THESIS REPORT

Ph.D.

Design and Performance of Trellis Codes for Wireless Channels

by S.A. Al-Semari
Advisor: T.E. Fuja

Ph.D. 95-10

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

Abstract

Title of Dissertation: DESIGN AND PERFORMANCE OF TRELLIS CODES FOR WIRELESS CHANNELS

Saud Ahmed Al-Semari, Doctor of Philosophy, 1995

Dissertation directed by: Associate Professor Thomas E. Fuja
Department of Electrical Engineering

Signal fading is one of the primary sources of performance degradation in mobile radio (wireless) systems. This dissertation addresses three different techniques to improve the performance of communication systems over fading channels, namely trellis coded modulation (TCM), space diversity and sequence maximum a posteriori decoding (MAP).

In the first part, TCM schemes that provide high coding gains over the flat, slowl Rayleigh distributed fading channel are presented. It is shown that the use of two encoders in parallel – used to specify the in-phase and quadrature components of the transmitted signal – results in large performance improvements in bit error rates when compared with conventional TCM schemes in which a single encoder is used. Using this approach – which we label "I-Q TCM" – codes with bandwidth efficiencies of 1,2, and 3 bits/sec/Hz are described for various constraint lengths. The performance of these codes is evaluated using tight upper bounds and simulation.

In the second part, the use of space diversity with three different combining schemes is investigated. Expressions for the cutoff rate parameter are shown

for the three combining schemes over the fully interleaved Rayleigh-distributed flat fading channel. Also, tight upper bounds on the pairwise error probability are derived for the three combining schemes. Examples of I-Q TCM schemes with diversity combining are shown. The cutoff rate and a tight upper bound on the pairwise error probability are also derived for maximal ratio combining with correlated branches.

In the last part the problem of reliably transmitting trellis coded signals over very noisy channels is considered. Sequence maximum a posteriori (MAP) decoding of correlated signals transmitted over very noisy AWGN and Rayleigh channels is presented. A variety of different systems with different sources, modulation schemes, encoder rates and complexities were simulated. Sequence MAP decoding proves to substantially improve the performance at very noisy channel conditions especially for systems with moderate redundancies and encoder rates. A practical example for coding the CELP line spectral parameters (LSP's) is also considered. Two source models are used. Coding gains as much as 4 dB are achieved.

DESIGN AND PERFORMANCE OF TRELLIS CODES FOR WIRELESS CHANNELS

by

Saud Ahmed Al-Semari

Dissertation submitted to the Faculty of the Graduate School
of The University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1995

Advisory Committee:

Associate Professor Thomas E. Fuja, Chairman/Advisor
Professor Nariman Farvardin
Professor Evaggelos Geraniotis
Associate Professor Steven Tretter
Professor Eric Slud

Dedication

To my parents and my wife

Acknowledgements

First of all, praises are due to God, whose help made it possible to accomplish this work.

I would like to express my deepest gratitude to my advisor, Professor Thomas Fuja, for his guidance, insights, suggestions, constructive criticisms, sustained encouragement and support. His profound thinking, depth of knowledge and nice personality made it a very pleasant experience. Gratitude is also expressed to my examining committee members, Dr. Farvardin, Dr. Geraniotis, Dr. Tretter and Dr. Slud. I would like also to thank Dr. Fady Alajaji of Queen's University for interesting discussions and help.

I am very grateful to all my friends and colleagues here at Maryland and Virginia for their assistance and help. The support from the EE departments at both King Fahd University of Petroleum and Minerals and the university of Maryland is gratefully acknowledged.

Finally, my most heartfelt gratitude goes to my parents and my wife. They have been instrumental in my pursuit of higher studies. They supported me with love, caring and pride. I can never thank them enough. This work is dedicated to them.

Table of Contents

Section	Page
List of Tables	**viii**
List of Figures	**xi**
1 Introduction	**1**
1.1 Typical System Model	1
1.2 Coded Modulation	4
1.3 Thesis Organization	8
2 Channel Models	**11**
2.1 The Additive Noise Channel	11
2.2 The Fading Channel	12
2.2.1 The Linear Time-Varying Model	12
2.2.2 Statistical Models	15
2.2.3 The Diversity Model	18
2.3 Correlation Models	19
2.3.1 Time Correlation	19
2.3.2 Space Correlation	21

2.4 The Rayleigh Channel Error Performance 23

3 I-Q Trellis Coded Modulation **29**

3.1 Introduction 29

3.1.1 Relevant Past Work 30

3.1.2 Proposed Codes 32

3.2 Background 34

3.2.1 Channel Model and Performance Criterion 35

3.3 The Structure of I-Q Trellis Coded Modulation 37

3.3.1 I-Q QPSK 39

3.3.2 I-Q 16-QAM 41

3.3.3 4-D I-Q 16-QAM 43

3.3.4 Comparisons 47

3.4 Performance over Correlated Rayleigh Fading 51

3.5 Other Considerations 55

3.5.1 Performance over Rician Channels 55

3.5.2 Effect of Channel State Information 57

3.6 Summary 60

4 Diversity Reception **61**

4.1 Introduction 61

4.2 System Model 63

4.3 Cutoff Rate for Diversity Reception 65

4.3.1 Maximal Ratio Combiner 66

4.3.2 Equal Gain Combining 68

4.3.3 Selection Combining 71

4.3.4 Comparisons and Numerical Results 73
4.4 Pairwise Error Probability 77
4.4.1 Maximal Ratio Combiner 78
4.4.2 Equal Gain Combiner 83
4.4.3 Selection Combining 87
4.5 Performance of Uncoded Systems 92
4.6 The Effect of Branch Correlation 95
4.6.1 Example 98
4.7 Comparisons and Complexity Tradeoffs 99
4.8 Summary 101

5 Sequence MAP Decoding **103**

5.1 Introduction 103
5.2 MAP decoding for Ideal Sources 106
5.2.1 System Model 106
5.2.2 MAP decoding rule 109
5.2.3 Bit Error Probability Upper Bound 110
5.3 Simulated Systems 111
5.3.1 BPSK Modulated Systems 112
5.3.2 QPSK Modulated Systems 113
5.3.3 8-PSK Modulated Systems 114
5.3.4 Observations and Discussions 114
5.3.5 The effect of Increasing the Signal Dimensionality on MAP Gains 127
5.4 CELP LSP's Coding 128
5.4.1 Coding of the LSP's 132

5.4.2 Simulation Results 133
5.4.3 Coding of CELP LSP's via I-Q QPSK 138
5.5 Summary 140

6 Summary and Conclusions **141**
6.1 Presentation Summary 141
6.2 Conclusions 142
6.3 Future Work 144

Bibliography **146**

List of Tables

Number		Page
3.1	Generator polynomials for the rate 1/2 codes.	40
3.2	A comparison between the proposed codes and conventional TCM schemes. The numbers inside parentheses correspond to Ungerboeck codes.	49
4.1	Improvement factor for MRC over EGC	70
4.2	The E_s/N_o per branch values (in dB) to achieve $R_o = 1$ for the QPSK constellation	74
4.3	The E_s/N_o per branch values (in dB) to achieve $R_o = 2$ for the 8-PSK constellation	75
4.4	The E_s/N_o per branch values (in dB) to achieve $R_o = 2$ for the 16-QAM constellation	75
4.5	The E_s/N_o per branch values (in dB) to achieve $R_o = 3$ for the 16-QAM constellation	75
5.1	The five source models and their redundancies	111
5.2	MAP decoding gains for different trellis codes over AWGN channels. Source I (with distribution $\Pr(0\|0) = .8, \Pr(1\|1) = .2$).	117

5.3 MAP decoding gains for different trellis codes over Rayleigh channels. Source I (with distribution $\Pr(0|0) = .8, \Pr(1|1) = .2$). 118
5.4 MAP decoding gains for different trellis codes over AWGN channels. Source II (with distribution $\Pr(0|0) = .8, \Pr(1|1) = .8$). . . . 119
5.5 MAP decoding gains for different trellis codes over Rayleigh channels. Source II (with distribution $\Pr(0|0) = .8, \Pr(1|1) = .8$). . . . 120
5.6 MAP decoding gains for different trellis codes over AWGN channels. Source III (with distribution $\Pr(0|0) = .88, \Pr(1|1) = .65$). . 121
5.7 MAP decoding gains for different trellis codes over Rayleigh channels. Source III (with distribution $\Pr(0|0) = .88, \Pr(1|1) = .65$). . 122
5.8 MAP decoding gains for different trellis codes over AWGN channels. Source IV (with distribution $\Pr(0|0) = .9, \Pr(1|1) = .8$). . . . 123
5.9 MAP decoding gains for different trellis codes over Rayleigh channels. Source IV (with distribution $\Pr(0|0) = .9, \Pr(1|1) = .8$). . . . 124
5.10 MAP decoding gains for different trellis codes over AWGN channels. Source V (with distribution $\Pr(0|0) = .2, \Pr(1|1) = .9$). . . . 125
5.11 MAP decoding gains for different trellis codes over Rayleigh channels. Source V (with distribution $\Pr(0|0) = .2, \Pr(1|1) = .9$). . . . 126
5.12 Sequence MAP decoding gains for the the CELP encoded speech with 4-state 4-D QPSK TCM schemes over both AWGN and Rayleigh channels 136
5.13 Sequence MAP decoding gains for the the CELP encoded speech with 8-state 4-D QPSK TCM schemes over both AWGN and Rayleigh channels 136

5.14 Sequence MAP 1 decoding gains for the the CELP encoded speech with 32-state I-Q QPSK TCM scheme over both AWGN and Rayleigh channels 138

List of Figures

Number		Page
1.1	A basic block diagram of a digital communication system.	3
1.2	A simple TCM scheme for encoding two or three bits per baud. .	5
2.1	The autocorrelation function	20
2.2	Cutoff rates for QPSK, 8-PSK and 16-QAM	27
3.1	The structure of the proposed system.	33
3.2	Trellis diagram for the 8-state I-Q QPSK code. (1 bit/sec/Hz) . .	40
3.3	BER performance for 1 bit/sec/Hz I-Q QPSK. The solid lines indicate the upper bound from Eqn. 3.8; dashed lines indicate simulation resullts.	41
3.4	Trellis diagram for the 8-state I-Q 16-QAM code. (2 bits/sec/Hz)	42
3.5	BER performance for 2 bits/sec/Hz I-Q 16-QAM. The solid lines indicate the upper bound from Eqn. 3.8; dashed lines indicate simulation resullts.	43
3.6	Signal partitioning for the 2D 4-AM constellation designed for the Rayleigh distributed channel	44
3.7	Trellis diagram for the 4-D I-Q 16-QAM 8-state code. (3 bits/sec/Hz)	45

3.8 Trellis diagram for the 4-D I-Q 16-QAM 16-state code (3 bits/s/Hz) 45
3.9 Trellis diagram for the 4-D I-Q 16-QAM 32-state code (3 bits/s/Hz) 46
3.10 BER performance for 3 bits/sec/Hz I-Q 4-D 16-QAM. The solid lines indicate the upper bound from Eqn. 3.8; dashed lines indicate simulation resullts. 47
3.11 A comparison between the proposed 1 bit/sec/Hz code and Gray-mapped rate 1/2 convolutional coding. The solid lines indicate I-Q TCM; dashed lines indicate Gray-mapped convolutional codes. 50
3.12 A comparison between the proposed 2 bit/sec/Hz codes and codes from [46]. The solid lines indicate I-Q TCM; dashed lines indicate [46]. 50
3.13 A comparison between the proposed 3 bit/sec/Hz code and codes from [49]. The solid lines indicate I-Q TCM; dashed lines indicate [49]. 51
3.14 BER for the 4-state I-Q 16-QAM encoded system with different decoder buffer sizes. 53
3.15 BER for the 4-state encoded I-Q 16-QAM with different interleaver dimensions. 54
3.16 BER for the interleaved 8-state encoded systems with different throughputs. 54
3.17 A comparison between the I-Q and the Gray-mapped QPSK 16-state codes with different Rician factors (a) $K = 5$ (b) $K = 9$ . . . 56
3.18 BER for the 16-state I-Q QPSK system over both the Rayleigh and Rician ($K = 9$) channels. Solid = Ideal CSI, dashed = NO CSI. 58

3.19 BER comparison between the 16-state I-Q QPSK code and the Gray mapped QPSK with a 16-state rate 1/2 convolutional code over both the Rayleigh and Rician ($K = 9$) channels. All curves assume NO CSI. Solid = I-Q code, dashed = rate 1/2 conv. encoded QPSK. 59

4.1 The cutoff rate of 16-QAM with maximal ratio combining and different diversity orders 68

4.2 The cutoff rate of 16-QAM with equal gain combining and different diversity orders 71

4.3 The cutoff rate of 16-QAM with selection combining and different diversity orders 74

4.4 The cutoff rate of 16-QAM with $M = 2$ and different combining schemes. 76

4.5 The cutoff rate of 16-QAM with $M = 4$ and different combining schemes. 76

4.6 A comparison between the new bound and the Chernoff bound for the 16-QAM I-Q TCM 8-state code with maximal ratio combining and double diversity. 81

4.7 BER of 16-QAM I-Q TCM 4,8, and 16-state codes with maximal ratio combining and different diversity orders. solid(bound), dashed(simulation), (a) $\nu = 2$ (b) $\nu = 3$ (c) $\nu = 4$ 82

4.8 BER of 16-QAM I-Q TCM 4,8, and 16-state codes with equal gain combining and different diversity orders. solid(bound), dashed(simulation), (a) $\nu = 2$ (b) $\nu = 3$ (c) $\nu = 4$ 86

