde{\mathbf{u}} = (u_{k_1}, \ldots, u_{k_L})^t$. Since $u_{k_i} = \frac{A}{\sqrt{N_o}}g_{a_i\beta+b_i+1}$, this implies the $(i,j)^{th}$ elements of $\tilde{\mathbf{\Phi}}_{\mathbf{uu}}$ is equal to the term $\phi(k_i, k_j)$ in (4.6). Now if all the L places that the codewords are different are confined to one single column, then all the b_i's are the same and $\phi(k_i, k_j)$ becomes $\rho(\beta(k_i - k_j))$. As mentioned earlier that in a Rayleigh fading channels, the error performance of a TCM scheme is dominated by the set of shortest error events. For example, the Ungerboeck's 8-state code in Figure 4.1 has a shortest error event of length 2. In an interleaver buffer of size α rows by β columns, there are only $\alpha - 1$ out of a total of $\alpha\beta - 1$ places where this error event can span over 2 columns. Consequently, most of the time we can assume this dominant event is confined to a single column with respect to the interleaver buffer. As a result, we let

83

$$\tilde{\Phi}_{uu} = \begin{pmatrix} \rho(0) & \cdots & \rho((k_1 - k_j)\beta) & \cdots & \rho((k_1 - k_L)\beta) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho((k_i - k_1)\beta) & \cdots & \rho((k_i - k_j)\beta) & \cdots & \rho((k_i - k_L)\beta) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho((k_L - k_1)\beta) & \cdots & \rho((k_L - k_j)\beta) & \cdots & \rho(0) \end{pmatrix}. \qquad (4.23)$$

Once $\tilde{\Phi}_{uu}$ is known, then we can show easily that

$$\tilde{\Phi}_{xx} = \tilde{\Phi}_{xu} = \tilde{\Phi}_{uu} \qquad (4.24)$$

and

$$\tilde{\Phi}_{yy} = \tilde{C}(\tilde{\Phi}_{uu} + I)\tilde{C}^{\dagger} \qquad (4.25)$$

where

$$\tilde{C} = \begin{pmatrix} c_{k_1} & & \\ & \ddots & \\ & & c_{k_L} \end{pmatrix}. \qquad (4.26)$$

By arranging these submatrices according to (2.20), we have the covariance matrix $\tilde{\Phi}_{rr}$. This matrix when combined with the matrix $\tilde{F}$ derived in (4.21) would allow us to compute the eigenvalues required in the error probability calculations.

4.3.3 Alternate Derivation of The Pairwise Error Probability

Although we have already obtained the error expression for Trellis-coded PSK with perfect CSI in the last section, we will derive below a simpler expression for this system. The new expression allows us to show the pairwise error probability for full interleaving and very slow fading in simple forms.

The optimal decoding metric for coded PSK with perfect CSI is shown in (4.12) and as shown in (4.13), a decoding error will occur if

$$
\begin{aligned}
D &= \sum_{i=1}^{L} y_{k_i} x_{k_i}^* (c_{k_i}^* - \hat{c}_{k_i}^*) + y_{k_i}^* x_{k_i} (c_{k_i} - \hat{c}_{k_i}) \\
&= \sum_{i=1}^{L} |h_i|^2 + h_i n_{k_i}^* + h_i^* n_{k_i},
\end{aligned}
\tag{4.27}
$$

is less than zero. Note that

$$
h_i = d_i u_{k_i},
\tag{4.28}
$$

where d_i is defined in (4.18) and the n_{k_i}'s are iid complex Gaussian random variables each having a zero mean and a unit variance. Equation (4.27) can also be written in matrix form as

$$
D = \mathbf{h}^\dagger \mathbf{h} + \tilde{\mathbf{n}}^\dagger \mathbf{h} + \mathbf{h}^\dagger \tilde{\mathbf{n}}
\tag{4.29}
$$

where $\mathbf{h} = (h_1, \dots, h_L)^t$ and $\tilde{\mathbf{n}} = (n_{k_1}, \dots, n_{k_L})^t$ are column vectors whose components are the h_i's and the n_{k_i}'s. The covariance matrix for the h_i's is

$$\mathbf{\Phi}_{hh} = \frac{1}{2}E\{hh^\dagger\} = \tilde{\mathbf{\Delta}}\tilde{\mathbf{\Phi}}_{uu}\tilde{\mathbf{\Delta}}^\dagger, \tag{4.30}$$

where $\tilde{\mathbf{\Delta}}$ is defined in (4.17) and $\tilde{\mathbf{\Phi}}_{uu}$ is defined in (4.23). Since all the elements of the matrix $\tilde{\mathbf{\Phi}}_{uu}$ in (4.23) depend only on the differences of the k_i's, we can assume, without loss of generality, that an error event always starts at the beginning of each codeword, i.e., $k_1 = 1$.

Since $\mathbf{\Phi}_{hh}$ is a Hermitian matrix, it can be written in the form

$$\mathbf{\Phi}_{hh} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger, \tag{4.31}$$

where

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_L \end{pmatrix} \tag{4.32}$$

is a diagonal matrix whose elements, the λ_i's, are the eigenvalues of $\mathbf{\Phi}_{hh}$, and $\mathbf{U}$ is a unitary matrix whose columns are orthonormal and $\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}$. Let $w_1, w_2, \ldots, w_L$ be a set of independent, zero mean, complex Gaussian random variables, with the variance of w_i equal to

$$\sigma_{w_i}^2 = \frac{1}{2}E\{w_i w_i^*\} = \lambda_i. \tag{4.33}$$

It can be shown that the h_i's can be obtained from the w_i's by applying the following transformation:

$$\mathbf{h} = \mathbf{Uw},\qquad\qquad (4.34)$$

where $\mathbf{w} = (w_1, \ldots, w_L)^t$. Moreover, it can be shown that the random variable D in (4.29) can now be written in terms of the w_i's as

$$
\begin{aligned}
D &= \mathbf{w}^\dagger(\mathbf{U}^\dagger\mathbf{U})\mathbf{w} + (\tilde{\mathbf{n}}^\dagger\mathbf{U})\mathbf{w} + \mathbf{w}^\dagger(\mathbf{U}^\dagger\tilde{\mathbf{n}}) \\
&= \mathbf{w}^\dagger\mathbf{w} + \mathbf{e}^\dagger\mathbf{w} + \mathbf{w}^\dagger\mathbf{e} \\
&= \sum_{i=1}^{L}\{|w_i|^2 + w_i e_i^* + w_i^* e_i\}.
\end{aligned}
\qquad (4.35)
$$

where

$$\mathbf{e} = \mathbf{U}^\dagger\tilde{\mathbf{n}} = (e_1, \ldots, e_L)^t. \qquad\qquad (4.36)$$

Using the fact that $\tilde{\mathbf{\Phi}}_{\mathbf{nn}} = \mathbf{I}$ and that $\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}$, it can be shown that the e_i's are a set of iid Gaussian random variables each having a zero mean and a unit variance. Also, the e_i's are independent of the w_i's. At this point, it becomes clear that the random variable D is simply a sum of independent quadratic forms of complex Gaussian variates. Using (4B.5) in [25], we can show that the characteristic function of D is

$$\Phi_D(s) = \left(\prod_{i=1}^{L}\frac{1}{\lambda_i}\right)\left(\prod_{i=1}^{L}\frac{-1}{(s - p_{1i})(s - p_{2i})}\right) \qquad (4.37)$$

where

87

$$\begin{bmatrix} p_{1i} \\ p_{2i} \end{bmatrix} = \frac{1}{2} \mp \sqrt{\frac{1}{4} + \frac{1}{\lambda_i}} \tag{4.38}$$

are the poles of the characteristic function. By summing the residues of $\Phi_D(s)/s$ as shown in (2.36), we can obtain the pairwise error probability. The numerical results are identical to the one obtained in the previous sub-section. As mentioned before, the advantage of deriving this expression is that error expression for full interleaving and very slow fading can be easily derived.

4.3.3.1 Full Interleaving

Consider the case when we have full interleaving, that is, the interleaving depth β is a very large number. The matrix $\tilde{\Phi}_{uu}$ in (4.23) becomes $(E_s/N_o)\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. This implies the matrix Φ_{hh} in (4.30) is a diagonal matrix, given by

$$\Phi_{hh} = \frac{E_s}{N_o} \tilde{\Delta}\tilde{\Delta}^\dagger = \frac{E_s}{N_o} \begin{pmatrix} |d_1|^2 & & \\ & \ddots & \\ & & |d_L|^2 \end{pmatrix} \tag{4.39}$$

where $|d_i|^2$ is the squared Euclidean distance between c_{k_i} and $\hat{c}_{k_i}$; see (4.18). Consequently, the eigenvalues are:

$$\lambda_i = \frac{E_s}{N_o}|d_i|^2. \tag{4.40}$$

By substituting these eigenvalues into (4.37) – (4.38), we can obtain the expressions for the characteristic function and for the poles. All of the resulting expressions are

identical to those found in [5].

4.3.3.2 Very Slow Fading

When fading is very slow, i.e., $f_D T \approx 0$, it is quite reasonable to assume that all the components of the matrix $\tilde{\Phi}_{uu}$ in (4.23) are the same and they are equal to $\rho(0) = \frac{E_s}{N_o}$. As a result, Φ_{hh} becomes

$$\Phi_{hh} = \frac{E_s}{N_o} \tilde{\Delta} \mathbf{J} \tilde{\Delta}^\dagger \tag{4.41}$$

where $\mathbf{J}$ is a matrix with all its elements equal to unity. Because of this property of the matrix $\mathbf{J}$, the matrix Φ_{hh} has a rank equal to 1. As a result, it has only one non-zero eigenvalue. This non-zero eigenvalue, denoted by λ_1, is equal to the trace of the matrix. In otherwords,

$$\lambda_i = \begin{cases} \left(\frac{E_s}{N_o}\right) \sum_{j=1}^{L} |d_j|^2 & i = 1 \\ 0 & otherwise \end{cases} \tag{4.42}$$

and the characteristic function becomes $\Phi_D(s) = -\lambda_1(s - p_{11})(s - p_{21})^{-1}$. In fact, λ_1 is the sum of squared Euclidean distance between the transmitted and the erroneous word. Using (4.37) and (4.38), we can show that

$$P(\mathbf{c} \to \hat{\mathbf{c}}) = \frac{1}{2} \left(1 - \sqrt{\frac{\lambda_1}{4 + \lambda_1}} \right). \tag{4.43}$$

This is an intuitively pleasing result because when fading is slow, the Rayleigh channel is equivalent to a Gaussian noise channel on a per codeword basis. As a

result, the error performance of a TCM system should be determined by the squared Euclidean distance between two codewords.