4.9 BER of 16-QAM I-Q TCM 4,8, and 16-state codes with selection combining and different diversity orders. solid(bound), dashed(simulation), (a) $\nu = 2$ (b) $\nu = 3$ (c) $\nu = 4$ 90

4.10 Simulated BER of 16-QAM I-Q TCM 8-state code with different combining schemes. dashed ($M = 2$), solid ($M = 4$). 91

4.11 Bit error probability of uncoded BPSK with different diversity orders and combining schemes (a) MRC (b) EGC (c) SC 94

4.12 A comparison between the upper bound and simulated BER of 16-QAM I-Q TCM 4-state code with MRC dual diversity and $\rho = 0.5$ 99

4.13 Analytical BER of 16-QAM I-Q TCM 4-state code with MRC dual diversity and different correlation values. 100

4.14 Simulated BER of 16-QAM I-Q TCM 8-state code and the 8-PSK TCM 8-state code with MRC and double diversity. 101

4.15 Copmarison between the 16-QAM I-Q TCM 4-state code (with double diversity, MRC and 50% branch correlation) and the 64-state code (with no diversity). 102

5.1 Block diagram of the system model. 107

5.2 Two-state binary Markov source model. 107

5.3 4-D signal points and the trellis diagram of the 4-D QPSK 8-state encoder. 129

5.4 BER of the 2-D and 4-D QPSK 8-state schemes with ML decoding. 130

5.5 BER of the 2-D and 4-D QPSK 8-state schemes with MAP decoding. 130

5.6 Spectral distortion vs. Eb/No for the 4-state 4-D QPSK scheme - solid(Rayleigh), dashed(AWGN). 137
5.7 Spectral distortion vs. Eb/No for the 8-state 4-D QPSK scheme - solid(Rayleigh), dashed(AWGN). 137
5.8 Spectral distortion vs. Eb/No for the 8-state 4-D QPSK scheme (dashed) - and the I-Q QPSK 32-state code (solid) over the Rayleigh channel. 139

Chapter 1

Introduction

In recent years there has been a growing demand for wireless communications services. This has resulted in a renewed interest in providing reliable voice and data transmission over wireless channels. Signals transmitted over wireless communication channels experience fading, imposing extreme variations on the envelope of the transmitted signal. As a result, a substantial degradation of the system error rate performance occurs.

Our emphasis in this thesis is on designing combined coding/modulation methods to improve the performance of wireless systems. Detailed descriptions of the technical issues will be carefully examined in the thesis chapters, where different methods to improve the reliability of transmission are proposed, evaluated, and compared to currently used ones.

1.1 Typical System Model

Figure 1.1 shows a block diagram (in discrete time notation) of a typical digital communication system operating over a frequency non-selective fading chan-

nel. The source encoder is used to reduce the redundancy of the signal, while the channel encoder introduces controlled redundancy to combat channel errors. Suppose at time instant i, x_i is transmitted and y_i is received; two kinds of distortions are imposed on the transmitted signal (multiplicative and additive). This is illustrated as follows

$$y_i = a_i \cdot x_i + n_i \tag{1.1}$$

where a_i is the fading amplitude and n_i is the two-dimensional additive white Gaussian noise with one-sided power spectral density N_o. The fading process has memory; i.e., the fading amplitudes a_i's are correlated, resulting in burst errors. Most channel coding schemes are designed for channels with random (non-bursty) errors. Hence, interleaving/deinterleaving is used. It can be viewed as a zero redundancy code that randomizes the distribution of errors.

The use of interleaving destroys the channel memory. However, a channel with memory has a larger capacity than the corresponding memoryless channel with the same marginal distribution. Hence, interleaving may be used in conjunction with a scheme for extracting the channel state information (CSI) - i.e., the attenuation of the fade - to recover some of the channel capacity lost due to interleaving. Moreover, with the availability of CSI, coherent detection is superior to non-coherent detection [1].

Different schemes for obtaining channel state information are available. The use of pilot tone techniques provides the receiver with an explicit amplitude and phase reference for detection. One of the best solutions among pilot tone techniques is the transparent tone-in-band technique[2]. Recently, an alternative method, called pilot symbol assisted modulation (PSAM), has been proposed [3, 4, 5]. With PSAM, the transmitter periodically inserts known symbols. It

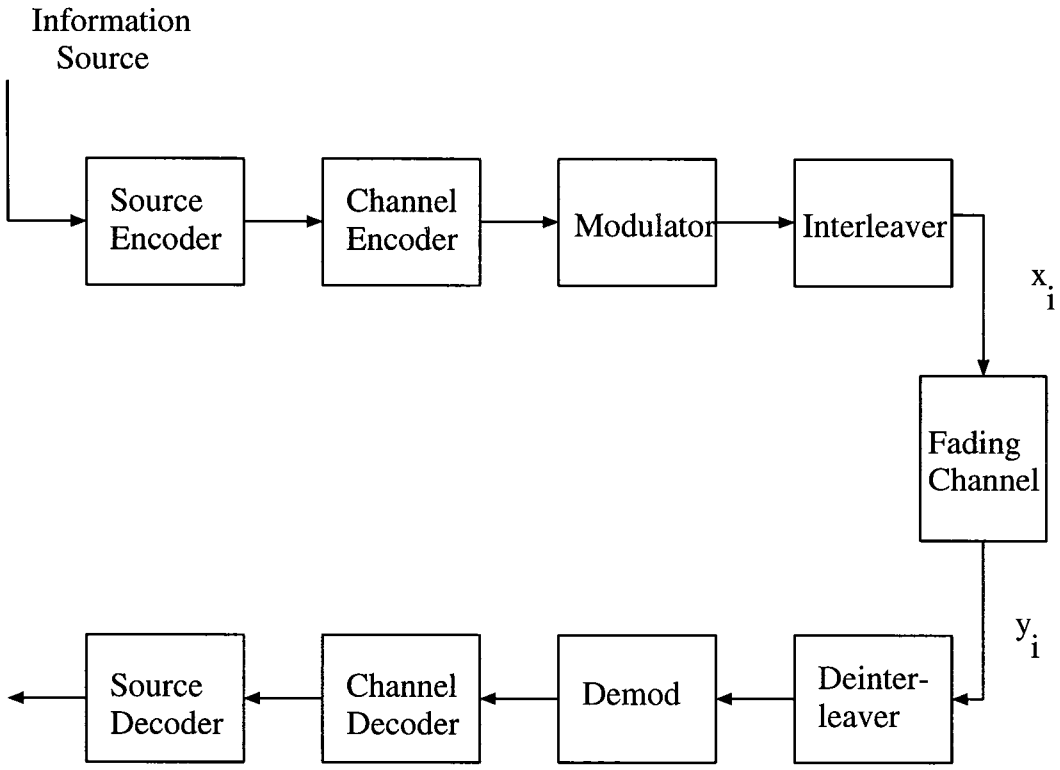


Figure 1.1: A basic block diagram of a digital communication system.

can be used with multilevel modulation. Also, it does not need any change in the pulse shape or the peak to average power ratios. Recently, a new algorithm to extract accurate estimates of the fade's amplitude and phase was proposed [6]. It combines decision-feedback and adaptive linear prediction (DFALP). It has also been shown that coherent systems with the DFALP algorithm outperform differentially coherent systems.

The channel encoder introduces diversity to the system. Diversity represents a way of classifying redundancy, and it is an effective method for reducing the error probability on fading channels [1, 7]. One of its simplest forms is time diversity. L^{th} order time diversity can be achieved by sending a symbol in L time slots where these slots are ideally affected by independent fading. This

method of retransmission L times can be viewed as a repetition code with a minimum Hamming distance of L. Therefore, well known block or convolutional codes with large minimum Hamming distances may be expected to yield large improvements over uncoded systems. However, this approach may result in bandwidth expansion. Such expansion can not be tolerated in bandwidth limited environments. Increased frequency efficiency can be obtained using higher-order modulations; however, a larger signal power is needed to maintain the same error probability. The use of coded modulation provides a solution to the bandwidth expansion problem.

1.2 Coded Modulation

Traditional channel coding schemes are designed independently from the modulator. In 1974, Massey [8] suggested the joint design of the modulator and encoder. Then, the notion of multilevel codes was introduced by Imai and Hirakawa in 1977 [9]. Ungerboeck [10], however, stimulated most of the research in the area of bandwidth efficient coding with his introduction of trellis coded modulation (TCM). TCM schemes are constructed from a convolutional code in conjunction with an expanded signal set. Decoding is done according to Euclidean distance rather than Hamming distance. This approach - called soft decision decoding - leads naturally to the use of convolutional codes with Viterbi decoding. Soft decision decoding has proven to substantially improve error performance, especially when fading exists [11].

There are two approaches to constructing coded modulation schemes. The first is known as trellis coded modulation (TCM) and was introduced by Unger-

boeck. It is concerned with the appropriate mapping of a convolutional (trellis) code onto an expanded signal constellation. A simple trellis coded modulation scheme is shown in Figure 1.2. One information bit enters a rate 1/2 encoder, and the two output bits together with an uncoded bit are mapped to a signal point from the 8-PSK signal constellation. Another scheme is shown where two input bits are uncoded and four bits are mapped to a signal point in the 16-QAM signal constellation.

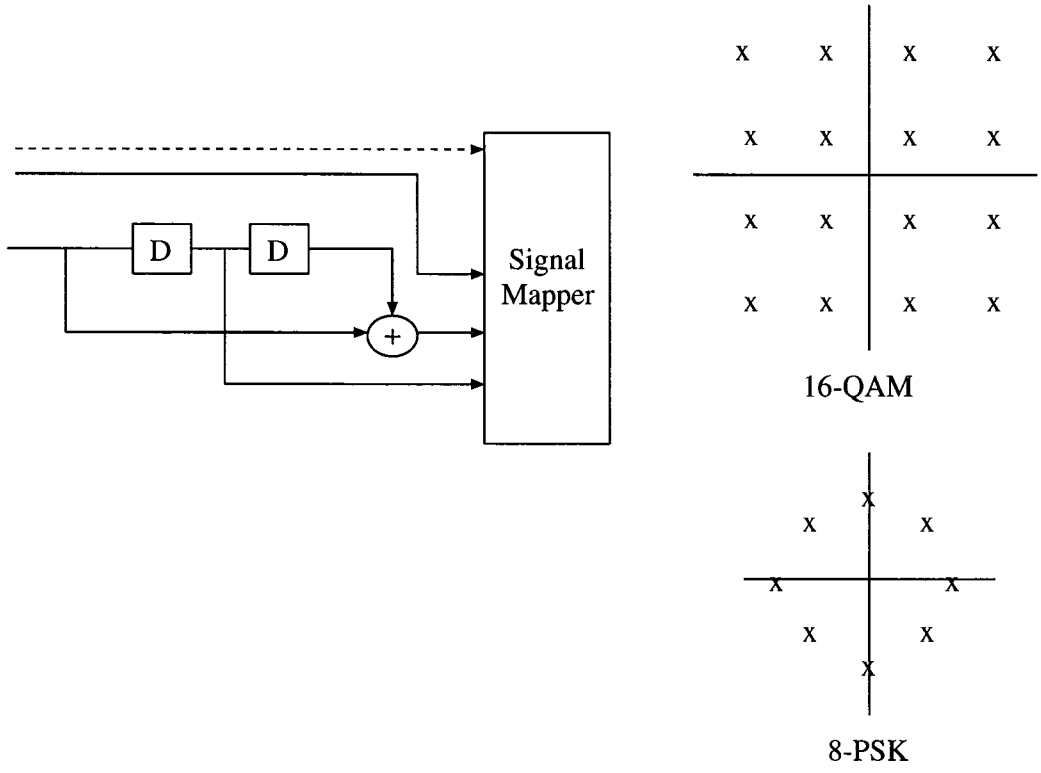


Figure 1.2: A simple TCM scheme for encoding two or three bits per baud.

Generally, to transmit n bits per signal with no redundancy, a signal constellation of cardinality 2^n is required. For the *AWGN non-fading channel*, Ungerboeck showed that most of the coding gain can be achieved by doubling the signal constellation size; i.e., if a signal constellation of size 2^{n+1} is used with appropri-

ate coding to transmit n bits per signal, large coding gains can be achieved. For AWGN channels, the code design goal is to maximize the minimum squared Euclidean distance between all code sequences. Ungerboeck achieved this through the use of "set partitioning", where the constellation is successively partitioned into subsets with increasing intra-subset minimum Euclidean distances. Partitioning is done $k+1$ times ($k \leq n$). At time instant i, k bits enter a rate $\frac{k}{k+1}$ convolutional encoder and $k+1$ bits are produced where they select one of 2^{k+1} possible subsets in the $(k+1)$-level partition while the $(n-k)$ uncoded bits are used to select one of the 2^{n-k} signals in the selected subset.

As a result of the advantages offered by TCM, a huge amount of literature has appeared on the subject. Many researchers have tried to develop a mathematical framework of the theory of TCM. Calderbank and Mazo [12] gave a new analytical description of TCM schemes, combining the coding and mapping into one step. Zehavi and Wolf [13] gave conditions on the regularity of trellis codes and provided a modified generating function approach to give an upper bound on the error probability. Biglieri and Mclane [14] also described the uniform error properties of trellis codes.