4.3.4 The Error Performance

We presents in this section the error performance of Ungerboeck's 8-state 8-PSK code in a Rayleigh fading channel, with the interleaving depth as a parameter. In nearly all applications, we are interested in the overall bit error probability of the TCM scheme rather than the individual pairwise error event probability. An approximation to the bit error probability can be obtained by summing the pairwise error event probabilities as follows [5]:

$$\mathbf{P_b} \approx \frac{1}{m} \sum_{\hat{\mathbf{c}} \neq \mathbf{c}} a(\mathbf{c}, \hat{\mathbf{c}}) \mathbf{P}(\mathbf{c} \rightarrow \hat{\mathbf{c}}).$$

In the above equation, m is the number of input bits per encoding interval, $a(\mathbf{c}, \hat{\mathbf{c}})$ is the number of bit-errors associated with each error event, and the summation is taken over the set of dominant error events listed in Table 4.1. The error events listed in Table 4.1 are taken from Table 4.2 in [20] and they correspond to the shortest paths in the modified error state diagram [20] obtained via the method of Zehavi and Wolf [38]. As defined earlier, the parameters L and S in Table 4.1 are the length[1] and the span[2] of an error event. It should be pointed out that in Table 4.1, the notation $\gamma_i, i = 0, 1, 2, 3, 4$ is used to represent the situation where

---

[1]The Hamming distance between two coded sequences, counted by the channel symbols.

[2]The number of transition branches associated with the error event.

90

L	S	Error Events					a(c, ĉ)
2	2	γ_2	γ_4				1
3	3	γ_4	γ_1	γ_2			1
3	3	γ_4	γ_1	γ_2			1
3	3	γ_4	γ_3	γ_2			1
3	3	γ_2	γ_2	γ_4			2
3	3	γ_2	γ_3	γ_2			2
3	3	γ_2	γ_1	γ_2			$\frac{3}{2}$
3	3	γ_2	γ_3	γ_2			$\frac{3}{2}$
3	4	γ_4	γ_1	γ_0	γ_4		2
3	4	γ_2	γ_0	γ_1	γ_2		2
3	4	γ_2	γ_0	γ_1	γ_2		$\frac{3}{2}$
3	4	γ_2	γ_0	γ_3	γ_2		$\frac{3}{2}$
3	4	γ_2	γ_3	γ_0	γ_4		3
3	5	γ_2	γ_0	γ_1	γ_0	γ_4	3

Table 4.1: List of dominant error events in Ungerboeck's 8-state 8-PSK code.

the transmitted symbol and the erroneous symbol have an absolute phase difference of $i\pi/4$ radian. Different error events are represented by different sequences of the γ_i's in Table 4.1. For example, in the shortest error event (the first sequence in Table 4.1), the transmitted and the erroneous words are different in two consecutive places, with the first and second symbols in the two codewords having absolute phase differences of $\frac{\pi}{2}$ and π respectively. Since non-ideal interleaving is used in the system, the complex gains experienced by different transmitted symbols are correlated. Therefore, the order of the γ_i's in each sequence is important. Also since there are many ways in which two symbols in the 8-PSK constellation can have an absolute phase difference of $i\pi/4$, the error events listed in Table 4.1 should only be used if the error probability depends only on the absolute phase differences between symbol pairs in the transmitted and the erroneous words. It is shown in Appendix F that the eigenvalues for the matrix Φ_{hh} in (4.30), and consequently the error probability, depend only on the squared Euclidean distance between symbol pairs in the two codewords. As a result, we can use the set of error events listed in Table 4.1 in our error probability calculation. Without loss of generality, we assume the transmitted codeword is the all zero-phase codeword, i.e., $c = (1, 1, \ldots, 1)$. The erroneous codeword in each error event can then be determined by the corresponding sequence of phase differences.

The approximate bit error performance of Ungerboeck's 8-state code in a Rayleigh fading channel with a normalized Doppler frequency, $f_D T$, of 0.01, is shown in Figure 4.7. As one would expect, as the interleaving depth increases, the bit error probability decrease. When the interleaving depth is equal to 20 symbols, the bit error probability is almost identical to that provided by full interleaving. Furthermore, it should be pointed out that the full interleaving curve is same as the one in Figure
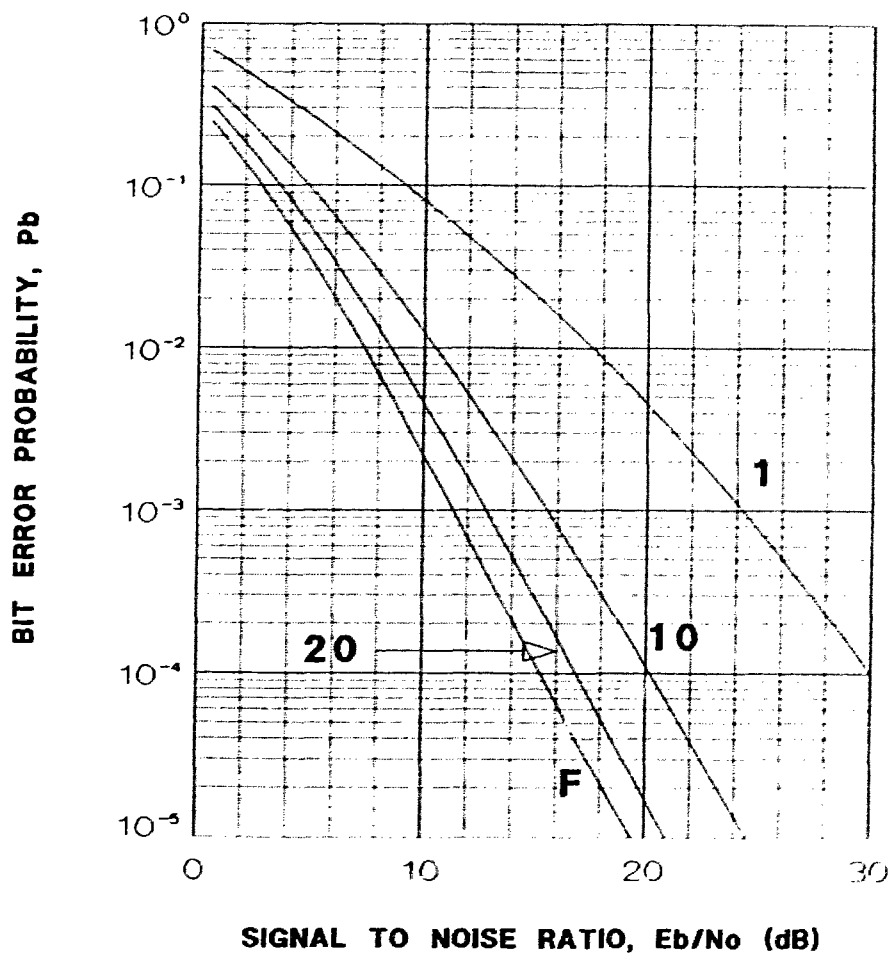
Figure 4.7: Bit error performance of Ungerboeck's 8-state code in a Rayleigh fading channel with perfect CSI. $f_D T = 0.01$. The number associated with each error curve is the interleaving depth. The curve labelled F represents the case of full interleaving.