It is clear that, as the dimensionality of the signal set increases, performance improvements are achieved. This is basically because more space is obtained to accommodate the signals and hence the distance between signals can be increased. Ungerboeck's codes were restricted to two-dimensional signal points. This motivated researchers to extend the TCM schemes to multi-dimensional constellations. Wilson [15], Calderbank and Sloane [16, 17], and Wei [18] have presented TCM codes with multi-dimensional constellations. Wei [18] described a simple decoding scheme that considerably simplifies the decoder. Calderbank

and Sloane [19] and Forney [20] applied lattice theory to TCM systems. They described the signal constellation as a finite set of points taken from an infinite lattice, and the partitioning of the constellation into subsets corresponds to the partitioning of the lattice into a sub-lattice and its cosets. A recent article [21] gives a good description of the performance evaluation methods of trellis coded modulation schemes.

The second approach to coded modulation, multilevel coding, was initially started by the work of Imai and Hirakawa [9]. The basic idea is to partition the signal constellation Ω_0 and produce a partition chain $\Omega_0/\Omega_1/\Omega_2/\ldots/\Omega_L$ where Ω_i is a refined version of Ω_{i-1}; i.e., it contains more subsets than the previous partition. This partition chain is used with an L-level code $C = (C_1, C_2, \ldots, C_L)$, where each component code C_i is applied at level i of the partition chain. This means that the component code C_i performs the coding upon the subsets that constitute the partition Ω_i.

Tanner [22] has related the minimum squared Euclidean distance of the overall multilevel code to the minimum Hamming distances of the component codes and the intra-subset minimum squared Euclidean distances of the sub-constellations in the partition. The minimum squared Euclidean distance of the multilevel code $d^2(C)$ can be bounded as

$$d^2(C) \geq \min(D_{H1}\delta_0^2, D_{H2}\delta_1^2, \ldots D_{HL}\delta_{L-1}^2, \delta_L^2) \tag{1.2}$$

where D_{Hi} is the minimum Hamming distance of the component code C_i, and δ_i^2 is the intra-subset minimum squared Euclidean distance at level i; i.e., it is the minimum squared Euclidean distance between any pair of signals in the same subset of Ω_i. The δ_i^2's are independent of the code and depend only on the partition chain of the signal constellation. In most of the cases, the relation

$\delta_0^2 \leq \delta_1^2 \leq \ldots \leq \delta_L^2$ is preserved.

Moreover, the minimum free Hamming distance D_H of the overall code is related to the minimum Hamming distances of the component codes as follows

$$D_H = \min(D_{H1}, D_{H2}, \ldots, D_{HL}). \tag{1.3}$$

More results on the design and performance of multilevel code can be found in [23, 24, 25, 26, 27]. A good overview for coded-modulation techniques for the Rayleigh fading channel can be found in [28].

Multistage decoding is a simple iterative decoding technique for multilevel codes that gives performance comparable to full maximum likelihood decoding in AWGN channels. However, it suffers from two problems: error propagation and error multiplicity. Moreover, significant decoding delay is required to effect near-optimal performance.

Often, communication over wireless channels involves the real time transmission of voice and images, requiring minimal decoding delay. Since the decoding of trellis coded modulation schemes requires less decoding delay than multilevel coding techniques, only trellis coded modulation schemes are considered in this thesis.

1.3 Thesis Organization

To increase the reliability of digital transmission over the Rayleigh distributed channel, the use of different techniques such as channel coding, space diversity and maximum a posteriori (MAP) decoding are proposed. The thesis is organized as follows. Chapter 2 gives a brief summary of fading channel models; diversity reception channel models are also described. Moreover, fading models

with space and time correlations are presented. Then, the error performance and coding gains limits for the Rayleigh channel are discussed.

The design and structure of channel codes depends primarily on the channel on which they will be used. Therefore, a trellis coding technique for the flat, slow Rayleigh fading channel is presented in Chapter 3. We call this approach in-phase-quadrature trellis coded modulation (I-Q TCM). I-Q TCM schemes have the advantage of higher minimum time diversities than conventional trellis coding schemes. I-Q TCM schemes for different bandwidth efficiencies (1,2 and 3 bit/sec/Hz) and encoder complexities are presented, analyzed (through simulations and tight upper bounds), and compared to conventionally designed schemes. I-Q TCM schemes are shown to have the highest coding gains and are the closest to the practical limits drawn from cutoff rate expressions.

Chapter 4 presents the effects of space diversity (i.e., multiple receiver antennas) on trellis coded modulation. The practical limits of channel coding gains are presented. Cutoff rate expressions for diversity reception with different combining schemes are derived and compared. The effect of increasing the diversity order is also shown. Cutoff rate expressions with correlated diversity branches and maximal ratio combining are derived to investigate the effect of space correlation on coding gains.

Tight upper bounds on the pairwise error probability for diversity reception assuming different combining schemes are derived. The bounds can be used to evaluate the error performance of both uncoded and coded systems. These upper bounds are expressed in product forms so that transfer function methods can be used to analyze the error performance of trellis codes. A comparison between I-Q TCM schemes and conventional TCM schemes in double diversity reception

is presented to demonstrate the superiority of I-Q TCM schemes even in double diversity reception.

In Chapter 5, a parametric study is presented to demonstrate the usefulness of MAP decoding of trellis codes. Different system configurations, with different sources and codes, are used. It is shown that MAP decoding results in significant gains, compared to ML decoding, in very noisy channel conditions. Also, the effect of the signal constellation size, encoder complexity, and channel type on the MAP decoding gains are presented. A practical example of coding the line spectral parameters (LSP's) of code excited linear predictive (CELP) encoded speech is also shown. Coding of the LSP's is performed using 4-dimensional trellis codes with QPSK modulation. Significant MAP decoding gains are obtained.

Chapter 2

Channel Models

In order to design reliable and spectrally efficient wireless communication systems, accurate characterization of such communication channels is required. In this chapter we will characterize different channel types. Since additive thermal noise is always present in communication systems, we first define the additive white Gaussian noise channel. Then, different fading channels are described using statistical models. The choice of a model depends on the transmission medium (satellite, land mobile radio, etc.). Moreover, diversity models in which two or more channel paths are used to carry the transmitted signal are also considered. Finally, we conclude with discussions on the error performance and cutoff rate for Rayleigh distributed channels.

2.1 The Additive Noise Channel

The additive noise channel is used to model thermal and other natural noise. It is often referred to as white noise since it affects all frequency bands equally. The amplitude of the noise is distributed according to a zero-mean Gaussian

distribution. Hence, it is commonly referred to as *additive white Gaussian noise* (AWGN). We consider the AWGN effect in conjunction with fading.

2.2 The Fading Channel

In typical mobile radio systems, radio signals experience extreme variations in both amplitude and phase. Such variation is a consequence of reflections, diffractions and scattering of the transmitted signal. This phenomenon is called fading.

Fading has both "long-term" and "short-term" effects on the channel [7]. Long-term fading affects the availability of the channel. It has a slow-varying mean that varies approximately according to a log-normal distribution. Hence, it is known as log-normal or power fading. Short-term fading, on the other hand, is the one that affects the details of the received waveform. It is known as multipath fading. We only consider short-term fading in this thesis and simply call it fading. It can be best described through a linear time-varying impulse response. In the following, baseband rather than bandpass notation is used since it significantly simplifies the analysis.

2.2.1 The Linear Time-Varying Model

Short-term fading is usually modeled as a time-varying linear filter [7] with a low-pass impulse response $h(\tau; t)$, where τ represents the propagation delay and t represents time. With the slow variation of the channel and considering the short-term fading effect, it can be assumed that the process is wide-sense stationary (WSS) [7]. Hence, we can define the autocorrelation function of the

impulse response of the channel as

$$R_h(\tau_1, \tau_2; \Delta t) = \frac{1}{2} E[h(\tau_1; t) h^*(\tau_2; t + \Delta t)]. \tag{2.1}$$

In most fading environments, the attenuation and phase shift of the channel associated with one path delay τ_1 is uncorrelated with that of another path delay τ_2. This is usually referred to as *wide sense stationary uncorrelated scattering* (WSSUS). This assumption can simplify the previous definition of $R_h(\tau_1, \tau_2; \Delta t)$ to

$$R_h(\tau_1, \tau_2; \Delta t) = R_h(\tau_1; \Delta t)\delta(\tau_1 - \tau_2) \tag{2.2}$$

where, $R_h(\tau; \Delta t) = \frac{1}{2} E[h(\tau; t) h^*(\tau; t + \Delta t)]$ An important quantity is the value of $R_h(\tau, \Delta t)$ at $\Delta t = 0$; i.e., $R_h(\tau; 0)$. It is referred to as the **multipath profile** (or delay power spectrum), and it represents the average output power of the channel at a relative path delay of τ. The range of values of τ over which $R_h(\tau; 0)$ is essentially non-zero is called the **multipath spread** of the channel and is denoted by T_m. In other words, $\frac{1}{T_m}$ gives the maximum frequency separation for which the received signals are still correlated. If the symbol period T does not satisfy $T \not\gg T_m$, then adjacent symbols are correlated and this gives rise to intersymbol interference and the channel is said to be frequency selective. In narrow band systems where the symbol period is usually much larger than T_m ($1/T \ll 1/T_m$), the channel is frequency non-selective and is usually referred to as a *flat* fading channel.

Another important function, denoted by $R_H(\Delta f; \Delta t)$, is called the spaced-tone autocorrelation function [7]. It is defined as

$$R_H(f_1, f_2; \Delta t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} R_h(\tau_1, \tau_2; \Delta t) e^{-j2\pi f_1 \tau_1} e^{j2\pi f_2 \tau_2} d\tau_1 d\tau_2. \tag{2.3}$$

Using the assumption of uncorrelated scattering (Eqn. 2.2), it is easy to show that

$$R_H(f_1, f_2; \Delta t) = R_H(\Delta f; \Delta t) \tag{2.4}$$

where $\Delta f = f_1 - f_2$. The importance of $R_H(\Delta f; \Delta t)$ is evident in its value at $\Delta t = 0$; i.e., $R_H(\Delta f; 0)$, which is the Fourier transform of the multi-path profile $R_h(\tau; 0)$. The spaced-tone autocorrelation function describes the selectivity of the channel. The **coherence bandwidth**, denoted by $(\Delta f)_c$, is defined where $R_H(\Delta f; 0)$ is nearly constant over $(\Delta f)_c$. From the Fourier transform relationship between the multipath profile and the spaced-tone autocorrelation function, the multipath spread and the coherence bandwidth can be related by $(\Delta f)_c \approx \frac{1}{T_m}$.

Considering the Δt-dependence of $R_h(\tau; \Delta t)$, the **scattering function**, denoted by $S_h(\tau; \Omega)$, is defined as the Fourier transform of $R_h(\tau; \Delta t)$ with respect to Δt; i.e.,

$$S_h(\tau; \Omega) = \int_{-\infty}^{\infty} R_h(\tau; \Delta t) e^{(-j2\pi\Omega\Delta t)} d(\Delta t) \tag{2.5}$$

where Ω denotes the frequency variable (Doppler frequency). Similarly, the Fourier transform of $R_H(\Delta f; \Delta t)$ with respect to Δt is defined as

$$S_H(\Delta f; \Omega) = \int_{-\infty}^{\infty} R_H(\Delta f; \Delta t) e^{(-j2\pi\Omega\Delta t)} d(\Delta t). \tag{2.6}$$

From the uncorrelated scattering assumption, it is easy to see that $S_h(\tau; \Omega)$ and $S_H(\Delta f; \Omega)$ form a Fourier transform pair. The importance of the scattering function is that it describes the dynamic changes of the channel characteristics. This is obvious in the **Doppler power spectrum** value, $V(\Omega)$, defined as

$$V(\Omega) = \int_{\tau} S_h(\tau, \Omega) d\tau \tag{2.7}$$

and it indicates how much the spectrum is spread if a pure sine wave is transmitted. The nominal width of $V(\Omega)$, denoted by f_D, is referred to as the **Doppler spread**. In other words, $\frac{1}{f_D}$ gives the time span over which the received symbol can be detected coherently. The Doppler spread indicates how fast the channel is. Typical values of f_D range between 20-100 Hz.

Another important quantity, denoted by $(\Delta t)_c$, is the coherence time. It is defined as the time span Δt over which $R_H(0; \Delta t)$ remains nearly constant. From the Fourier transform relation, it can be deduced that $(\Delta t)_c \approx \frac{1}{f_D}$.

Based on the previous definitions, four classes of fading channels can be identified based on the transmitted symbol duration T :

1. Slow, frequency non-selective fading : $T_m \ll T \ll (\Delta t)_c$.
2. Slow, frequency selective fading : $T_m \not\ll T$ and $T \ll (\Delta t)_c$.
3. Fast, frequency non-selective : $T_m \ll T$ and $T \not\ll (\Delta t)_c$.
4. Fast, frequency selective fading : $T_m \not\ll T$ but $T \not\ll (\Delta t)_c$.

The spread factor (L) is defined to be the product of the multipath spread and Doppler spread; i.e., $L = T_m f_D$. In most fading channels, $L \ll 1$ and the channel is said to be **underspread**.

2.2.2 Statistical Models

Different statistical models are used to describe fading channels. The choice of a model depends on the nature of the transmission media. The most commonly used models are described below

- **The Rayleigh Model**

  Because of the multipath nature of the propagation media, the received signal is composed of a sum of transmitted signals each with different amplitude and phase. If there is only a diffuse collection of multipath signals – i.e., there is no significant "line-of-sight" path – we can conclude from the central limit theorem that the in-phase and quadrature multiplicative fades are independent Gaussian random processes with zero mean and a variance of σ^2. This means that the multiplicative envelope of the signal is Rayleigh distributed with the following pdf :

$$f(a) = \frac{a}{\sigma^2} \exp\left(-\frac{a^2}{2\sigma^2}\right), \quad a \geq 0. \tag{2.8}$$

  The phase will have a uniform distribution on $[0, 2\pi]$ but since the fading is slow, phase tracking by the synchronization circuit is usually assumed.

  Rayleigh distributed fading channels represent "the worst case" of the Rician family of distributions (see below) used to characterize the channel behavior of land mobile radio (LMR) systems [1, 7, 11, 28]. If the coherence bandwidth is much greater than the signal bandwidth, the channel is considered as a frequency non-selective. This means that the channel imposes a multiplicative distortion on the signal. An important example of this channel is that of the narrow-band cellular mobile radio.

- **The Rician Model**

  In the Rician fading model, a line-of-sight (LOS) is assumed to exist between the transmitter and the receiver. Also, additional diffuse paths resulting from multipath reflections exist. Such a situation occurs most in mobile satellite communications. The multiplicative envelope of the signal

is accordingly Rician distributed with the following pdf :

$$f(a) = 2a(1+K)\exp -(K + a^2(1+K))I_o(2a\sqrt{K(K+1)}), \quad a \geq 0 \tag{2.9}$$

where $I_o(x) = \sum_{n=0}^{\infty}\left(\frac{x^n}{2^n n!}\right)^2$ is the zero-order modified Bessel function of the first kind and K is the Rician parameter defined as ratio of the energy of the direct component to the energy of the diffuse components.

- **Nakagami Model**

  Another fading model is the one introduced by Nakagami [29]. It was empirically derived by fitting envelope statistics around the mean or median rather than near the zero region. The envelope in Nakagami fading is distributed according to

$$f(a) = \frac{2}{\Gamma(m)}\left(\frac{m}{\Omega_m}\right)^m a^{(2m-1)}e^{-ma^2/\Omega_m} \tag{2.10}$$

  where $\Gamma(x) = \int_0^\infty y^{x-1}e^{-y}dy(x > 0)$ is the Gamma function, m is the Nakagami parameter ($m \geq 0.5$) and $\Omega_m = E[a^2]$.

  The Rayleigh distribution is a special case of the Nakagami model when $m = 1$. When $m = 1/2$ the Nakagami model reduces to the Gaussian distribution, and when $m \to \infty$ the channel becomes non-fading. Moreover, there is a close fit between the Rician and Nakagami distributions where K and m are related by

$$K = \frac{\sqrt{m^2-m}}{m-\sqrt{m^2-m}} \quad m > 1. \tag{2.11}$$

  This fit is accurate for small values of E_S/N_o. However, at large values the difference between the models becomes large [30].

2.2.3 The Diversity Model

Diversity combining is a well-known and effective method for improving the performance of digital communication systems over fading channels [1, 7, 31]. The basic principle of M-fold diversity is to use M independent channels so that the probability of a "deep fade" on *all* channels is low. (These independent channels can be created in a number of ways, including separation in frequency, time, and/or polarization; if multiple antennas are used to receive multiple independent versions of the received signal, the approach is called *spatial* diversity.) A combining circuit is used to form a single resultant signal from the M different "branch" signals.

A diversity reception system can be described as follows. Suppose the complex signal x_i is transmitted at time i. At the receiver, M corresponding signals $\underline{y}_i = \{y_{i,1}, y_{i,2}, \ldots, y_{i,M}\}$ exist – i.e.,

$$\begin{aligned} y_{i,1} &= a_{i,1}x_i + n_{i,1} \\ y_{i,2} &= a_{i,2}x_i + n_{i,2} \\ &\vdots \\ y_{i,M} &= a_{i,M}x_i + n_{i,M}. \end{aligned} \tag{2.12}$$

Here, $\underline{a}_i = \{a_{i,1}, a_{i,2}, \ldots, a_{i,M}\}$ are the fading amplitudes (typically Rayleigh or Rician) and $\underline{n}_i = \{n_{i,1}, n_{i,2}, \ldots, n_{i,M}\}$ are the two-dimensional additive white Gaussian noise samples with one-sided power spectral density N_o. We will focus our interest on the performance of diversity systems employing trellis coded modulation assuming a variety of different signal combining methods.

2.3 Correlation Models

In this section we discuss two types of correlation that are of concern when dealing with fading channels. Since fading has memory, adjacent signal symbols are affected by correlated fade amplitudes. The correlation function depends on the transmission media and the receiver. Moreover, when space diversity is used at the receivers, it may be difficult to separate the antennas sufficiently far apart so that their received signals are uncorrelated. In the following we discuss these two types of correlation.

2.3.1 Time Correlation

In time correlated fading, the successive fade amplitudes (a_i's) are correlated. In land mobile radio, assuming an omnidirectional antenna at the mobile and that the received plane waves are uniformly distributed in arrival angle, the autocorrelation function $R_H(0;\Delta t)$ can be represented as [31, 32]

$$R_H(0;\Delta t) = J_0(2\pi f_D \Delta t) \tag{2.13}$$

where $J_0(x) = \sum_{n=0}^{\infty}(-1)^n(\frac{x^n}{2^n n!})^2$ is the zero-order Bessel function of the first kind, and f_D is maximum Doppler frequency and is given by

$$f_D = \frac{v}{c} f_c \tag{2.14}$$

where v is the mobile unit speed, f_c is the carrier frequency and c is the speed of light. Figure 2.1 shows a plot of the autocorrelation function. It can be seen that the correlation decay is very slow and that it goes negative.

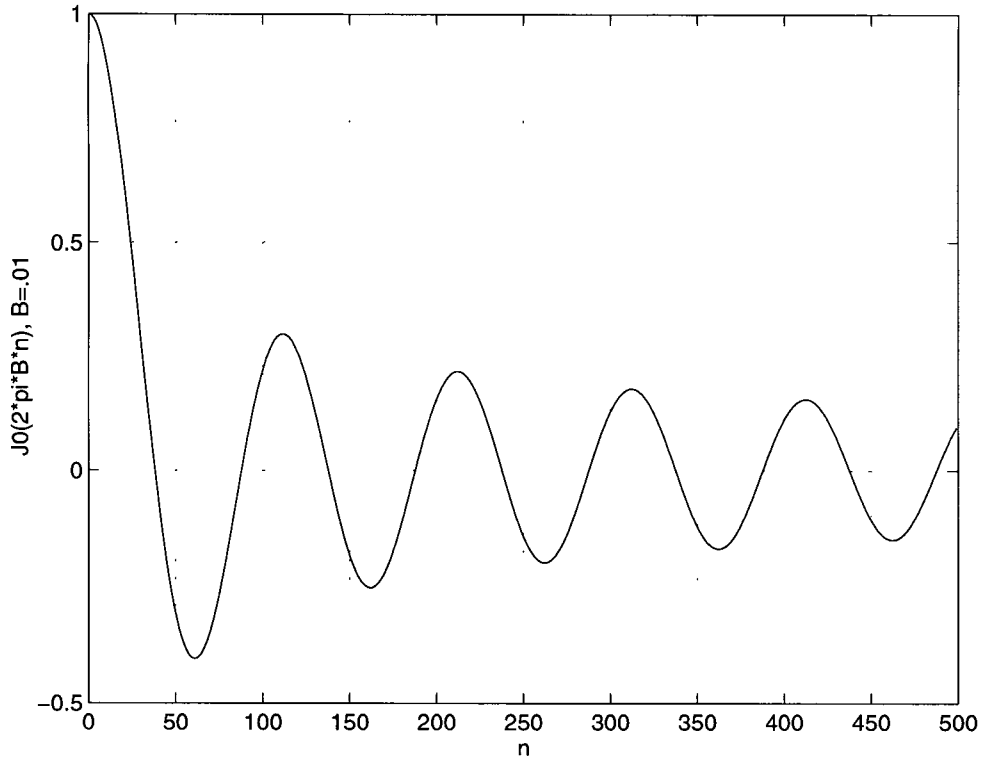


Figure 2.1: The autocorrelation function .

The Doppler power spectrum can be written as as

$$V(f) = \begin{cases} \frac{1}{2\pi\sqrt{{f_D}^2 - f^2}} & \text{if } |f| < |f_D| \\ 0 & \text{otherwise.} \end{cases} \tag{2.15}$$

This fading model can be simulated by generating first two sequences of i.i.d. Gaussian random variables, passing each sequence through a low pass filter with a spectrum specified by Eqn. 2.15. The Rayleigh fade will be the envelope of the two Gaussian random variables, one from each sequence. The low pass filter is constructed using an FIR filter designed to approximate the response of the ideal filter. In our construction, a 200-tap FIR was used.

2.3.2 Space Correlation

If two antennas are separated by a distance d, then the correlation coefficient between the signals of the two antennas is given by [31]

$$\rho = J_0(\frac{2\pi d}{\lambda}) \tag{2.16}$$

where λ is the carrier wavelength. The first zero of the Bessel function occurs when $d = 0.38\lambda$. In practice, a separation distance of $d = 0.5\lambda$ is used [31]. For example, with a carrier frequency $f_c = 900$ MHz, a separation distance between the receiving antennas of 6–7 inches will result in practically negligible correlation. For higher carrier frequencies less separation is required.

In many systems employing spatial diversity, it may be difficult to achieve uncorrelated branches. This may be due to improper antennas positioning or receiver space limitations. Branch correlation can be simulated using a linear transformation of independent Gaussian variables. Observe that the Rayleigh fade amplitude, $a_{i,l}$, can be written as

$$a_{i,l}^2 = |g_{i,l}|^2 \tag{2.17}$$

where $g_{i,l}$ is a complex Gaussian random variable with zero mean and a variance of one-half for both the real and imaginary parts, and $|\cdot|$ denotes the envelope. Therefore, we can write

$$\sum_{l=1}^{M} a_{i,l}^2 = \underline{g}_i \underline{g}_i^* \tag{2.18}$$

where $\underline{g}_i = \{g_{i,1}, \ldots, g_{i,M}\}$ and $(\cdot)^*$ denotes the Hermitian transpose. The probability density function of $\underline{g}_i$ is expressed as

$$f(\underline{g}_i) = \frac{1}{\pi^M \det K_{\underline{g}_i}} \exp(-\underline{g}_i K_{\underline{g}_i}^{-1} \underline{g}_i^*) \tag{2.19}$$

where $K_{\underline{g}_i}$ is an $M \times M$ covariance matrix with entries $(K_{\underline{g}_i})_{l,k} = E(g_{i,l}g^*_{i,k})$. Since the $g_{i,l}$'s are uncorrelated, $K_{\underline{g}_i}$ will be simply the identity matrix.

We are interested in generating a new vector $\underline{h}_i = \{h_{i,1}, \ldots, h_{i,M}\}$ with a specified covariance matrix, $K_{\underline{h}_i}$. This can be accomplished using the transformation

$$\underline{g}_i = A\underline{h}_i \tag{2.20}$$

Using this transformation, $K_{\underline{h}_i}$ is expressed as follows

$$K_{\underline{h}_i} = AA^t \tag{2.21}$$

where A^t is the transpose of A. Since $K_{\underline{h}_i}$ is a symmetric matrix, it can be represented in the form

$$K_{\underline{h}_i} = P\Lambda P^t, \tag{2.22}$$

where Λ is a diagonal matrix that consists of the eigenvalues of $K_{\underline{h}_i}$ and P is a matrix whose columns are the orthonormal set of the eigenvectors of $K_{\underline{h}_i}$. Therefore, we can write

$$AA^t = P\Lambda^{1/2}\Lambda^{1/2}P^t \tag{2.23}$$

i.e., if we let $A = P\Lambda^{1/2}$, we get the desired covariance matrix. For example, to generate two branches with the following covariance matrix

$$K_{\underline{h}_i} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{2.24}$$

it is straightforward to show that

$$A = \begin{bmatrix} \sqrt{(1-\rho)/2} & \sqrt{(1+\rho)/2} \\ -\sqrt{(1-\rho)/2} & \sqrt{(1+\rho)/2} \end{bmatrix} \tag{2.25}$$

2.4 The Rayleigh Channel Error Performance

The research conducted in this thesis considers only short-term fading. The fading is assumed to be slow with a signaling interval much less than the coherence time of the channel. The fading is also assumed to be flat – i.e., frequency non-selective. This work concentrates on land mobile radio channels. Hence, Rayleigh distributed fading is assumed throughout most of the thesis. However, some results for the Rician model are given in Chapter 3.

In the following we discuss the error performance of signals transmitted over the flat Rayleigh distributed channel. Suppose at time instant i, x_i ($x_i \in A$) is transmitted and y_i is received - i.e,

$$y_i = a_i \cdot x_i + n_i \tag{2.26}$$

where a_i is the fading amplitude having the Rayleigh distribution and n_i is the two-dimensional additive white Gaussian noise with one-sided power spectral density N_o

The sequence of signals $\mathbf{x}_N = (x_1, x_2, \ldots, x_N)$ is transmitted. At the receiver, two sequences, the received signal $\mathbf{y}_N = (y_1, y_2, \ldots, y_N)$ and the channel fade amplitude sequence $\mathbf{a}_N = (a_1, a_2, \ldots, a_N)$, are the inputs to the TCM decoder which performs maximum likelihood (ML) decoding, with ideal channel state information. The decoding metric is

$$m(\mathbf{y}_N, \mathbf{x}_N; \mathbf{a}_N) = \sum_{l=1}^{N} \ln P(y_l|x_l, a_l) = \sum_{l=1}^{N} -|y_l - a_l x_l|^2. \tag{2.27}$$

The pairwise error probability $P(\mathbf{x}_N \rightarrow \hat{\mathbf{x}}_N)$ is the conditional probability that the decoder estimates the transmitted signal to be $\hat{\mathbf{x}}_N$, given that it is in fact $\mathbf{x}_N$. This happens only if

$$m(\mathbf{y}_N, \hat{\mathbf{x}}_N; \mathbf{a}_N) > m(\mathbf{y}_N, \mathbf{x}_N; \mathbf{a}_N) \tag{2.28}$$

which can be written as

$$\sum_{l=1}^{N} -|y_l - a_l\hat{x}_l|^2 > \sum_{l=1}^{N} -|y_l - a_l x_l|^2. \tag{2.29}$$

To obtain an upper bound on the pairwise error probability, we first condition on the fade amplitude sequence $\mathbf{a}_N$. Then, using the Chernoff bound and optimizing the Chernoff factor, $P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N)$ can be simplified to

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) \leq \prod_{l=1}^{N} \exp\left(\frac{-a_l^2|x_l - \hat{x}_l|^2}{4N_o}\right). \tag{2.30}$$

Finally, averaging over the fading amplitude reveals

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \prod_{l\in\eta} \frac{1}{1 + \frac{|x_l - \hat{x}_l|^2}{4N_o}} \tag{2.31}$$

where $\eta = \{l : x_l \neq \hat{x}_l\}$. From the above expression we see that the pairwise error probability is mainly affected by the symbol-wise Hamming distance of the code sequence and so the error event probability will be dominated by the code's minimum symbol-wise Hamming distance, also known as *the minimum time diversity* of the code. A secondary parameter is the minimum product distance among all error events with minimum time diversity. This means that in finding good trellis codes for the Rayleigh distributed fading channel, the minimum time diversity is the most important parameter, and not the minimum free Euclidean distance of the code as it is for AWGN channels.