93

4 in [5]. To see if other TCM schemes also require a similar interleaving depth, we show in Figure 4.8 the approximate bit error performance of Ungerboeck's 16-state code with the interleaving depth as a parameter. As mentioned in Section 4.1.1, to find the set of dominant error event for this 16-state code is a time consuming task. We will, therefore, obtain the bit error probability by only considering the shortest error event. This error event is shown in Figure 4.5 and has a length $L = 3$. The corresponding branch Euclidean distances in (4.18) are $d_1 = 1 - exp(j4\frac{\pi}{4})$, $d_2 = 1 - exp(j\frac{\pi}{4})$, and $d_3 = 1 - exp(j2\frac{\pi}{4})$. By comparing Figure 4.7 and 4.8, we see that for each of the interleaving depth considered, the 16-state code has a lower bit error probability than the 8-state code. As in the case of the 8-state code, an interleaving depth of 20 symbols provides almost the same performance as full interleaving. Let the fade cycle be defined as the reciprocal of the normalized Doppler frequency, $f_D T$. Then, by doing the appropriate normalization, we conclude that for both the 8 and 16-state coded, an interleaving depth roughly equivalent to one-fifth the duration of a fade cycle is sufficient to produce the same effect as full interleaving. In order to verify the above statement, the bit error performance of Ungerboeck's 8-state code with a normalized Doppler frequency of 0.03 is plotted in Figure 4.9. These results indicated that error performance of interleaving depth between 5 to 10 symbols is close to the case of ideal interleaving. As a confirmation of the analytical results and the assumption that the bit error probability is dominated by a set of short error events, simulation results (dash curves) for interleaving depth of 5 and 10 symbols are also provided in Figure 4.9. The simulation of the Rayleigh fading channel was done by generating samples of the time varying complex channel gain u_k. The real and imaginary parts of u_k were generated by filtering two statistically independent Gaussian random number sequences with a FIR digital low-pass filter.
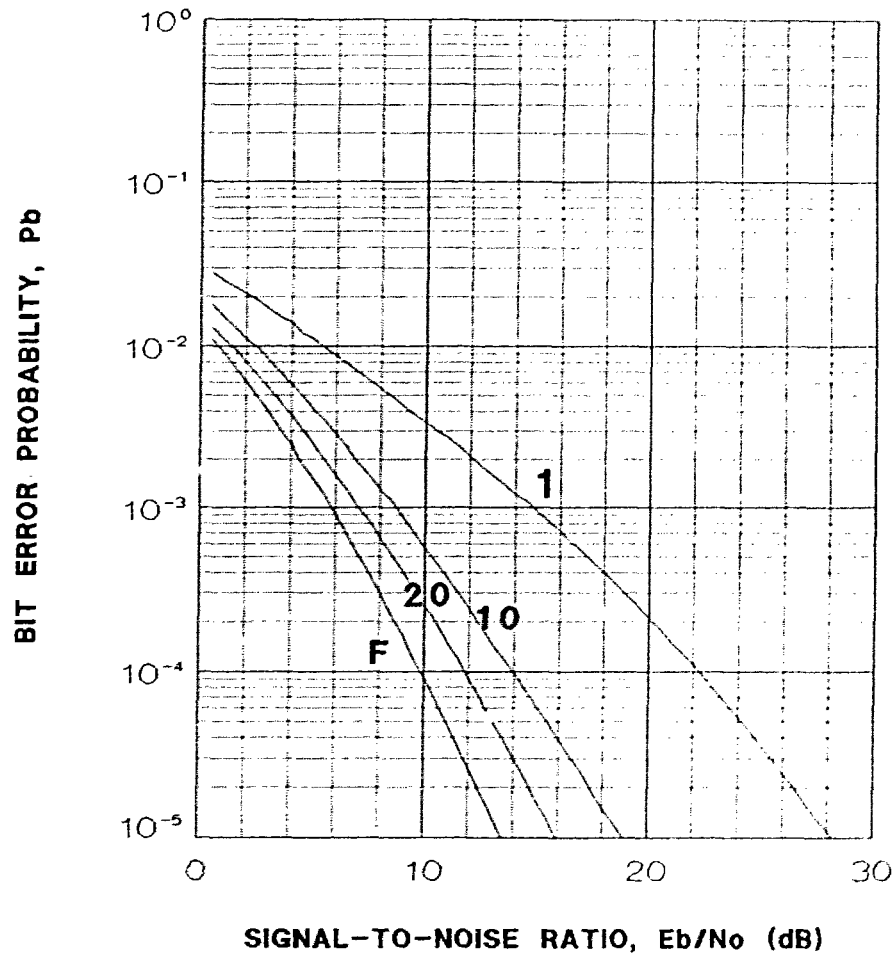
Figure 4.8: Bit error performance of Ungerboeck's 16-state code in a Rayleigh fading channel with coherent detection and perfect CSI. $f_D T = 0.01$.
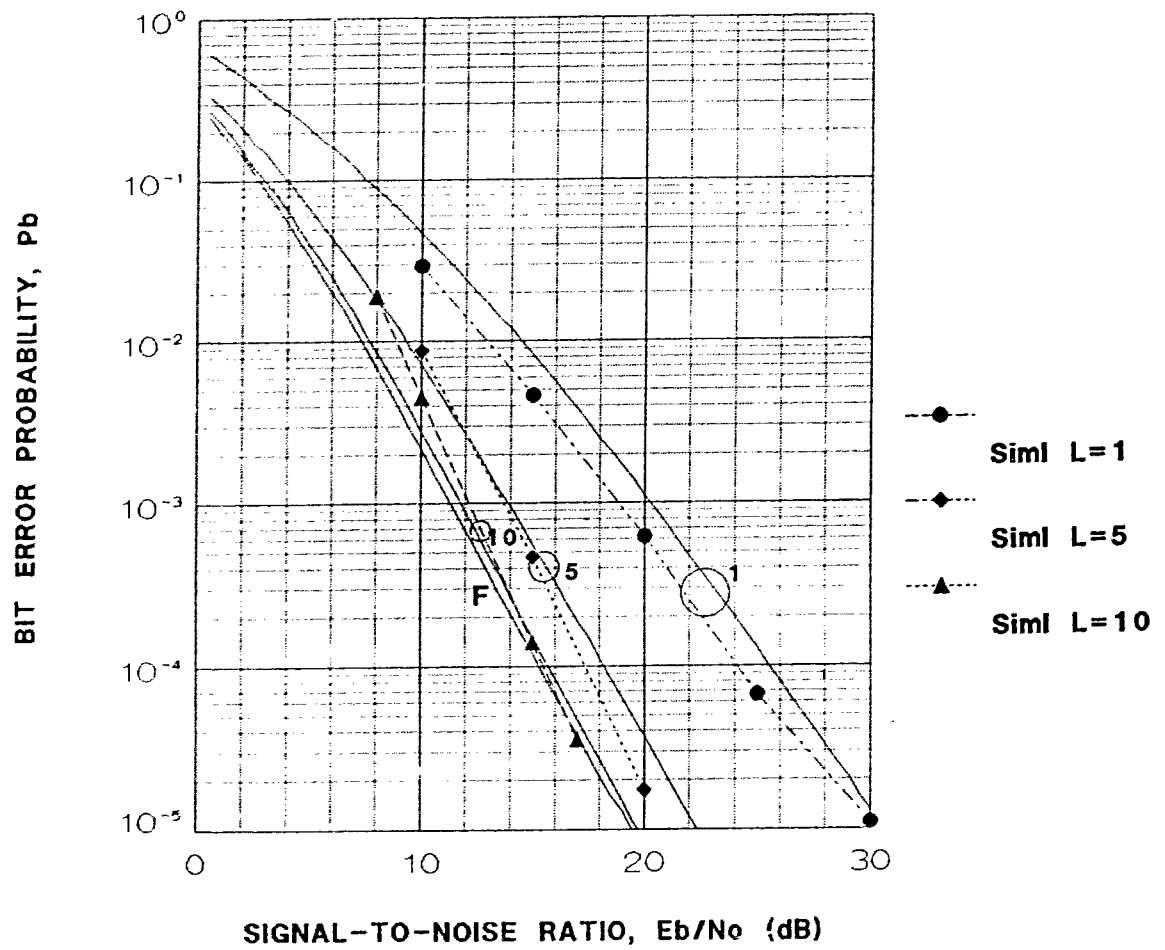
Figure 4.9: Bit error performance of Ungerboeck's 8-state code in a Rayleigh fading channel with coherent detection and perfect CSI. $f_D T = 0.03$.

The simulation of the error performance was based on Monte Carlo error counting techniques and the number of samples for each point of the simulation curve was chosen to be larger than $100/P_b$. It is observed that the simulation results agree well with the analytical results at large SNR.

4.4 Analysis of interleaved, Trellis-coded DPSK

In Section 4.2, we see that for coded PSK with perfect CSI, the optimal decoding metric is simply a linear sum of branch metrics; see (4.12). It is not too difficult to show that the Viterbi algorithm can no longer be employed if the decoding metric does not have such an additive property. For this reason, we cannot adopt the optimum quadratic decoder from Section 2.2 for coded DPSK. We use instead the following metric [9]:

$$M(\hat{c}) = -\sum_k \{y_k x_k^* \hat{c}_k^* + y_k^* x_k \hat{c}_k\} \qquad (4.44)$$

where y_k and x_k is defined in (2.10) and (4.10) respectively. Given that the transmitted codeword is $c = (c_1, \ldots, c_k, \ldots, c_N)$, error will occur if the random variable

$$
\begin{aligned}
D &= M(\hat{c}) - M(c) \\
&= \sum_{k=1}^{N} \{y_k x_k^* (c_k^* - \hat{c}_k^*) + y_k^* x_k (c_k - \hat{c}_k)\} \qquad (4.45)
\end{aligned}
$$

is less than 0. As in the case of coded PSK with perfect CSI, we let $(k_1, k_2, \ldots, k_L)$ denotes the set of index k for which $c_k \neq \hat{c}_k$, where $k_i > k_j$ if $i > j$. In addition, we let each k_i be written as $k_i = b_i \alpha + a_i + 1$, where $0 \leq a_i \leq \alpha - 1$ and $0 \leq b_i \leq \beta - 1$.

After substituting these information as well as (2.10), (4.4), (4.5) and (4.10) into (4.45), we have, in the case of coded DPSK, the following expression for the random variable D:

$$D = \sum_{i=1}^{L} \{ y_{k_i} x_{k_i}^* d_i^* + y_{k_i}^* x_{k_i} d_i \} \qquad (4.46)$$

where

$$y_{k_i} = \frac{A}{\sqrt{N_o}} g_{a_i\beta+b_i+1} s_{a_i\beta+b_i} c_{k_i} + n'_{a_i\beta+b_i+1} \qquad (4.47)$$

$$x_{k_i} = \frac{A}{\sqrt{N_o}} g_{a_i\beta+b_i} s_{a_i\beta+b_i} + n'_{a_i\beta+b_i}, \qquad (4.48)$$

and the d_i's are defined in (4.18). If the error event spans no more than α symbols, i.e., if $k_L - k_1 < \alpha$, then the set of $2L$ noise terms

$$\{ n'_{a_i\beta+b_i+1}, n'_{a_i\beta+b_i} \}$$

are iid complex Gaussian random variables, each having a zero mean and a unit variance. In otherwords, as long as the span of an error event is shorter than the number of rows in the interleaver buffer, the set of $2L$ noise terms will not contain duplicated items.

Let $\tilde{\mathbf{x}} = (x_{k_1}, \ldots, x_{k_i}, \ldots, x_{k_L})^t$, $\tilde{\mathbf{y}} = (y_{k_1}, \ldots, y_{k_i}, \ldots, y_{k_L})^t$ and $\tilde{\mathbf{r}} = (\tilde{\mathbf{x}}^t, \tilde{\mathbf{y}}^t)^t$. Then, the random variable D can be written as

$$D = \tilde{\mathbf{r}}^\dagger \tilde{\mathbf{F}} \tilde{\mathbf{r}}, \qquad (4.49)$$

where $\tilde{\mathbf{F}}$ is defined in (4.21). To calculate the error probability, we need to determine the matrix $\tilde{\boldsymbol{\Phi}}_{\mathbf{rr}} = \frac{1}{2}E\{\tilde{\mathbf{r}}\tilde{\mathbf{r}}^\dagger\}$. As shown in Chapter 2, this matrix is given in terms of the submatrices $\tilde{\boldsymbol{\Phi}}_{\mathbf{uu}}$ and $\tilde{\boldsymbol{\Phi}}_{\mathbf{xu}}$. If we assume that the error event is confined to one single column with respect to the interleaver buffer in Figure 1.2, then all the b_i's are the same and hence from (4.47), we obtain

$$\tilde{\boldsymbol{\Phi}}_{\mathbf{uu}} = \tilde{\mathbf{S}}(\tilde{\boldsymbol{\Phi}}_{22} + \mathbf{I})\tilde{\mathbf{S}}^\dagger \qquad (4.50)$$

where

$$\tilde{\mathbf{S}} = \begin{pmatrix} S_{a_1\beta+b} & & \\ & \ddots & \\ & & S_{a_L\beta+b} \end{pmatrix}, \qquad (4.51)$$

and $\tilde{\boldsymbol{\Phi}}_{22}$ is same as the $\tilde{\boldsymbol{\Phi}}_{\mathbf{uu}}$ in (4.23) where $\tilde{u}$ is defined on page 83. Similarly, we can show that

$$\tilde{\boldsymbol{\Phi}}_{\mathbf{xx}} = \tilde{\mathbf{S}}(\tilde{\boldsymbol{\Phi}}_{22} + \mathbf{I})\tilde{\mathbf{S}}^\dagger \qquad (4.52)$$

and

$$\tilde{\boldsymbol{\Phi}}_{\mathbf{ux}} = \tilde{\mathbf{S}}\boldsymbol{\Phi}_1\tilde{\mathbf{S}}^\dagger, \qquad (4.53)$$

where $\boldsymbol{\Phi}_1$ is a matrix whose $(i,j)^{th}$ element, denoted by $\phi_1(i,j)$, is equal to

$$\phi_1(i,j) = \rho(\beta(k_i - k_j) + 1). \qquad (4.54)$$

99

After substitution these submatrices into (2.20), we can rewrite $\tilde{\Phi}_{rr}$ as

$$\tilde{\Phi}_{rr} = \mathbf{LKL}^\dagger \qquad (4.55)$$

where

$$\mathbf{K} = \begin{pmatrix} \tilde{\Phi}_{22} + \mathbf{I} & \Phi_1^\dagger \\ \Phi_1 & \tilde{\Phi}_{22} + \mathbf{I} \end{pmatrix}, \qquad (4.56)$$

and

$$\mathbf{L} = \begin{pmatrix} \tilde{\mathbf{S}} & \mathbf{0} \\ \mathbf{0} & \check{\mathbf{C}}\tilde{\mathbf{S}} \end{pmatrix}. \qquad (4.57)$$

It should be pointed out that $\mathbf{LL}^\dagger = \mathbf{I}$. Although $\mathbf{L}$ is data dependent, we can remove $\tilde{\mathbf{S}}$ from the matrix $\mathbf{L}$ without affecting the final results; see Appendix G. Hence, (4.55) is equivalent to:

$$\tilde{\Phi}_{rr} = \mathbf{EKE}^\dagger \qquad (4.58)$$

where

$$\mathbf{E} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \check{\mathbf{C}} \end{pmatrix}. \qquad (4.59)$$

We are now in the position to find the pairwise error probability. As showed in (4.49), the random variable D is in quadratic form and the matrices $\tilde{\mathbf{F}}$ and $\tilde{\Phi}_{rr}$ are defined in (4.21) and (4.58) respectively. We can follow step 3 and 4 on page 31 to

100

obtain the pairwise error event probability of interleaved, coded DPSK signals. The key is to determine the eigenvalues of either the matrix $\tilde{\mathbf{B}}$ in (2.33) or the eigenvalues of $\tilde{\mathbf{\Phi}}_{rr}\tilde{\mathbf{F}}$ in (2.37).