The channel coding theorem given by Shannon [33] states that given a particular channel model, there exists a maximum rate of reliable transmission called the capacity of the channel; as long as the transmission rate does not exceed the channel capacity, there exists some coding scheme which can be used to achieve an arbitrary degree of reliability. Knowledge of the channel capacity tells us how much potential gain can be achieved.

Some early work on the capacity of the Rayleigh channel was considered by Ericsson [34] and Lee [35]. Recently, some information theoretic limits for fading channels was given by Buz [36]. He showed that the asymptotic loss due to independent Rayleigh fading (with CSI) with respect to the AWGN non-fading channel is 2.51 dB. He also showed that using the pilot tone technique for channel estimation is essentially ideal for high SNR. Moreover, it was shown that the use of space diversity reclaims a significant amount of the loss experienced due to Rayleigh fading. Most of the gain is achieved by using two antennas.

To compare different modulation schemes over a certain channel type, the cutoff rate parameter, R_o, is the appropriate design criterion. R_o is the largest number such that there exists sequence codes with rate R and increasing block-length N such that

$$P(E) \leq 2^{-N(R_o - R)} \tag{2.32}$$

for N sufficiently large. Here, $P(E)$ is the average error probability. R_o establishes a limit on the rate at which one can communicate with arbitrary small error probability; i.e., when $R < R_o$, $P(E) \to 0$ as the code block length $N \to \infty$. Although it is not the ultimate limit - that would be the capacity - it is considered to be the practical limit for coded systems.

The cutoff rate was first introduced by Wozencraft and Kennedy in 1966 [37] using random coding bound arguments. It received more attention when it was introduced by Massey in 1974 where he promoted its use as a criterion for joint modulation and coding systems [8]. The higher the R_o value for a given E_b/N_o the better the modulation system is. For a continuous valued fading channel,

the cutoff rate, in bits/transmitted signal, is given by [38]

$$R_o = \lim_{N\to\infty} \max_{\Pr(\mathbf{x}_N)} \left[-\frac{1}{N} \log_2 \left(\int \int [\sum_{\mathbf{x}_N} \Pr(\mathbf{x}_N) \sqrt{f(\mathbf{y}_N|\mathbf{x}_N, \mathbf{a}_N) f(\mathbf{a}_N)}]^2 d\mathbf{a}_N d\mathbf{y}_N \right) \right] \tag{2.33}$$

where $\Pr(\mathbf{x}_N)$ is the a priori probability of transmitting the sequence $\mathbf{x}_N$, $f(\mathbf{y}_N|\mathbf{x}_N, \mathbf{a}_N)$ is the conditional pdf of $\mathbf{y}_N$ given the input sequence and the fade depth sequence, and $f(\mathbf{a}_N)$ is the fade sequence distribution. For the fully interleaved channel (i.e., memoryless), this expression reduces to

$$R_o = \max_{\Pr(x_i)} \left[-\log_2 \left(\sum_{x_i \in A} \sum_{x_j} \Pr(x_i) P(x_j) \min_{\nu} C(x_i, x_j, \nu) \right) \right] \tag{2.34}$$

where $\Pr(x_i)$ is the probability of transmitting x_i, and $C(x_i, x_j, \nu)$ is the Chernoff factor of x_i and x_j, defined as

$$C(x_i, x_j, \nu) = E_{a_i} \left[E_{y_i|x_i} [\exp(\nu\{m(y_i, \hat{x}_i; a_i) - m(y_i, x_i; a_i)\})] \right]. \tag{2.35}$$

The minimal Chernoff factor is given by

$$C(x_i, x_j) = \min_{\nu} C(x_i, x_j, \nu) = \frac{1}{1 + \frac{|x_i - x_j|^2}{4N_o}}. \tag{2.36}$$

Also, the cutoff rate is maximized with a uniform input distribution [38]. Therefore, it can be written as

$$R_o = 2\log_2(|A|) - \log_2 \left(\sum_{x_i \in A} \sum_{x_j \in A} \frac{1}{1 + \frac{|x_i - x_j|^2}{4N_o}} \right). \tag{2.37}$$

Cutoff rate curves, in Figure 2.2, show that a significant coding gain can be achieved using a larger signal constellation. For example, if a system with two information bits per transmitted signal is to be designed, a QPSK uncoded system requires an $E_s/N_o = 50$ dB to provide a bit error probability of 10^{-5}. However, if an 8-PSK signal constellation is used with a powerful coding scheme,

it can provide the same error performance at $E_s/N_o = 12$ dB. This represents a coding gain of 38 dB. Moreover, if 16-QAM is used the cutoff rate limit is reduced to $E_s/N_o = 11$ dB. In the next chapter we will introduce highly reliable trellis codes that achieve coding gains very close to the cutoff rate limit.

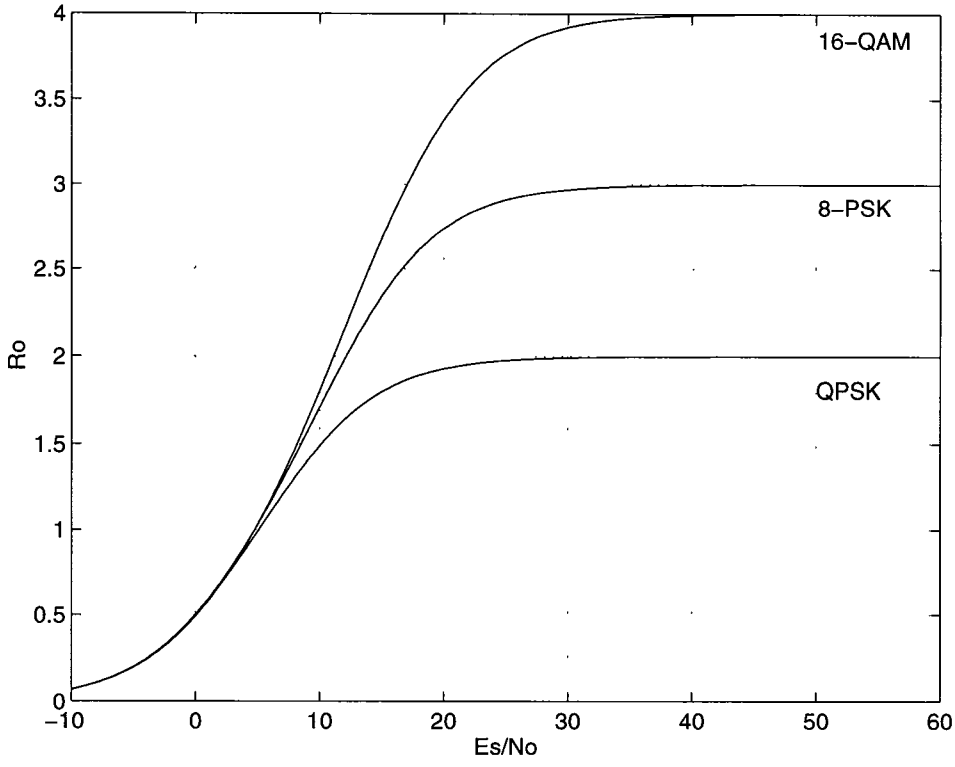


Figure 2.2: Cutoff rates for QPSK, 8-PSK and 16-QAM

In [38], it is shown that the performance of coded systems significantly degrades over correlated fading channels, relative to channels with i.i.d. fading. Therefore, interleaving/deinterleaving is used to break up error bursts caused by amplitude fades of duration greater than one symbol. It is reported that interleaving with reasonably long interleaving depths is almost as good as ideal infinite interleaving.

The interleaver can be regarded as a buffer with d rows and s columns. Here,

d represents the depth of interleaving while s represents the span. The interleaver size is therefore $d \times s$. The signal points are fed into the buffer in successive rows and transmitted out of the buffer column-by-column. The deinterleaver performs the reverse operation. In practice the interleaver depth should be chosen such that the signal amplitudes received in any two signaling intervals are subjected to fading that is as close to independent as possible. In the case of cellular radio, the signal amplitudes received at two different locations separated by a distance of a half carrier-wavelength are independent [31]. For example, for a vehicle with a speed of v, the corresponding distance associated with one symbol duration is vT. Therefore, the interleaver depth should be designed such that $dvT > 0.5\lambda$, where λ is the carrier wavelength.

The interleaver span, s, should be chosen greater than the decoding depth of the code so that correlated amplitudes will not affect the decoding of the current symbol.

Chapter 3

I-Q Trellis Coded Modulation

3.1 Introduction

In cellular mobile radio channels, the transmitted signal typically suffers distortion from fading. One of the most commonly assumed models is that of a frequency non-selective slowly Rayleigh distributed fading channel; this model has been used to characterize the channel behavior of mobile radio systems [1, 11, 28, 31]. The effect of fading is a substantial degradation of the system error rate performance.

Diversity is a popular method for mitigating the effects of fading. One of its simplest forms is time diversity, in which each symbol is repeated in L different time slots such that the slots are affected by independent fades. This approach – called L-fold time diversity – can be viewed as an $(L, 1)$ repetition code with a minimum Hamming distance of L. Therefore, it may be expected that more sophisticated block or convolutional codes with large minimum Hamming distance could yield significant gains over uncoded systems.

However, channel coding raises the possibility of bandwidth expansion; in bandlimited environments, such expansion might not be acceptable. Ungerboeck [10] stimulated most of the research in bandwidth-efficient coding with his introduction of trellis coded modulation (TCM). While originally applied to AWGN channels, the use of TCM for fading channels has received considerable attention in recent years.

3.1.1 Relevant Past Work

Divsalar and Simon [39, 40] were the first to evaluate the performance of trellis coded modulation for Rayleigh/Rician channels. They showed that substantial performance improvement can be achieved using simple TCM schemes combined with interleaving; the interleaving can be viewed as a zero redundancy code that "randomizes" the distribution of the errors, breaking up bursts and destroying the memory of the channel. Interleaving may be used in conjunction with a scheme for extracting channel state information (CSI) – i.e., the attenuation of the fade – to recover some of the channel capacity lost due to the interleaving.

Divsalar and Simon [39] also derived Chernoff upper bounds on the pairwise error probability, assuming independent fading (i.e., infinite interleaving) and perfect recovery of the channel state information. They pointed out that the effective minimum time diversity of the code is equivalent to its minimum Hamming distance (in signal symbols); moreover, the performance of a code over the Rayleigh-distributed fading channel depends strongly on the code's minimum time diversity and very weakly on its minimum squared Euclidean distance – the most important performance parameter for non-fading AWGN channels. This means that trellis codes that are optimal for non-fading AWGN channels are, ex-

cept for a few cases, sub-optimal for fading channels. An exact characterization of the pairwise error-probability in the Rayleigh fading environment was given in [41] and [42]. Also, an upper bound on the bit error probability for Rician channels is given in [43].

A number of papers dealing with the optimization of TCM codes for the Rayleigh channel have appeared [44]-[49]. Most of these coding schemes use the traditional Ungerboeck approach in their design – i.e., they involve doubling the constellation size over what is required for uncoded transmission and the use of a rate $k/(k+1)$ encoder to describe valid symbol sequences. However, if an encoder with k input bits and one output channel symbol is used, the achievable minimum time diversity L is upper bounded by $L \leq \lfloor \nu/k \rfloor + 1$, where ν is the number of memory elements in the encoder. Therefore, the larger the number of the encoder input bits, the more memory elements are needed for a given minimum time diversity. This suggests that if the input bits are distributed to different convolutional encoders, the minimum time diversity could be increased. This is one motivation behind using multilevel codes for the Rayleigh channel [50, 51, 52, 53, 54]. However, multilevel codes with multistage decoding require a large delay at the decoder to get near-optimal performance.

The problem of transmitting two bits/symbol over the Rayleigh channel with 8-PSK modulation was considered by Zehavi [55]; he showed that using three bit interleavers with a sub-optimal decoding metric can provide additional coding gain. The use of multidimensional trellis codes for the Rayleigh fading channel was investigated in [56] and [57]. Multidimensional TCM permits the design of systems with non-integer bandwidth efficiencies and improves the error performance for low constraint length codes. However, a disadvantage of multidimen-

sional TCM is the increased number of input bits to the encoder, which limits the minimum time diversity of the code. Also, more interleaving and a larger decoder buffer size are needed; this increases the delay, which may prohibit the use of such codes in real-time applications.

All of the results described above fall far short of the performance promised by R_0, the computational cutoff rate limit. For example, for 8-PSK TCM schemes with two information bits per baud over a channel with independent Rayleigh fading, cutoff rate curves show that reliable communication can be achieved at $E_b/N_o = 9$ dB, while the best code currently known requires $E_b/N_o = 14.5$ dB to provide a BER of 10^{-5} [55].

3.1.2 Proposed Codes

To achieve a larger minimum time diversity, we propose distributing the data bits to two parallel encoders; the first encoder encodes the in-phase component of the signal while the second encodes the quadrature component. (See Figure 3.1) This approach requires only two encoders/decoders and permits the two decoders to work independently; by comparison, multilevel codes typically require more than two encoders/decoders and incur substantial delay to approach optimal performance. In this chapter, examples of codes with bandwidth efficiencies of 1,2, and 3 bits/sec/Hz and different constraint lengths are described. Their minimum time diversities are shown to be greater than that of the corresponding conventionally designed codes with the same complexity.

The idea of I-Q TCM first appeared in the "pragmatic" TCM design of Viterbi *et. al.* in which they proposed two off-the-shelf rate 1/2 64-state encoders to encode the in-phase and quadrature components of a QAM constella-

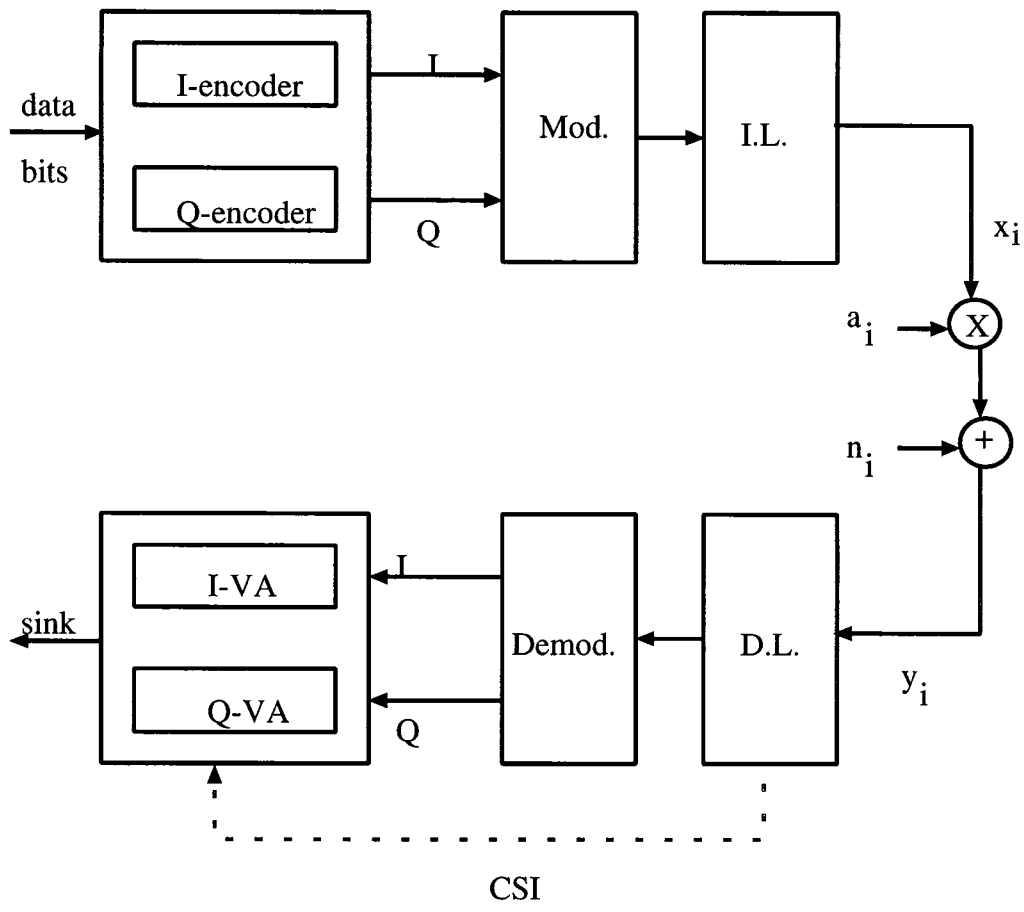


Figure 3.1: The structure of the proposed system.

tion [58]. This approach results in quadrupling the QAM constellation over what is required for uncoded transmission. Another approach that uses off-the-shelf convolutional codes with QAM modulation was suggested by Heegard, Lery, and Paik; in [59] they demonstrated how coded QAM modulation can be implemented using QPSK-based coded modulation, eliminating the need to quadruple the constellation.

Moher and Lodge [3] used the I-Q approach for transmitting 2 bits/sec/Hz on a satellite channel modeled with a Rician distribution; they showed that pilot symbol aided techniques could effectively extract channel state information. Ho,

Cavers, and Varaldi [60] used the I-Q approach to compare the performance of a single 8-state octal PSK trellis code with that of two parallel 4-state encoders – the latter resulting in 16-QAM signaling; they show that a large coding gain over the Rayleigh distributed channel is achieved, even if the ratio of peak-to-average power is deducted from the coding gain. However, no codes were given for a higher number of states or for different bandwidth efficiencies.

Our work here is a generalization of this idea. Systems with a throughput of 1, 2, and 3 bits/sec/Hz and different constraint lengths are described and compared to their corresponding conventional codes. If the throughput is not an even number, then the encoder is operated every two signaling intervals and thus produces 4-dimensional coded signals. Quadrupling the signal constellation is only required for the 2 bits/sec/Hz system.

Section 3.2 gives a summary of the channel model, the performance criteria, upper bounds on the bit error rate, and the coding gain limits drawn from the cutoff rate curves. Then, a description of the codes, their analytical and simulated performance over the ideal Rayleigh channel, and complexity comparisons are presented in Section 3.3. Performance results for systems with finite interleaving on a cellular mobile correlated fading model are given in Section 3.4. Some other considerations on the performance of these codes are discussed in Section 3.5. Finally, Section 3.6 summarizes the work in this chapter.

3.2 Background

In this section we briefly review the problem of coded modulation over fading channels.

3.2.1 Channel Model and Performance Criterion

Let x_i be the complex signal input to the channel at time i. The channel introduces two kinds of distortion – multiplicative and additive. The received signal at time i is given by

$$y_i = a_i \cdot x_i + n_i \tag{3.1}$$

where a_i is a Rayleigh distributed random variable with $E(a_i^2) = 1$, and n_i is two dimensional additive white Gaussian noise with single sided power spectral density N_0. We assume that the effect of fading on the transmitted signal phase is completely compensated for by the synchronization circuitry – i.e., perfect coherent detection is assumed.

The trellis encoder produces a sequence of signals $\mathbf{x}_N = (x_1, x_2, \ldots, x_N)$ that are transmitted over the fading channel. At the receiver, the received sequence $\mathbf{y}_N = (y_1, y_2, \ldots, y_N)$ is observed; furthermore, we assume that the fading amplitudes $\mathbf{a}_N = (a_1, a_2, \ldots, a_N)$ can be perfectly recovered at the receiver. The decoder then performs maximum likelihood (ML) decoding using the Viterbi algorithm and decoding metric

$$m(\mathbf{y}_N, \mathbf{x}_N; \mathbf{a}_N) = \sum_{l=1}^{N} \ln P(y_l | x_l, a_l) \tag{3.2}$$

which can be simplified (up to multiplication by a constant) as

$$m(\mathbf{y}_N, \mathbf{x}_N; \mathbf{a}_N) = -\sum_{l=1}^{N} |y_l - a_l x_l|^2. \tag{3.3}$$

The pairwise error probability $P(\mathbf{x}_N \to \hat{\mathbf{x}}_N)$ is the probability that the decoder chooses as its estimate the sequence $\hat{\mathbf{x}}_N = (\hat{x_1}, \hat{x_2}, \ldots, \hat{x}_N)$ when the transmitted sequence was in fact $\mathbf{x}_N = (x_1, x_2, \ldots, x_N)$. This occurs only if

$$\sum_{l=1}^{N} |y_l - a_l \hat{x_l}|^2 < \sum_{l=1}^{N} |y_l - a_l x_l|^2. \tag{3.4}$$

If the fading amplitudes $\{a_l\}$ are independent – i.e., the channel is *fully interleaved* – then a bound on the pairwise error probability can be obtained as in [39]:

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \prod_{l \in \eta} \frac{1}{1 + \frac{1}{4N_o}|x_l - \hat{x}_l|^2} \tag{3.5}$$

where $\eta = \{l : x_l \neq \hat{x}_l\}$. From Eqn. 3.5, it's clear that the pairwise error probability is affected primarily by the symbol-wise Hamming distance between the code sequences; so the error event probability will be dominated by the minimum symbol-wise Hamming distance of the code – its *minimum time diversity*, denoted by L. A parameter of secondary importance is d_p^2, the minimum product distance among code sequences at distance L apart.

Generally, the error performance of a trellis code is computed via the generalized transfer function of the (pairwise) super-state diagram [61]. This diagram has $2^{2\nu}$ states, where ν is the number of memory elements in the encoder; thus this approach is impractical even for codes with few states. However, the codes presented here satisfy the quasi-regularity property of Zehavi and Wolf [13], allowing the use of the transfer function of a modified state diagram with only 2^{ν} states. To describe this modification, define the weight profile function

$$F(B_i, e, D) = \sum_d A_d D^{a^2 d^2[s(c), s(c \oplus e)]} \quad i = 1, 2. \tag{3.6}$$

Here, B_1 and B_2 are the subsets forming the first-level partition of the signal constellation, e is a binary error vector, $s(c)$ is the signal label corresponding to the binary encoder output vector c, A_d is the number of signals with $d^2[s(c), s(c \oplus e)] = d$ and a is the fading attenuation. For codes satisfying the Zehavi-Wolf condition, $F(B_1, e, D) = F(B_2, e, D)$.

Having defined the weight profile function, the branches in the state diagram

are labeled with

$$\frac{I^r \overline{F}(B_1, \tilde{e}, D)}{2^k} \mid_{D=\exp(-\frac{E_s}{4N_o})} \tag{3.7}$$

where $\tilde{e}$ is the label of the signal relative to a transition, r is the number of input "ones" associated with $\tilde{e}$, k is the number of input bits entering the encoder, E_s is the average signal energy, and the overbar denotes averaging over the fading distribution.

The bit error probability of the TCM system can be tightly upper bounded [62] by

$$P_b \leq \frac{1}{k}\left[\frac{1}{2^{2L}}\sum_{j=1}^{L}\begin{pmatrix} 2L-j-1 \\ L-1 \end{pmatrix}\left(\frac{2}{1+x_{min}}\right)^j\right]\frac{\partial}{\partial I}T(D,I) \mid_{I=1} \tag{3.8}$$

where L is the minimum time diversity of the code, $T(D, I)$ is the code transfer function and

$$\begin{aligned} x_{min} &= \sqrt{\frac{d_x^2/4N_o}{1+d_x^2/4N_o}} \\ d_x^2 &= min\{|x_i - x_j|^2, x_i \neq x_j\}. \end{aligned} \tag{3.9}$$

3.3 The Structure of I-Q Trellis Coded Modulation

Recall the structure of the proposed system in Figure 3.1. The basic idea is to use two independent encoders in parallel to select the in-phase and quadrature components of the transmitted signal. Moreover, two independent decoders are used to recover the data associated with the in-phase and quadrature components of the received signal. Hence, this coding scheme is called I-Q TCM.

As an example, suppose we use two rate 1/2 encoders, with the outputs of each mapped to a 4-AM signal set. The result is a 16-QAM constellation for

transmitting 2 bits/sec/Hz. This code structure appears in [60], where it is shown that a 4-state code thus designed has a minimum time diversity of three, while the best 8-state/8-PSK code with a throughput of 2 bits/sec/Hz has a minimum time diversity of two.

Regarding the complexity of the proposed structure: Similar to [61], we measure complexity by the total number of non-parallel paths leaving all the states divided by the number of information bits associated with a transition through the trellis*. Consider a conventional TCM scheme in which $k = k_1 + k_2$ data bits enter the encoder every signaling interval, causing a transition; here, k_1 is the number of "encoded" bits and k_2 is the number of "uncoded" bits – i.e., there are 2^{k_1} non-parallel "bundles" of 2^{k_2} parallel paths leaving each state. If the encoder has ν memory elements, the complexity is $2^{\nu+k_1}/k$.

By comparison, consider the following I-Q TCM schemes:

- If k is even, we split up the k data bits into two blocks of $k/2$, and use two $k/2$-input encoders every signaling interval to encode the in-phase and quadrature components of the transmitted signal; assume $k_1/2$ of the $k/2$ bits are "encoded" and $k_2/2$ are "uncoded". If each of the two encoders has ν memory elements, then there are (in both encoders combined) $2^{\nu+1}$ states with $2^{k_1/2}$ non-parallel edges emanating from each state – thus a complexity of $2^{\nu+1+(k_1/2)}/k$.

- If k is odd we use two k-input encoders every *two* signaling intervals; assuming k_1 of the k bits are "encoded", the result is $2^{\nu+1}$ states with 2^{k_1}

* In [61] the complexity is normalized per transmitted two-dimensional signal; we opt here to normalize per information bit.

non-parallel edges emanating from each state to encode $2k$ data bits – or a complexity of $2^{\nu+1+k_1}/2k = 2^{\nu+k_1}/k$.