4.4.1 Full Interleaving

The matrix $\mathbf{K}$ in (4.56) is a Hermitian matrix which can be written alternatively as

$$\mathbf{K} = \mathbf{U_K \Lambda_K U_K^\dagger} \tag{4.60}$$

where $\mathbf{\Lambda_K}$ is a diagonal matrix whose elements are the eigenvalues of the matrix $\mathbf{K}$ and $\mathbf{U_K}$ is a unitary matrix whose columns are eigenvectors of the matrix $\mathbf{K}$. As a result, $\tilde{\mathbf{\Phi}}_{rr}$ in (4.58) can be expressed as

$$\tilde{\mathbf{\Phi}}_{rr} = (\mathbf{E U_K \Lambda_K^{\frac{1}{2}}})(\mathbf{\Lambda_K^{\frac{1}{2}} U_K^\dagger E^\dagger}). \tag{4.61}$$

Consider the case of full interleaving, i.e., when β is a very large number. In this situation, the matrix $\mathbf{K}$ in (4.56) becomes

$$\mathbf{K} = \begin{pmatrix} (\rho(0) + 1)\mathbf{I} & \rho(1)\mathbf{I} \\ \rho(1)\mathbf{I} & (\rho(0) + 1)\mathbf{I} \end{pmatrix}, \tag{4.62}$$

where $\rho(0) = E_s/N_o$ and $\rho(1) = (E_s/N_o)J_o(2\pi f_D T)$; see (4.7). The eigenvalues of this matrix are

101

$$\lambda_{K,i} = \begin{cases} \rho(0) + 1 + \rho(1) & 1 \leq i \leq L \\ \rho(0) + 1 - \rho(1) & L + 1 \leq i \leq 2L \end{cases} \qquad (4.63)$$

This implies the matrix Λ_K can be written as

$$\Lambda_K = \begin{pmatrix} \lambda_1 I & 0 \\ 0 & \lambda_2 I \end{pmatrix}, \qquad (4.64)$$

where

$$\lambda_j = \begin{cases} \rho(0) + 1 + \rho(1) & j = 1 \\ \rho(0) + 1 - \rho(1) & j = 2 \end{cases} \qquad (4.65)$$

It can be shown that the corresponding unitary matrix U_K is

$$U_K = \frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ I & -I \end{pmatrix}. \qquad (4.66)$$

Substituting (4.64) and (4.66) into $\tilde{B} = \Lambda_K^{\frac{1}{2}} U_K^{\dagger} E^{\dagger} \tilde{F} E U_K \Lambda_K^{\frac{1}{2}}$ in (2.33), and after some algebra, we can show that the matrix $\tilde{B}$ can be written as

$$\tilde{B} = \frac{1}{2} \begin{pmatrix} \lambda_1(\Theta + \Theta^{\dagger}) & \sqrt{\lambda_1 \lambda_2}(\Theta - \Theta^{\dagger}) \\ -\sqrt{\lambda_1 \lambda_2}(\Theta - \Theta^{\dagger}) & -\lambda_1(\Theta + \Theta^{\dagger}) \end{pmatrix}, \qquad (4.67)$$

where

$$\Theta = \tilde{\Delta}^{\dagger} \tilde{C} = \begin{pmatrix} c_{k_1} d_1^* & & \\ & \ddots & \\ & & c_{k_L} d_L^* \end{pmatrix} \qquad (4.68)$$

and $d_i = c_{k_i} - \hat{c}_{k_i}$. The eigenvalues of the matrix $\tilde{B}$ are

$$\begin{bmatrix} \lambda_{g,i} \\ \lambda_{g,i+L} \end{bmatrix}_{i=1}^{L} = \frac{|d_i|^2 \rho(1)}{2} \left(1 \pm \sqrt{1 + \frac{4[\rho(0) + 1 + \rho(1)][\rho(0) + 1 - \rho(1)]}{|d_i|^2 \rho^2(1)}} \right). \quad (4.69)$$

This implies the poles of the characteristic function $\Phi_D(s)$ are

$$\begin{bmatrix} p_i \\ p_{i+L} \end{bmatrix}_{i=1}^{L} = \frac{1}{2} \frac{\rho(1)}{(\rho(0) + 1)^2 - \rho^2(1)} \left(\frac{1}{2} \mp \sqrt{\frac{1}{4} + \frac{\rho(0)(1 - |\mu|^2) + 1}{\rho(0)|d_i|^2|\mu|^2}} \right), \quad (4.70)$$

where

$$|\mu|^2 = \frac{J_o^2(2\pi f_D T)}{1 + \left(\frac{E_s}{N_o}\right)^{-1}} \quad (4.71)$$

is the squared magnitude of the correlation coefficient of y_{k_i} and x_{k_i} in (4.47) and (4.48). Note that the poles given in (4.70) are identical to those found in [5] except for the scaling factor of $\frac{1}{2}\rho(0)\{(\rho(0) + 1)^2 - \rho^2(1)\}^{-1}$. However, it can be shown easily that such a scaling factor has no effect on the pairwise error event probability calculation.

4.4.2 The Error Performance

We report in this section the error performance of Ungerboeck's 8-state Trellis-coded 8-DPSK scheme in a Rayleigh fading channel. Similar to the case of perfect CSI, we obtain an approximation to the bit error probability by taking the sum of

103

the probabilities of the dominant error events listed in Table 4.1. As discussed in Section 4.3.4, in order to make use of the set of events listed in Table 4.1 in the bit error probability calculation, we have to make sure that the eigenvalues of the matrix $\tilde{\Phi}_{rr}\tilde{F}$, and consequently the error probability, depend only on the squared Euclidean distance between symbol pairs in the transmitted and the erroneous words. It is easy to see that the eigenvalues depend only on the phase differences between different pairs of c_{k_i} and $\hat{c}_{k_i}$. However, we are not able to demonstrate analytically if the signs of the phase differences affect the eigenvalues. We have found numerically, though, the signs of the phase differences has no effect on the probabilities of the error events listed in Table 4.1.

We show in Figure 4.10 the approximate bit error probability of Ungerboeck's 8 state code with differential detection in a Rayleigh fading channel. The normalized fade rate is only 0.003, which corresponds to rather slow fading. It is observed that as the interleaving depth increases, the bit error probability decreases. When the interleaving depth is equal to 80 symbols, or equivalently, one-quarter the duration of a fade-cycle, the error performance approaches that provided by full interleaving. When the normalized fade rate is equal to 0.03, we have the bit error probability shown in Figure 4.11. Because of the relative fast fading, the CSI provided by the previously received sample has a lower correlation with the fading experienced by the symbol transmitted in the current interval. As a result, we have the *irreducible* error floors shown in Figure 4.11. Moreover, it is observed that full interleaving does not provide the best error performance. As a matter of fact, it appears that an interleaving depth of 10 symbols is close to the optimal choice at this fade rate. The *irreducible* error floor at this interleaving depth is about 3 times lower than that of full interleaving. When the interleaving depth is equal to 20 symbols, the