It is clear, then, that using two encoders – each with ν memory elements – to implement I-Q TCM is no more complex than using a single encoder with ν memory elements at the same throughput.

The proposed codes are detailed in the following sections.

3.3.1 I-Q QPSK

To transmit 1 bit/sec/Hz, each of the two encoders must encode 0.5 bit/sec/Hz. Hence, each encoder will be rate 1/2 and will operate every two signaling intervals; furthermore, each of the two encoded bits coming out of an encoder will be mapped to a 2-AM signal. The output of the first encoder specifies the in-phase components of two consecutive signals, while the output of the second encoder specifies the quadrature components; thus two QPSK signals are generated every two signaling intervals. This approach dictates the use of convolutional codes optimized in terms of minimum Hamming distance [63]. Table 3.1 shows the generator polynomials and the minimum Hamming distances of the codes thus used. Figure 3.2 shows the trellis diagram of the 8-state code, and Figure 3.3 shows the BER of each code – both the analytical upper bound from Eqn. 3.8 and error rates obtained via simulation. Note that, for the 64 state code, a BER of 10^{-5} can be achieved at $E_b/N_o = 7.5$ dB, while the cutoff rate limit is 5.5 dB. For this code the complexity is equivalent to that of a Gray mapped QPSK encoded system with 64 states.

ν	$g_1(D)$	$g_2(D)$	L
2	$1+D^2$	$1+D+D^2$	5
3	$1+D+D^3$	$1+D+D^2+D^3$	6
4	$1+D^3+D^4$	$1+D+D^2+D^4$	7
5	$1+D+D^3+D^5$	$1+D^2+D^3+D^4+D^5$	8
6	$1+D^2+D^3+D^5+D^6$	$1+D+D^2+D^3+D^6$	10

Table 3.1: Generator polynomials for the rate 1/2 codes.

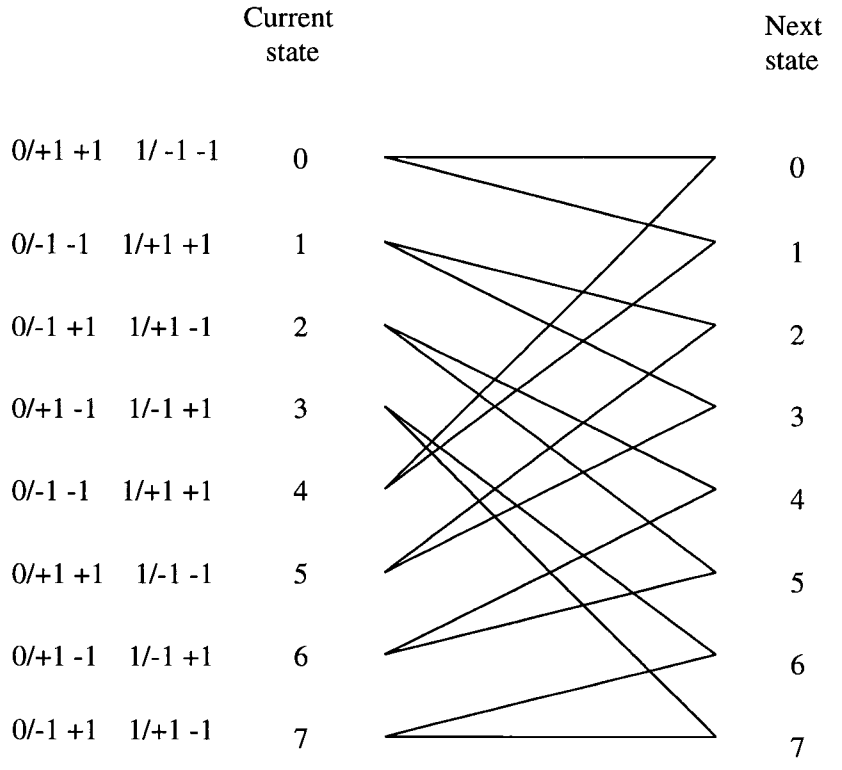


Figure 3.2: Trellis diagram for the 8-state I-Q QPSK code. (1 bit/sec/Hz)

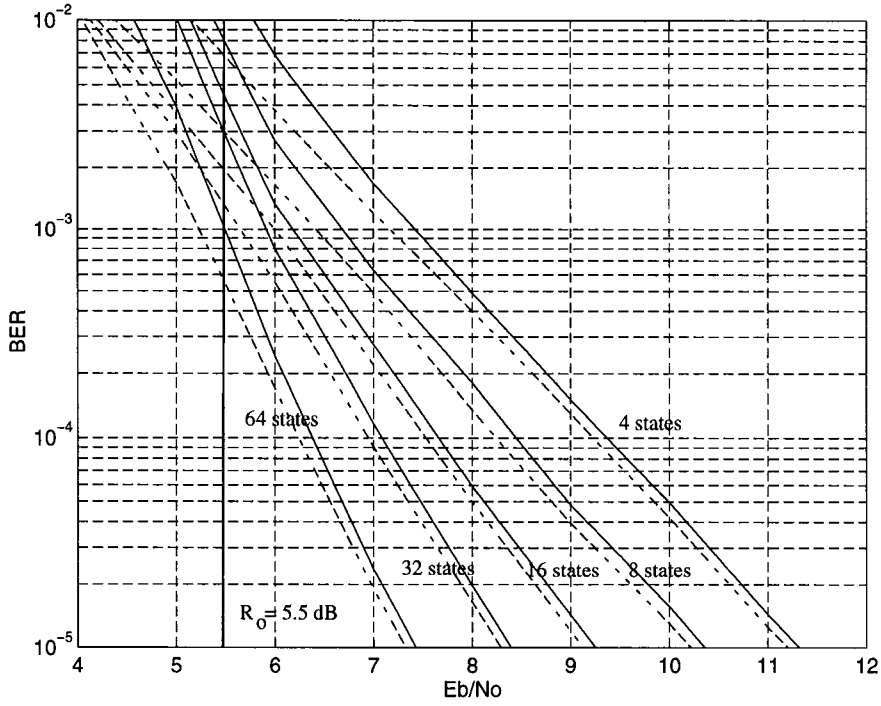


Figure 3.3: BER performance for 1 bit/sec/Hz I-Q QPSK. The solid lines indicate the upper bound from Eqn. 3.8; dashed lines indicate simulation resullts.

3.3.2 I-Q 16-QAM

To transmit 2 bits/sec/Hz, two rate 1/2 encoders are used each signaling interval, with the output of each mapped to a 4-AM signal set; the result is a 16-QAM signal selected every signaling period. Gray mapping between the output bits of each encoder and the 4-AM signal set allows the use of convolutional codes optimized for Hamming distance; once again, the codes in Table 3.1 are used. With this approach, the minimum time diversity is bounded by $L \leq \nu + 1$, where ν is the number of memory elements in each of the two encoders. By comparison, conventional 8-PSK codes sending two bits per baud have a minimum time diversity at most $\lfloor \nu/2 \rfloor + 1$. Figure 3.4 shows the trellis diagram of the 8-state

code with $L = 4$.

Since the I and Q channels operate independently, the bit error performance for I-Q 16-QAM is identical to that of 4-AM. Figure 3.5 shows the BER performance for the different codes. Note that a BER of 10^{-5} can be achieved at $E_b/N_o = 10.5$ dB using the 64-state code; this is close to the cutoff rate limit of 8 dB for 16-QAM signaling when 2 bits/sec/Hz are transmitted.

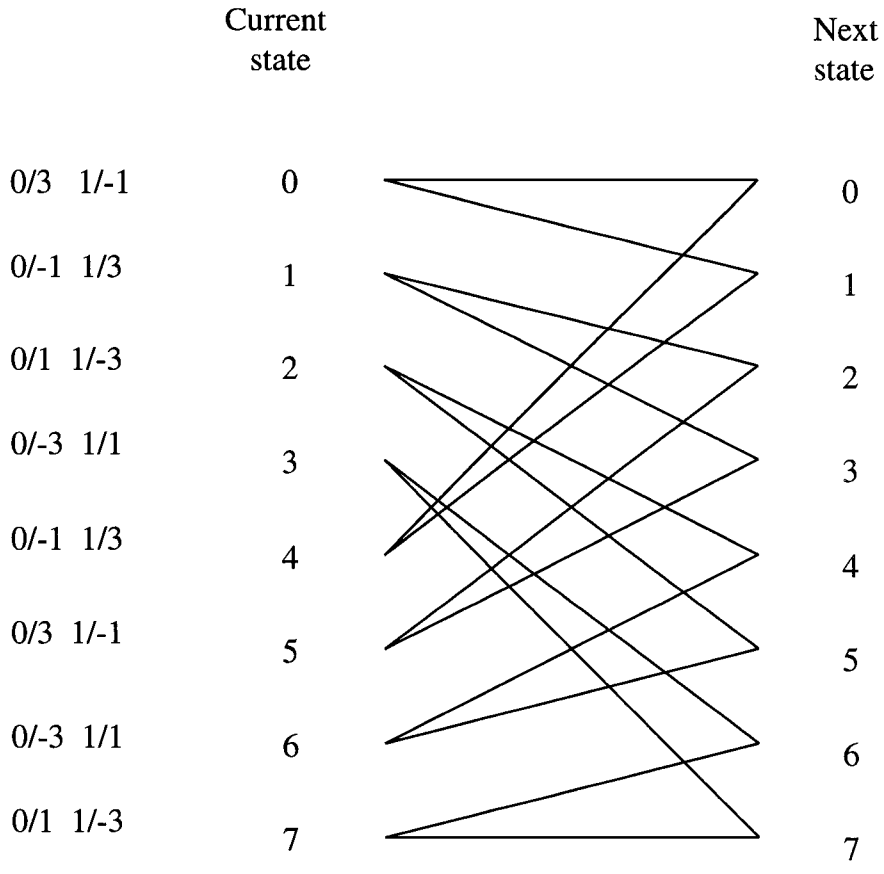


Figure 3.4: Trellis diagram for the 8-state I-Q 16-QAM code. (2 bits/sec/Hz)

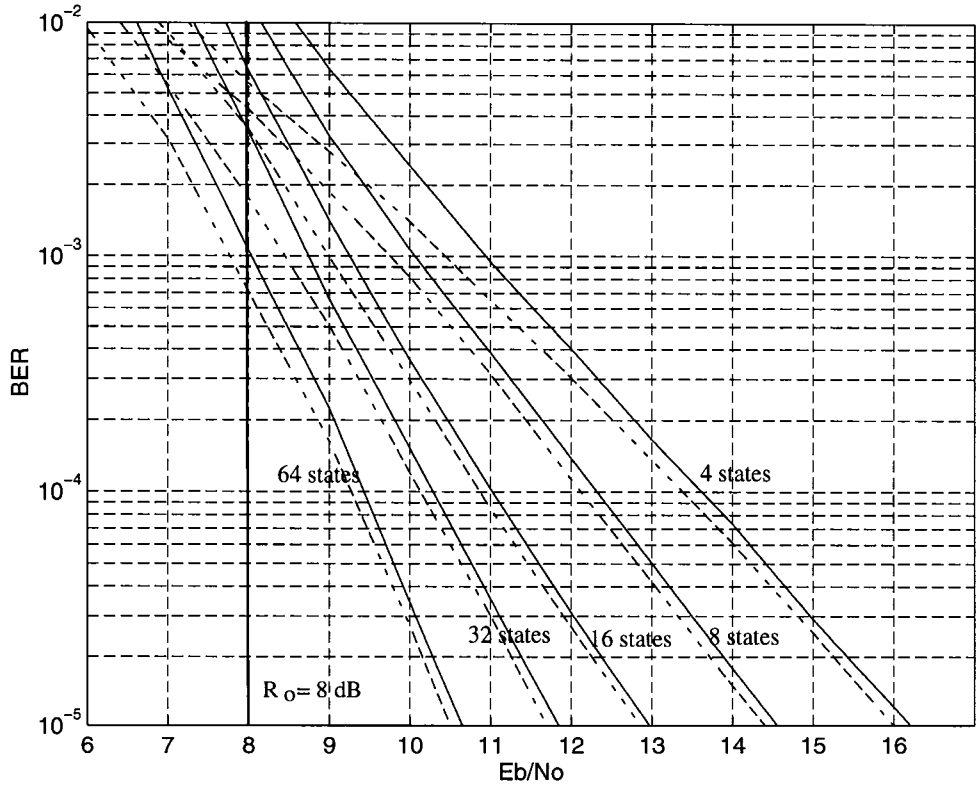


Figure 3.5: BER performance for 2 bits/sec/Hz I-Q 16-QAM. The solid lines indicate the upper bound from Eqn. 3.8; dashed lines indicate simulation resullts.

3.3.3 4-D I-Q 16-QAM

To transmit 3 bits/sec/Hz, two rate 3/4 encoders are used every two signaling intervals. The four bits from each encoder select a pair of 4-AM signals; the output of the first encoder specifies the in-phase components of two consecutive signals, while the output of the second encoder specifies the quadrature components. Thus two 16-QAM signals are generated every two signaling intervals. Figure 3.6 shows the signal partitioning of the 4-AM signal set; this partitioning maximizes the symbol Hamming distance and the minimum squared product

distance. An eight state code's trellis diagram is shown in Figure 3.7. For this code the minimum time diversity is two and the minimum squared product distance (MSPD) is $16(\Delta_0^2)^2$ where Δ_0 is the spacing between the signal symbols in the 16-QAM signal set. Figure 3.8 shows the trellis diagram of a 16-state code. Also, a 32-state code, shown in Figure 3.9, with a minimum time diversity of 4 is designed. Figure 3.10 shows the performance of these three codes. No codes with higher number of states are designed because of their large complexity and the small gain in performance due to the limited minimum time diversity.