Figure 4.10: Bit error probability of Ungerboeck's 8-state 8-PSK code in a Rayleigh fading channel with differential detection. The normalized Doppler frequency is **0.003.**
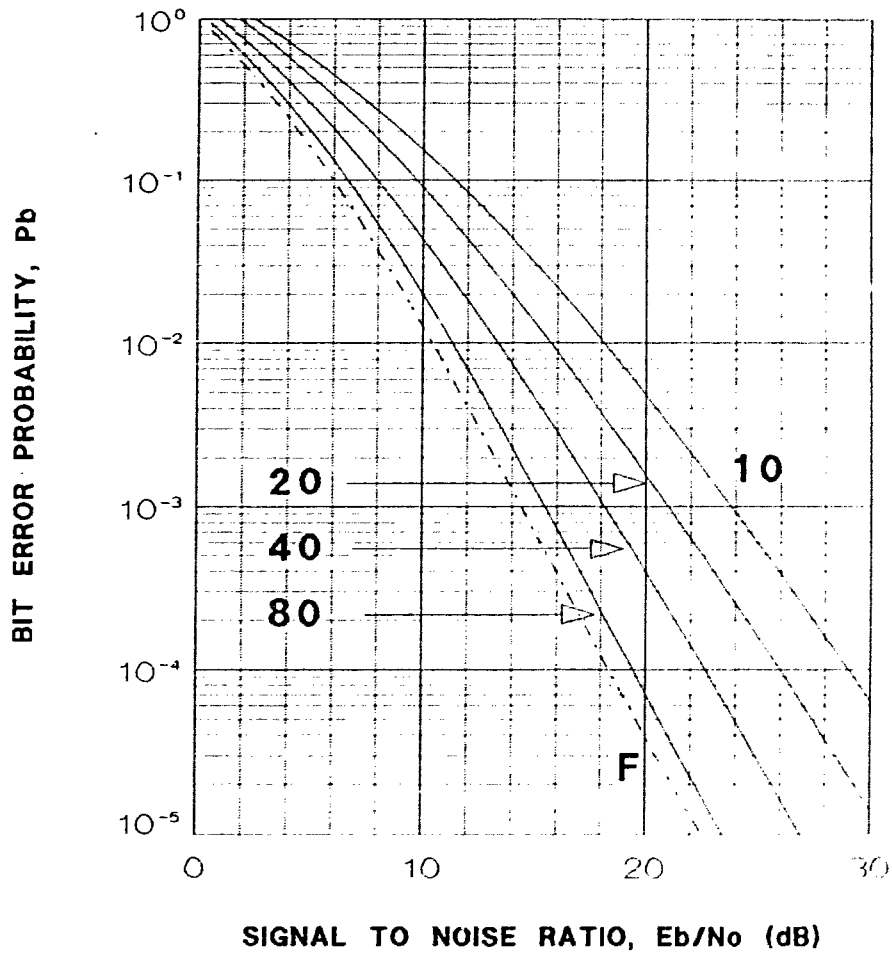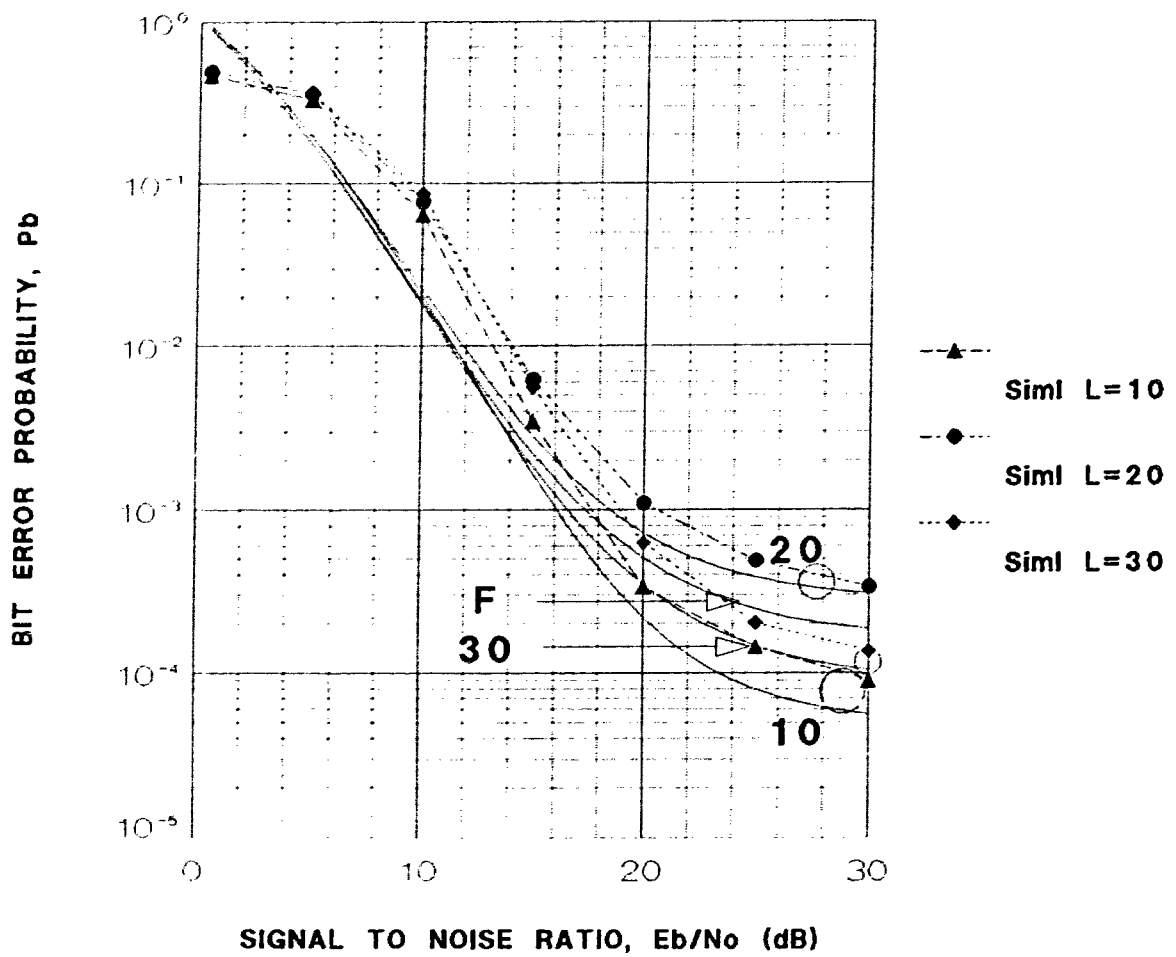
Figure 4.11: Bit error probability of Ungerboeck's 8-state 8-PSK code in a Rayleigh fading channel with differential detection. The normalized Doppler frequency is **0.03**.

bit error performance is worse than that of full interleaving. However, if the interleaving depth is increased to 3^n, the performance becomes once again better than full interleaving. This phenomenon is further confirmed by computer simulation; see dash curves on Figure 4.11. It appears that the bit error probability, as a function of the interleaving depth, oscillates around the bit error probability provided by full interleaving. This is probably caused by the oscillation in the Bessel function representing the autocorrelation function of the channel fading process. When the autocorrelation function of the channel fading process is changed to an exponential function, and specifically, if $\rho(k)$ in (4.7) is replaced by $\frac{E_s}{N_o} exp(-|k f_D T|)$, we have the error probability shown in Figure 4.12. The fade rate in Figure 4.12 is again equal to 0.03. It appears that if the autocorrelation function decrease monotonically to zero, increasing the interleaving depth always improves the error performance (with diminishing returns).

4.5 Summary

In this chapter, we have reviewed the concept of TCM. Specifically, we have shown the structure of Ungerboeck's 8-state and 16-state code. The exact pairwise error event probability expressions for interleaved, coded PSK with perfect CSI and interleaved, coded DPSK have been obtained via the 4 steps on page 31. For full interleaving, those pairwise error event expressions have been proven analytically the same as those report in [5]. Numerical calculations of the approximate bit error probability at various interleaving depths and fade rates are presented. It is found that generally an interleaving depth roughly equivalent to one-fifth or one-quarter the duration of a fade cycle is sufficient to produce the same effect as full interleav-
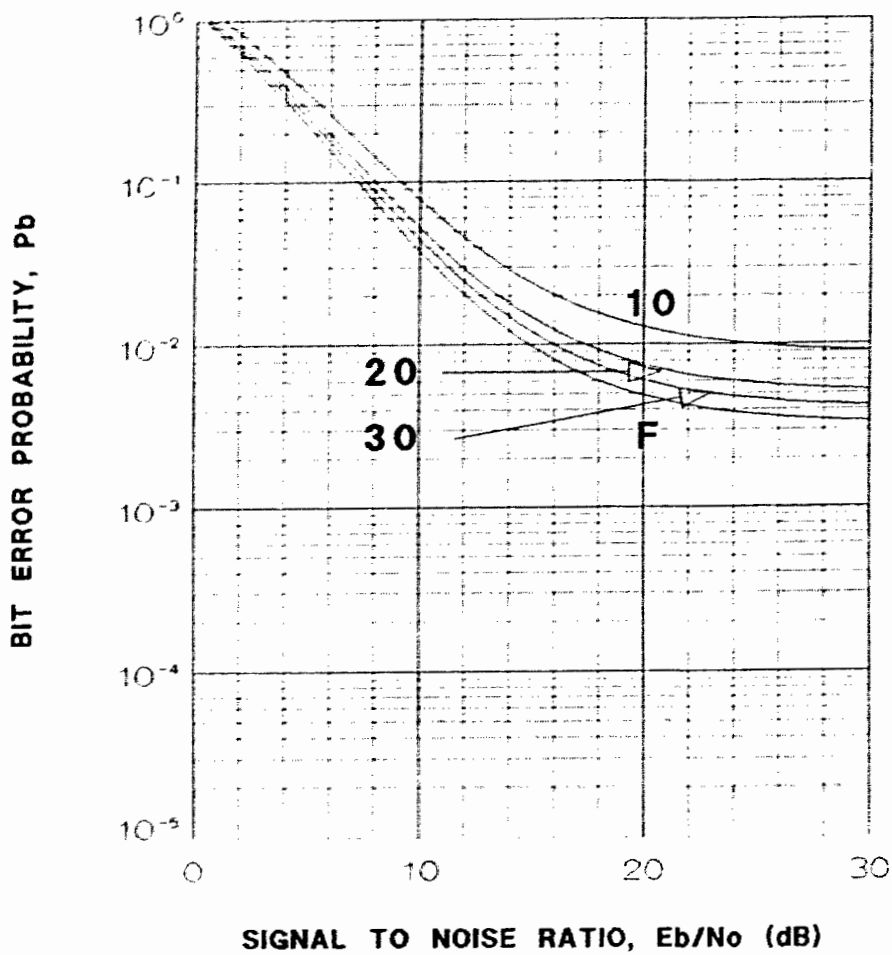
Figure 4.12: Bit error probability of Ungerboeck's 8-state 8-PSK code in a Rayleigh fading channel with differential detection. The normalized Doppler frequency is **0.03** and the autocorrelation function of the fading process is an exponential function.

ing. However, DPSK with a fade rate of 0.03 or higher, increasing the interleaving depth does not necessarily enhance the error performance. As a confirmation of the analytical error performance, Monte-Carlo simulation results have been provided for both coherent PSK and DPSK with a 0.03 fade rate.

CHAPTER 5

CONCLUSIONS

5.1 Conclusions

In this thesis, we have presented a general technique to find the pairwise error event probability of PSK signals transmitted over correlated Rayleigh fading channels. This technique is suitable for analyzing a wide range of modulation schemes including multiple-symbol differential detection, pilot symbol assisted modulation (PSAM), Trellis coded PSK with perfect CSI, and Trellis coded DPSK.

As the first example to illustrate the general technique, we analyze the error performance of multiple-symbol differential detection of PSK signals transmitted over Rayleigh fading channels. It is found that this detection strategy is very effective in eliminating the irreducible error probability associated with a conventional differential detector. For a 4-PSK system and a channel with a maximum Doppler frequency equal to 3% of the signaling rate, a detector with a word length of $N = 2$ symbols

can practically eliminate the irreducible error floor associated with a conventional detector ($N = 1$) at the bit error rate above 10^{-5}. When the length is increased to 5, the degradation in energy efficiency relative to perfect coherent detection is only 4 dB at an error rate of 10^{-4}. It is true that the larger the word length N, the better is the error performance. In the limiting case when N is extremely large, a multiple-symbol differential detector will probably have the error performance close to that of the perfect coherent detector. In reality though, the performance of a multiple-symbol differential detector will be limited by the decoding complexity. With exhaustive search, the decoding complexity grows exponentially with N.

Similar to multiple-symbol differential detection, PSAM can also remove the irreducible error floor at the bit error rate above 10^{-5}. The performance of uncoded BPSK with PSAM has been studied in this thesis. It is observed that in fast fading (such as $f_D T = 0.05$), block decoding strategy achieves slightly better results than those reported in [6]. Besides, if the size of the data block is reasonably large, the difference between using 4 pilot symbols and 20 pilot symbols is almost the same. Hence, a smaller decoding delay is resulted.

Using the general analytical technique developed in Chapter 2, we study the error performance of interleaved Trellis coded PSK modulations transmitted over correlated Rayleigh fading channels. Both coded PSK with perfect CSI as well as coded DPSK were considered. In the case of full interleaving, our results coincide with those found in [5]. For coded PSK with perfect CSI, we found that increasing the interleaving depth always provides an improvement in error performance (of course with diminishing return). On the other hand, we found that for coded DPSK, increasing the interleaving depth does not necessarily improve the error performance.

111

This is especially true when the Doppler frequency is around 3% of the signaling rate. At this fade rate, the optimum interleaving depth is 10 symbols. We suspect the existence of an optimum interleaving depth is probably due to the sinusoidal nature of the autocorrelation function of the channel's fading process. When an exponential autocorrelation function is used instead, we found once again that full interleaving provides the best performance. Finally it should be pointed out that at relatively slow fade rate (such as $f_D T = 0.003$), coded DPSK behaves similar to coded PSK with perfect CSI, i.e., increasing the interleaving depth does improve the error performance.

5.2 Suggestions for Further Research

Some suggestions for further work are as follows:

1. The analysis of PSAM could be extended to coded system as well as other forms of modulation.

2. With such encouraging results obtained in uncoded multiple-symbol differential detection system, we believe that these detectors could produce higher performance when used in conjunction with a soft decision channel coding system with interleaving to disperse the channel deep fades. However, suboptimal decoding algorithms with significant reduction in computational complexity are required if these detectors are to achieve an error performance close to that of perfect coherent detection.

Appendix A

THE GENERAL OPTIMUM

DECODING METRIC

In this appendix, we simplify the conditional probability density function in (2.26) and consequently, a very simple form of the optimum decoding metric is obtained.

Consider (2.20), the determinant of the covariance matrix Φ_{rr} is related to the determinant of Φ_{xx} as

$$\|\Phi_{rr}\| = \|\Phi_{xx}\| \cdot \|\Phi_{yy} - \Phi_{yx}\Phi_{xx}^{-1}\Phi_{xy}\|. \tag{A.1}$$

By making use of the partitioned multiplication in [3], the inverse of the matrix Φ_{rr} can be obtained as

$$\Phi_{rr}^{-1} = \begin{pmatrix} I & -\Phi_{xx}^{-1}\Phi_{xy} \\ 0 & I \end{pmatrix} \begin{pmatrix} \Phi_{xx}^{-1} & 0 \\ 0 & (\Phi_{yy} - \Phi_{yx}\Phi_{xx}^{-1}\Phi_{xy})^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -\Phi_{yx}\Phi_{xx}^{-1} & I \end{pmatrix},$$

which implies

$$\mathbf{r}^{\dagger}\mathbf{\Phi}_{\mathbf{rr}}^{-1}\mathbf{r} = \left(\begin{array}{cc} \mathbf{x}^{\dagger} & \mathbf{y}^{\dagger} \end{array} \right) \mathbf{\Phi}_{\mathbf{rr}}^{-1} \left(\begin{array}{c} \mathbf{x} \\ \mathbf{y} \end{array} \right)$$

$$= \mathbf{x}^{\dagger}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{x} + \mathbf{y}^{\dagger}(\mathbf{\Phi}_{\mathbf{yy}} - \mathbf{\Phi}_{\mathbf{yx}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xy}})^{-1}\mathbf{y}$$

$$-\mathbf{x}^{\dagger}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xy}}(\mathbf{\Phi}_{\mathbf{yy}} - \mathbf{\Phi}_{\mathbf{yx}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xy}})^{-1}\mathbf{y}$$

$$-\mathbf{y}^{\dagger}(\mathbf{\Phi}_{\mathbf{yy}} - \mathbf{\Phi}_{\mathbf{yx}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xy}})^{-1}\mathbf{\Phi}_{\mathbf{yx}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{x}$$

$$+\mathbf{x}^{\dagger}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xy}}(\mathbf{\Phi}_{\mathbf{yy}} - \mathbf{\Phi}_{\mathbf{yx}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xy}})^{-1}\mathbf{\Phi}_{\mathbf{yx}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{x}. \qquad (A.2)$$

Furthermore, using (2.21), (2.22) and (2.23), we can show that conditioned on the data sequence $\hat{\mathbf{c}}$,

$$\mathbf{\Phi}_{\mathbf{yy}} - \mathbf{\Phi}_{\mathbf{yx}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xy}} = \hat{\mathbf{C}} \left[\mathbf{\Phi}_{\mathbf{uu}}' - \mathbf{\Phi}_{\mathbf{ux}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xu}} \right] \hat{\mathbf{C}}^{\dagger}, \qquad (A.3)$$

where $\mathbf{\Phi}_{\mathbf{uu}}'$ is defined in (2.22). Since each $\hat{c}_k$ has a unit magnitude, the determinant of the above matrix is independent of the sequence $\hat{\mathbf{c}}$ and it is equal to

$$||\mathbf{\Phi}_{\mathbf{yy}} - \mathbf{\Phi}_{\mathbf{yx}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xy}}|| = \frac{||\mathbf{\Phi}_{\mathbf{rr}}||}{||\mathbf{\Phi}_{\mathbf{xx}}||} = ||\mathbf{\Phi}_{\mathbf{uu}}' - \mathbf{\Phi}_{\mathbf{ux}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xu}}||. \qquad (A.4)$$

Subsequently, the constant term in front of the exponential function in (2.26) is independent of $\hat{\mathbf{c}}$ and therefore, can be ignored in the derivation of the optimal decoding metric. Now, if we let

$$\mathbf{a} = \left[\mathbf{\Phi}_{\mathbf{uu}}' - \mathbf{\Phi}_{\mathbf{ux}}\mathbf{\Phi}_{\mathbf{xx}}^{-1}\mathbf{\Phi}_{\mathbf{xu}} \right]^{-1},$$

the determinant of $\mathbf{\Phi}_{\mathbf{rr}}$ in (A.1) becomes

114

$$\|\Phi_{rr}\| = \frac{\|\Phi_{xx}\|}{\|a\|}. \tag{A.5}$$

Moreover, substituting (2.21), (2.23) and (A.3) into (A.2), the equation can be further reduced to

$$
\begin{aligned}
r^\dagger \Phi_{rr}^{-1} r &= x^\dagger \Phi_{xx}^{-1} x + y^\dagger \hat{C} a \hat{C}^\dagger y \\
&\quad - x^\dagger \Phi_{xx}^{-1} \Phi_{xu} a \hat{C} y \\
&\quad - y^\dagger \hat{C} a \Phi_{ux} \Phi_{xx}^{-1} x \\
&\quad + x^\dagger \Phi_{xx}^{-1} \Phi_{xu} a \Phi_{ux} \Phi_{xx}^{-1} x.
\end{aligned} \tag{A.6}
$$

Now, we can rearrange (A.6) so that it is once again in a matrix form

$$
r^\dagger \Phi_{rr}^{-1} r = x^\dagger \Phi_{xx}^{-1} x + \begin{pmatrix} x^\dagger & y^\dagger \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & \hat{C} \end{pmatrix} \begin{pmatrix} b^\dagger a b & -b^\dagger a \\ -ab & a \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & \hat{C}^\dagger \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{A.7}
$$

where

$$b = \Phi_{ux} \Phi_{xx}^{-1}$$

Finally, combining (2.26), (A.5), and (A.7) and then taking the natural log of (2.26), the optimum decoder is equivalent to the one which selects the sequence $\hat{c}$ such that the decoding metric

$$
M(\hat{C}) = \begin{pmatrix} x^\dagger & y^\dagger \end{pmatrix} \begin{pmatrix} 0 & -b^\dagger a \hat{C}^\dagger \\ -\hat{C} a b & \hat{C} a \hat{C}^\dagger \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
$$

is the smallest. Note that the term $\mathbf{b}^\dagger \mathbf{ab}$ is independent of $\hat{\mathbf{c}}$ and thus, can be removed from the decoding metric.

Appendix B

THE LINEAR

TRANSFORMATION FOR THE

VECTOR r IN CHAPTER 2

We present in this appendix the linear transformation to the random vector r in (2.12) so that the random variable D in (2.30) can be expressed in terms of independent Gaussian random variates. The particular transformation we use is taken from Appendix B in [27] and is reproduced here for completeness.

Consider the covariance matrix $\mathbf{\Phi_{rr}}$ in (2.20), since it is a Hermitian matrix, it can be written as

$$\mathbf{\Phi_{rr}} = \mathbf{U_r}\mathbf{\Lambda_r}\mathbf{U_r^{\dagger}} \qquad (\text{B.1})$$

where $\mathbf{U_r}$ is a unitary matrix [7] whose columns are orthonormal, and

$$\Lambda_r = \begin{pmatrix} \lambda_{r,1} & & & & \\ & \lambda_{r,2} & & & \\ & & \ddots & & \\ & & & \lambda_{r,k} & \\ & & & & \ddots \end{pmatrix} \qquad (B.2)$$

is a diagonal matrix whose elements, the $\lambda_{r,k}$'s, are the eigenvalues of the matrix $\mathbf{\Phi_{rr}}$. From the definition of the unitary matrix, it should be clear that $\mathbf{U_r^{-1} = U_r^\dagger}$. Since $\mathbf{\Phi_{rr}}$ is a positive semi-definite matrix, all its eigenvalues are real and non-negative. This implies the matrix $\mathbf{\Lambda_r}$ can also be decomposed into two matrices as:

$$\mathbf{\Lambda_r = \Lambda_r^{\frac{1}{2}} \Lambda_r^{\frac{1}{2}}} \qquad (B.3)$$

where

$$\mathbf{\Lambda_r^{\frac{1}{2}}} = \begin{pmatrix} \sqrt{\lambda_{r,1}} & & & & \\ & \sqrt{\lambda_{r,2}} & & & \\ & & \ddots & & \\ & & & \sqrt{\lambda_{r,k}} & \\ & & & & \ddots \end{pmatrix}. \qquad (B.4)$$

Now consider the transformation

$$\mathbf{h = \Lambda_r^{-\frac{1}{2}} U_r^\dagger r} \qquad (B.5)$$

where

118

$$
\Lambda_{\mathbf{r}}^{-\frac{1}{2}} = \begin{pmatrix} \frac{1}{\sqrt{\Lambda_{r,1}}} & & & & \\ & \frac{1}{\sqrt{\Lambda_{r,2}}} & & & \\ & & \ddots & & \\ & & & \frac{1}{\sqrt{\Lambda_{r,k}}} & \\ & & & & \ddots \end{pmatrix} \tag{B.6}
$$

is the inverse of $\Lambda_{\mathbf{r}}^{\frac{1}{2}}$. Then the covariance matrix $\Phi_{hh} = \frac{1}{2} E\{hh^\dagger\}$ becomes

$$
\begin{aligned}
\Phi_{hh} &= \Lambda_{\mathbf{r}}^{-\frac{1}{2}}(U_{\mathbf{r}}^\dagger \Phi_{\mathbf{rr}} U_{\mathbf{r}})\Lambda_{\mathbf{r}}^{-\frac{1}{2}} \\
&= \Lambda_{\mathbf{r}}^{-\frac{1}{2}} \Lambda_{\mathbf{r}} \Lambda_{\mathbf{r}}^{-\frac{1}{2}} \\
&= I
\end{aligned} \tag{B.7}
$$

where I is an identity matrix. With the above transformation of $\mathbf{r}$, the random variable $D = \mathbf{r}^\dagger F \mathbf{r}$ in (2.30) can be written as

$$
D = h^\dagger B h \tag{B.8}
$$

where

$$
B = \Lambda_{\mathbf{r}}^{\frac{1}{2}} U_{\mathbf{r}}^\dagger F U_{\mathbf{r}} \Lambda_{\mathbf{r}}^{\frac{1}{2}}. \tag{B.9}
$$

Since B is Hermitian matrix, it can be written as

$$
B = U_b \Lambda_b U_b^\dagger, \tag{B.10}
$$

where $\mathbf{U_b}$ is a unitary matrix whose columns are orthonormal, and

$$\Lambda_b = \begin{pmatrix} \lambda_{b,1} & & & & & \\ & \lambda_{b,2} & & & & \\ & & \ddots & & & \\ & & & \lambda_{b,k} & & \\ & & & & \ddots & \end{pmatrix} \tag{B.11}$$

is a diagonal matrix whose elements, the $\lambda_{b,k}$'s, are the eigenvalues of the matrix $\mathbf{B}$. Now, consider the transformation

$$\mathbf{h} = \mathbf{U_b q}. \tag{B.12}$$

The random variable D in (B.8) can be written as

$$\begin{aligned} D &= \mathbf{q}^\dagger \mathbf{U}_b^\dagger \mathbf{B} \mathbf{U_b q} \\ &= \mathbf{q}^\dagger \Lambda_b \mathbf{q} \\ &= \sum_k \lambda_{b,k} |q_k|^2. \end{aligned} \tag{B.13}$$

Since the covariance matrix

$$\begin{aligned} \Phi_{qq} &= \frac{1}{2} E\{\mathbf{q}\mathbf{q}^\dagger\} \\ &= \frac{1}{2} E\{\mathbf{U}_b^\dagger \mathbf{h}\mathbf{h}^\dagger \mathbf{U_b}\} \\ &= \mathbf{I}, \end{aligned} \tag{B.14}$$

the q_k's are iid complex Gaussian variables each having a zero mean and a unit variance. This completes the task of transforming D into a sum of independent Gaussian quadratic forms.

Appendix C

SHORT CUT TO FIND $\lambda_{b,k}$

We present in this appendix a faster way to find the eigenvalues of the matrix $\mathbf{B}$ in (B.9). Consider (2.30), $D = \mathbf{r}^{\dagger}\mathbf{Fr}$, with the transformation given in (B.5), $\mathbf{h} = \Lambda_{\mathbf{r}}^{-\frac{1}{2}}\mathbf{U}_{\mathbf{r}}^{\dagger}\mathbf{r}$, the random variable D can be expressed as

$$D = \mathbf{h}^{\dagger}\Lambda_{\mathbf{r}}^{\frac{1}{2}}\mathbf{U}_{\mathbf{r}}^{\dagger}\mathbf{F}\mathbf{U}_{\mathbf{r}}\Lambda_{\mathbf{r}}^{\frac{1}{2}}\mathbf{h}$$

or

$$D = \mathbf{h}^{\dagger}\mathbf{Bh}$$

where $\mathbf{B}$ was defined in (B.9). According to (B.5) and (B.7), the h_k's are a set of independent random variables. Furthermore, for any eigenvalue $\lambda_{b,k}$ of the matrix $\mathbf{B}$, there exists an eigenvector $\mathbf{m}_b$ which must satisfy the following relation,

$$\mathbf{Bm}_b = \lambda_{b,k}\mathbf{m}_b. \tag{C.1}$$

Now, we can substitute (B.9) into (C.1) and obtain the following equation:

$$\Lambda_r^{\frac{1}{2}} U_r^{\dagger} F U_r \Lambda_r^{\frac{1}{2}} m_b = \lambda_{b,k} m_b. \tag{C.2}$$

By multiplying both sides by $U_r \Lambda_r^{\frac{1}{2}}$, (C.2) becomes

$$U_r \Lambda_r^{\frac{1}{2}} \Lambda_r^{\frac{1}{2}} U_r^{\dagger} F U_r \Lambda_r^{\frac{1}{2}} m_b = \lambda_{b,k} U_r \Lambda_r^{\frac{1}{2}} m_b. \tag{C.3}$$

Using (B.1) and (B.3), we find that $U_r \Lambda_r^{\frac{1}{2}} \Lambda_r^{\frac{1}{2}} U_r^{\dagger}$ is in fact the covariance matrix Φ_{rr}. Consequently, (C.3) is simply

$$\Phi_{rr} F m_{rb} = \lambda_{b,k} m_{rb}, \tag{C.4}$$

where

$$m_{rb} = U_r \Lambda_r^{\frac{1}{2}} m_b. \tag{C.5}$$

This implies that the eigenvalues of the matrix B are also the eigenvalues of the matrix

$$G' = \Phi_{rr} F.$$

Appendix D

THE MATRIX ã FOR MULTIPLE-SYMBOL DIFFERENTIAL DETECTION IN STATIC FADING CHANNELS

We derive in this appendix the matrix $\tilde{a}$ in (3.21) for multiple symbol differential detection in static fading channels. Starting with the matrices Φ_{12}, Φ_{21} and Φ_{22} in (3.16) and (3.10), when $f_D T = 0$, we can rewrite them as:

$$\Phi_{12} = \rho \mathbf{I}^t \tag{D.1}$$

$$\Phi_{21} = \rho \mathbf{l} \tag{D.2}$$

$$\Phi_{22} + \mathbf{I} = \rho \mathbf{J} + \mathbf{I} \tag{D.3}$$

where $\mathbf{l}$ and $\mathbf{J}$ are $N \times 1$ and $N \times N$ matrices whose entries are all equal to unity, $\mathbf{I}$ is an identity matrix and ρ is $\phi(0,0)$ defined in (3.11). Then, substituting (D.1), (D.2), (D.3) and (3.19) into (3.21), we have

$$
\tilde{\mathbf{a}} = [\Phi_{22} + \mathbf{I} - \Phi_{21}\Phi_{\mathbf{xx}}^{-1}\Phi_{12}]^{-1} = \left[\rho\mathbf{J} + \mathbf{I} - \frac{\rho^2}{\rho+1}\mathbf{J}\right]^{-1}
$$

$$
= \left[\mathbf{I} + \left(\frac{\rho}{\rho+1}\right)\mathbf{J}\right]^{-1}. \tag{D.4}
$$

For any non-singular matrix $\mathbf{A}$, there exists a relationship that $\mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$. Hence, the matrix $\tilde{\mathbf{a}}$ can be found by

$$
\mathbf{I} = \tilde{\mathbf{a}}^{-1}\tilde{\mathbf{a}}
$$

$$
= \left[\mathbf{I} + \left(\frac{\rho}{\rho+1}\right)\mathbf{J}\right]\left[\mathbf{I} + \left(\frac{a\rho+b}{c\rho+d}\right)\mathbf{J}\right]
$$

$$
= \mathbf{I} + \frac{\rho}{\rho+1}\mathbf{J} + \frac{a\rho+b}{c\rho+d}\mathbf{J} + \frac{Na\rho^2 + Nb\rho}{(\rho+1)(c\rho+d)}\mathbf{J}
$$

$$
= \mathbf{I} + \frac{c\rho^2 + a\rho^2 + Na\rho^2 + d\rho + a\rho + b\rho + Nb\rho + b}{(\rho+1)(c\rho+d)}\mathbf{J} \tag{D.5}
$$

where a, b, c and d are arbitrary constants. Obviously, if the numerator of the coefficient of matrix $\mathbf{J}$ in (D.5) equals to zero, we can obtain the inverse matrix in (D.4). To process, we have to solve the following simultaneous equations:

$$Na + a + c = 0,$$

$$Nb + a + b + d = 0,$$

$$b = 0.$$

Since there are many solutions for the above equations, we choose a arbitrarily and use a to determine the values of c and d that satisfy the given equations. If $a = -1$, then $d = 1$ and $c = N + 1$. Note that the variable d is always equal to zero. As a result, $\tilde{a}$ in (3.21) becomes

$$
\begin{aligned}
\mathbf{a} &= \mathbf{I} - \frac{\rho}{(N+1)\rho + 1}\mathbf{J} \\
&= \mathbf{I} - f\mathbf{J}
\end{aligned}
\tag{D.6}
$$

where $f = \frac{\rho}{(1+(N+1)\rho)}$ and ρ is $\phi(0,0)$ in (3.11).

Appendix E

MESSAGE SEQUENCE IN MULTIPLE-SYMBOL DIFFERENTIAL DETECTION

We want to show in this appendix that the pairwise error probability is independent of the message sequence. To proceed, we note that by using the **Cholesky decomposition** method, the matrix $\Phi_{\mathbf{rr}}$ in multiple symbol differential detection can be written as

$$\Phi_{\mathbf{rr}} = \mathbf{ZMM}^\dagger \mathbf{Z}^\dagger \qquad (\text{E.1})$$

where

$$\mathbf{Z} \;=\; \begin{pmatrix} 1 & & & \\ & z_1 & & \\ & & \ddots & \\ & & & z_N \end{pmatrix}, \tag{E.2}$$

$$\mathbf{MM}^\dagger \;=\; \begin{pmatrix} \phi(0,0)+1 & \boldsymbol{\Phi}_{12} \\ \boldsymbol{\Phi}_{21} & \boldsymbol{\Phi}_{22}+\mathbf{I} \end{pmatrix}, \tag{E.3}$$

and $\phi(0,0)+1$, $\boldsymbol{\Phi}_{12}$, $\boldsymbol{\Phi}_{22}$ are defined in Section 3.1.1. As showed in Appendix C, the eigenvalues of the matrix $\mathbf{G}' = \boldsymbol{\Phi}_{\mathbf{rr}}\mathbf{F}$ are identical to the eigenvalues of the matrix

$$\mathbf{B} \;=\; \mathbf{M}^\dagger(\mathbf{Z}^\dagger \mathbf{F}\mathbf{Z})\mathbf{M} \tag{E.4}$$

$$=\; \mathbf{M}^\dagger \begin{pmatrix} 0 & \tilde{\mathbf{b}}^\dagger \tilde{\mathbf{a}}(\mathbf{I}-\boldsymbol{\Theta}^\dagger) \\ (\mathbf{I}-\boldsymbol{\Theta})\tilde{\mathbf{a}}\tilde{\mathbf{b}} & \boldsymbol{\Theta}\tilde{\mathbf{a}}\boldsymbol{\Theta}^\dagger - \tilde{\mathbf{a}} \end{pmatrix} \mathbf{M}, \tag{E.5}$$

where $\mathbf{F}$ was defined in (3.26), and

$$\boldsymbol{\Theta} = \mathbf{z}^\dagger \hat{\mathbf{z}} \tag{E.6}$$

is a diagonal matrix whose k^{th} element is equal to $\hat{z}_k z_k^* = \prod_{j=1}^{k} \hat{c}_j c_j^*$. It should be clear from these results that as long as the set of phase differences between the symbols c_k and $\hat{c}_k, k = 1, \ldots, N$, are fixed, the eigenvalues of the matrix $\mathbf{G}'$ are independent of the sequence $\mathbf{c} = (c_1, \ldots, c_N)$.

Appendix F

MESSAGE SEQUENCE IN COHERENT PSK

We want to show in this appendix that the eigenvalues for the matrix Φ_{hh} in (4.30) depends only on the magnitudes of the d_i's and not their phases. To proceed, let

$$\Phi'_{hh} = T\Phi_{hh}T, \tag{F.1}$$

where T is a permutation matrix. The matrix T is chosen in such a way that the ith row and the ith column of Φ_{hh} now becomes the first row and the first column of Φ'_{hh} and vice versa. The other rows and columns of Φ_{hh} are left unchanged. The matrix Φ'_{hh} has the same eigenvalues as Φ_{hh}. In addition, it can be written in the form:

$$\Phi'_{hh} = \begin{pmatrix} \rho(0)|d_i|^2 & d_i M_{12} \\ d_i^* M_{12}^\dagger & M_{22} \end{pmatrix}. \qquad \text{(F.2)}$$

where $d_i M_{12}$ and M_{22} are submatrices of Φ'_{hh}, and M_{12} and M_{22} are independent of d_i. The sizes of these two submatrices are $1 \times L$ and $(L-1) \times (L-1)$ respectively. The eigenvalues of Φ_{hh} are the roots of the determinant of $\Phi'_{hh} - \lambda I$. It can be shown that

$$\|\Phi'_{hh} - \lambda I\| = (\rho(0)|d_i|^2 - \lambda) \cdot \left\| (M_{22} - \lambda I) - \frac{|d_i|^2}{\rho(0)|d_i|^2 - \lambda} M_{12}^\dagger M_{12} \right\| \qquad \text{(F.3)}$$

The above equation tells us that the eigenvalues remain unchanged if we replace d_i by its conjugate. In otherwords, the eigenvalues of Φ_{hh} depend only on the magnitudes of the d_i's, or equivalently only on the squared Euclidean distances between symbol pairs in the transmitted and the erroneous words.

Appendix G

THE MATRIX $\tilde{\Phi}_{rr}$ FOR DPSK

We want to show in this appendix that the eigenvalues for the matrix $\tilde{\Phi}_{rr}\tilde{F}$ in coded DPSK do not depend on the matrix $\tilde{S}$. To proceed, we use the Cholesky decomposition method to rewrite the matrix $\tilde{\Phi}_{rr}$ in (4.55) as:

$$\tilde{\Phi}_{rr} = LMM^\dagger L^\dagger \tag{G.1}$$

where

$$MM^\dagger = K \tag{G.2}$$

and L is defined in (4.57). As shown in Appendix C, the eigenvalues of $\tilde{\Phi}_{rr}\tilde{F}$ is also the eigenvalues of the matrix

$$\bar{B} = M^\dagger L^\dagger \tilde{F} L M$$

$$
\begin{aligned}
&= \mathbf{M}^\dagger \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}}^\dagger \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{S}}^\dagger & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{S}}^\dagger \end{pmatrix} \begin{pmatrix} \mathbf{0} & \tilde{\boldsymbol{\Delta}}^\dagger \\ \tilde{\boldsymbol{\Delta}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{S}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{S}} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}} \end{pmatrix} \mathbf{M} \\
&= \mathbf{M}^\dagger \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}}^\dagger \end{pmatrix} \begin{pmatrix} \mathbf{0} & \tilde{\mathbf{S}}^\dagger \tilde{\boldsymbol{\Delta}}^\dagger \tilde{\mathbf{S}} \\ \tilde{\mathbf{S}}^\dagger \tilde{\boldsymbol{\Delta}} \tilde{\mathbf{S}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}} \end{pmatrix} \mathbf{M} \\
&= \mathbf{M}^\dagger \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}}^\dagger \end{pmatrix} \begin{pmatrix} \mathbf{0} & \tilde{\boldsymbol{\Delta}}^\dagger \\ \tilde{\boldsymbol{\Delta}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}} \end{pmatrix} \mathbf{M}.
\end{aligned} \tag{G.3}
$$

It is clear from (G.3) that $\tilde{\mathbf{B}}$ is independent of $\tilde{\mathbf{S}}$. This implies $\tilde{\mathbf{S}}$ can be removed the matrix $\mathbf{L}$ in (4.57).

REFERENCES

[1] Arredondi G. A., Chriss W. H., and Walker E. H., "A Multipath Fading Simulator for Mobile Radio", *IEEE Transactions on Communications*, Vol. COM-21, No. 11, pp. 1325-1328, November 1973.

[2] Arredondo G. A. and Smith J. I., "Voice and Data Transmission in a Mobile Radio Channel at 850 MHz", *IEEE Transactions on Vehicular Technology*, Vol. VT-26, No. 1, pp. 88-93, February 1977.

[3] Barnett S., "Matrix Methods for Engineers and Scientists", *London: McGraw-Hill*, 1979.

[4] Caples E. L., Massad K. E., and Minor T. R., " A UHF Channel Simulator for Digital Mobile Radio", *IEEE Transactions on Vehicular Technology*, Vol. VT-29, No. 2, pp. 281-289, May 1980.

[5] Cavers Jim and Ho Paul, "Analysis of the Error Performance of Trellis-Coded Modulations in Rayleigh Fading Channel", to appear in *IEEE Transactions on Communications*.

[6] Cavers Jim, "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels", to appear in *IEEE Transactions on Vehicular Technology*.

[7] Deif A. S., "Advanced Matrix Theory for Scientists and Engineers", *London: Abacus Press*, 1982.

[8] Divsalar D. and Simon M., "Trellis-Coded Modulation for 4800-9600 bits/s Transmission over a Fading Mobile Satellite Channel", *IEEE Journal on Selected Areas in Communications*, Vol. SAC-5, No. 2, pp. 162-175, February 1987.

133

[9] Divsalar D. and Simon M., "The Performance of Trellis Codes Multilevel DPSK on a Fading Mobile Satellite Channel", *IEEE Transactions on Vehicular Technology*, Vol. 37, No. 2, pp. 78–91, May 1988.

[10] Divsalar D. and Simon M., "The Design of Trellis Coded MPSK for Fading Channels: Performance Criteria", *IEEE Transactions on Communications*, Vol. 36, No. 9, pp. 1004–1012, September 1988.

[11] Divsalar D. and Simon M., "Multiple-Symbol Differential Detection of MPSK", *IEEE Transactions on Communications*, Vol. 38, No. 3, pp. 300–308, March 1990.

[12] Edbauer Franz, "Performance of Interleaved Trellis-Coded Differential 8-PSK Modulation over Fading Channels", *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 9, pp. 1340–1346, December 1989.

[13] Jakes W. C., "Microwave Mobile Communications", *New York: Wiley*, 1974.

[14] Kaplan Wilfred, "Advanced Mathematics for Engineers", *Massachusetts: Addison-Wesley Publishing Company*, 1981.

[15] Karim M. R., "Transmission of Digital Data over a Rayleigh Fading Channel", *IEEE Transactions on Vehicular Technology*, Vol. VT-31, No. 1, pp. 1–6, Feb 1982.

[16] Lee C. Y. William, Mobile Communications Engineering, *New York: McGraw Hill*, 1982.

[17] Lee C. M. Albert and McLane Peter J., " Convolutionally Interleaved PSK and DPSK Trellis Codes for Shadowed, Fast Fading Mobile Satellite Communication Channels", *IEEE Transactions on Vehicular Technology*, Vol. VT 39, No. 1, pp. 37-47, February 1990.

[18] Makrakis D. and Mathiopoulos P. T., "Optimal Decoding in Fading Channels: A Combined Envelope, Multiple Differential and Coherent Detection Approach", *Conference Proceedings, IEEE Globecom 89*, Dallas, November 1989, pp. 1551-1557.

[19] Makrakis D., Yongaçoğlu A., and Feher K., "A Sequential Decoder for Differential Detection of Trellis coded PSK Signals", *Conference Proceedings, ICC 88*, Philadelphia, June 1988, pp.1433-1438.

[20] McKay Ross, "**Error Bounds for Trellis Coded Modulation on Mobile Satellite Fading Channels**", M.Sc.(Eng.) Thesis, Dept. of Electrical Engineering, Queen's University, Kingston, Ontario, Canada.

[21] McLane P.J., Wittke P.H., Ho P, and Loo C., "**The Performance of Trellis Codes for Mobile Satellite Communications**", *IEEE Transactions on Communications*, Vol. 36, No. 9, pp. 1242-1246, November 1988.

[22] Miller M.J. and Ahamed S.V., "**Digital Transmissions Systems and Networks**", Volume 1: Principles, *Maryland: Computer Science Press*, 1987

[23] Nadia S. Adawi. *et al.*, "**Coverage Prediction for Mobile Radio Systems Operating in the 800/900 MHz Frequency Range**", *IEEE Trans. on Vehicular Technology*, VT-37, pp. 3-19, February 1988.

[24] Papoulis, "**Probability, Random Variables, and Stochastic Processes**", Second Edition, New York: McGraw-Hill, 1984.

[25] Proakis J. G., "**Digital Communications**", *New York: McGraw Hill*, 1983.

[26] Sampei S. and Sunaga T., "**Rayleigh Fading Compensation Method for 16QAM in Digital Land Mobile Radio Channels**", *Conference Proceedings, IEEE Vehicular Technology Conference*, pp. 640-646, San Francisco, May 1989.

[27] Schwartz M., Bennett W. and Stein. S, " **Communication Systems and Techniques**", *New York: McGraw Hill*, 1966.

[28] Simon M. K., "**Dual Pilot Tone Calibration Technique (DTCT)**", *IEEE Transactions on Vehicular Technology*, Vol. VT-35, pp. 63-70, May 1986.

[29] Sklar Bernard, "**Digital Communications Fundamentals and Applications**", *New Jersey: Prentice Hall*, 1988.

[30] Smith John, "**A Computer Generated Multipath Fading Simulator for Mobile Radio**", *IEEE Transactions on Vehicular Technology*, Vol. VT-24, No. 3, pp. 39-40 August 1975.

[31] Townsend A. A. R., Mobile Communications Engineering, *New York: Prentice Hall*, 1938.

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

[33] Ungerboeck G., "Trellis-Coded Modulation with Redundant Signal Sets Part I: Introduction", *IEEE Communications Magazine*, Vol. 25, No. 2, pp. 5-11, February 1987.

[34] Viterbi A. J. and Omura J. K., "Principles of Digital Communication and Coding", *New York: McGraw-Hill*, 1979.

[35] Walker John, Mobile Information Systems, *Boston: Artech House*, 1990.

[36] Wilson S. G., Freebersyser J., and Marshall C., "Multiple-Symbol Detection of M-DPSK", *Conference Proceedings, IEEE Globecom 89*, Dallas, November 1989, pp. 1692-1697.

[37] Wilson S. G. and Leung Y. S., "Trellis-coded Phase Modulation on Rayleigh Channels", *Conference Proceedings*, ICC '87, Seattle, WA., June 1987, pp. 21.3.1-21.3.5.

[38] Zehavi E. and Wolf J. K., "On the Performance Evaluation of Trellis Codes", *IEEE Transactions on Information Theory*, Vol. IT-33, No. 2, pp.196-202, March 1987.