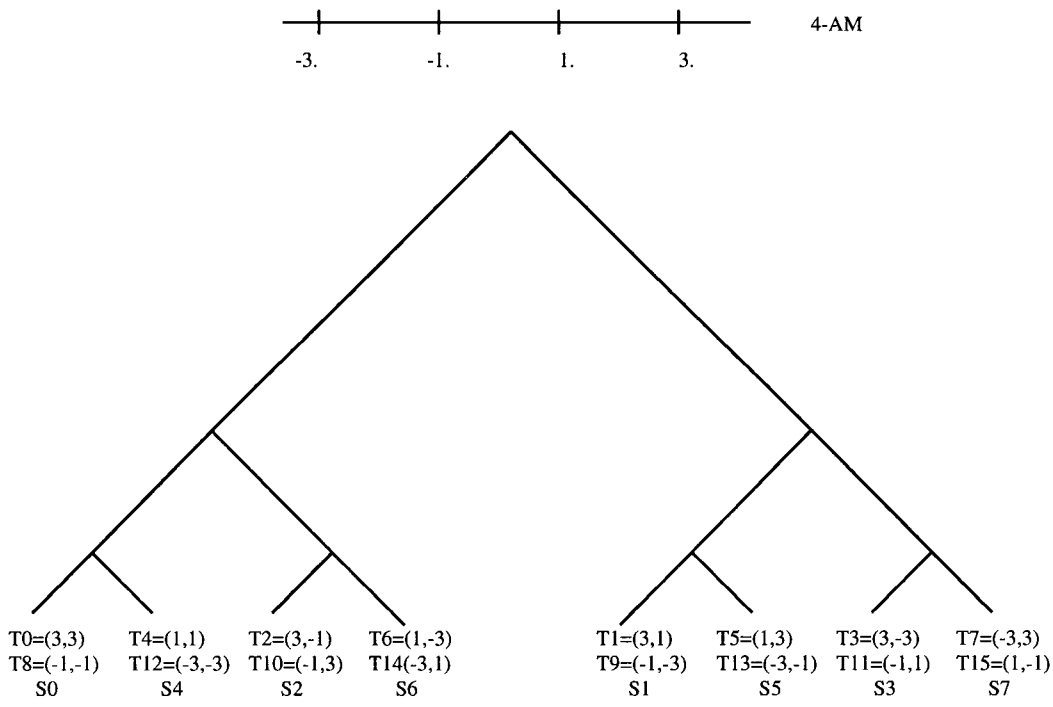


Figure 3.6: Signal partitioning for the 2D 4-AM constellation designed for the Rayleigh distributed channel

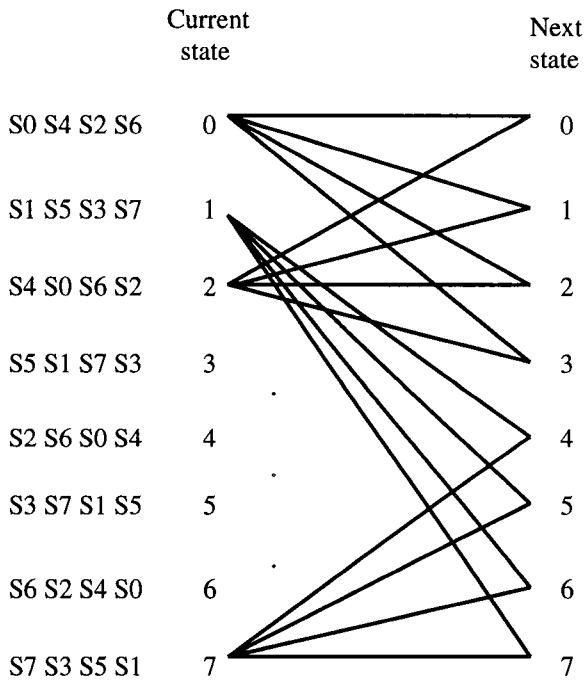


Figure 3.7: Trellis diagram for the 4-D I-Q 16-QAM 8-state code. (3 bits/sec/Hz)

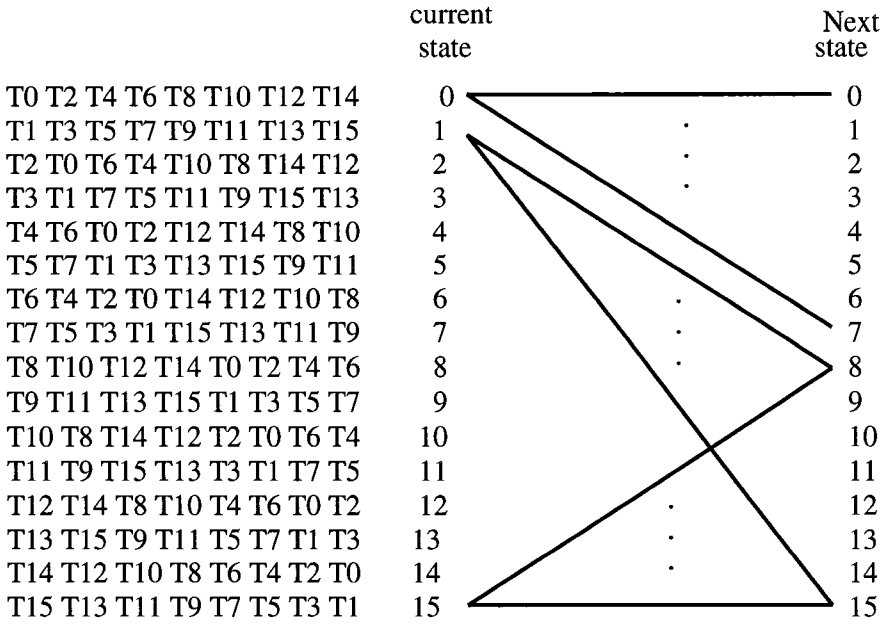


Figure 3.8: Trellis diagram for the 4-D I-Q 16-QAM 16-state code (3 bits/s/Hz)

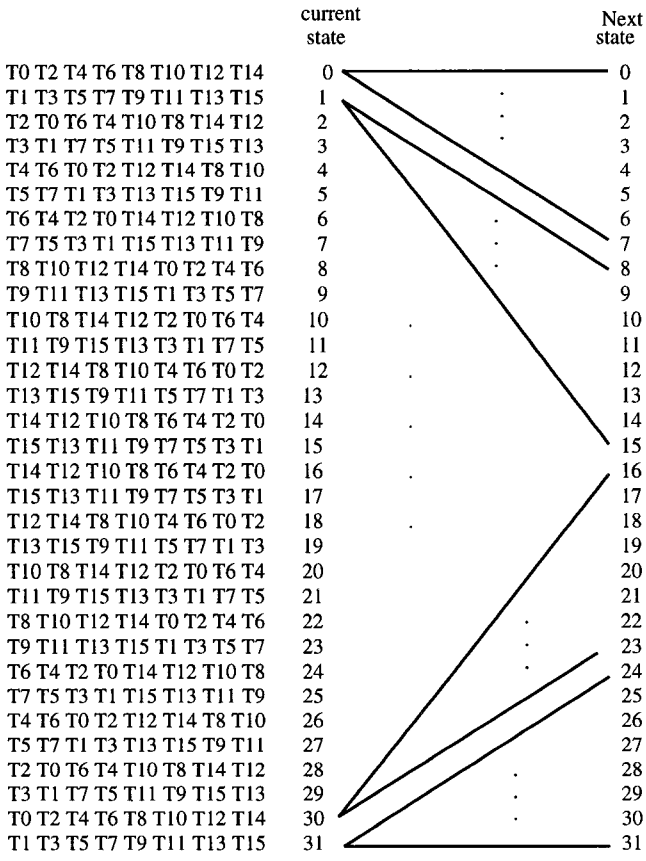


Figure 3.9: Trellis diagram for the 4-D I-Q 16-QAM 32-state code (3 bits/s/Hz)

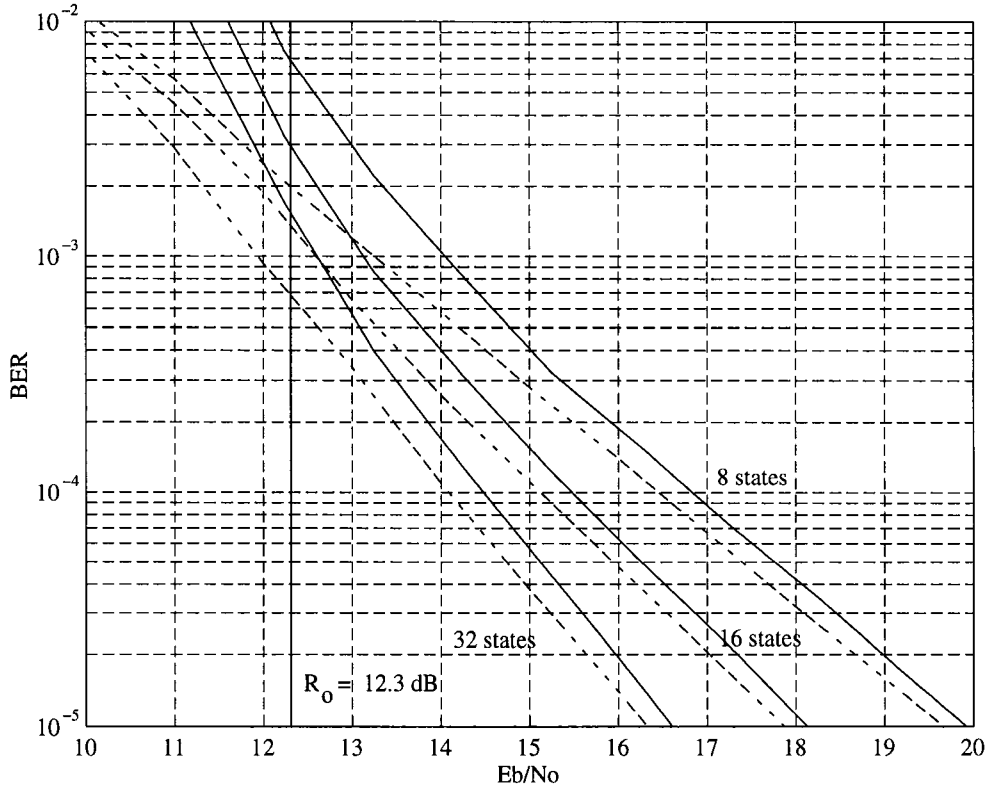


Figure 3.10: BER performance for 3 bits/sec/Hz I-Q 4-D 16-QAM. The solid lines indicate the upper bound from Eqn. 3.8; dashed lines indicate simulation resullts.

3.3.4 Comparisons

Table 3.2 compares the minimum time diversities and the product distances of the proposed codes with those of conventionally-designed TCM schemes.

- The 1 bit/sec/Hz I-Q QPSK system is compared with conventional Gray-mapped QPSK used with a rate 1/2 convolutional code.
- The 2 bits/sec/Hz I-Q 16-QAM system is compared with TCM schemes based on 8-PSK designed by Schlegel and Costello [46].

- The 3 bits/sec/Hz four-dimensional I-Q 16-QAM system is compared with codes employing 16-QAM signaling proposed by Du and Vucetic [49].

In each case a comparison is made between two codes of the same complexity, and in each case the I-Q code has a substantially higher minimum time diversity than the conventional TCM scheme. (The only exception: An eight-state 3 bits/sec/Hz code from [49] is compared with an eight-state 3 bits/sec/Hz I-Q TCM code; the time diversities are identical, but the complexity of the I-Q TCM code is one-half that of the code from [49]. In addition, the MSPD of the I-Q TCM code is $16(\Delta_0^2)^2$ while the MSPD of the code from [49] is $5(\Delta_0^2)^2$). Figures 3.11, 3.12 and 3.13 show the results of simulations that compare the I-Q TCM codes with conventional codes. The coding gains at a BER of 10^{-5} range from 2 dB (for I-Q QPSK) to more than 5 dB (for 4-D I-Q 16-QAM).

Moreover, a comparison between the I-Q and the traditionally designed codes' minimum free squared Euclidean distances are also shown in Table 3.2. As mentioned in Chapter 1, the primary design goal in AWGN non-fading channels is to maximize the code's minimum free squared Euclidean distance d_{free}^2. To make a fair comparison, d_{free}^2 and d_p^2 are normalized by the bit energy E_b. Also, the values of d_{free}^2/E_b for Ungerboeck codes (i.e., the codes optimized for the AWGN non-fading channel) are included in parentheses. The I-Q QPSK codes achieve the same d_{free}^2/E_b values. The I-Q 16-QAM codes have little degradation in d_{free}^2/E_b, while the loss in d_{free}^2/E_b for the 4-D I-Q 16-QAM codes is more evident.

Code	bits/s/Hz	ν	L	$d_p^2/(E_b)^L$	Complexity	d_{free}^2/E_b
QPSK-Gray	1	2	3	32	8	10 $(10)^8$
I-Q QPSK	1	2	5	32	8	10
QPSK-Gray	1	3	4	64	16	12 (12)
I-Q QPSK	1	3	6	64	16	12
QPSK-Gray	1	4	5	128	32	14 (14)
I-Q QPSK	1	4	7	128	32	14
QPSK-Gray	1	5	6	256	64	16 (16)
I-Q QPSK	1	5	8	256	64	16
QPSK-Gray	1	6	7	1024	128	20 (20)
I-Q QPSK	1	6	10	1024	128	20
8-PSK	2	2	2	9.38	8	6.34 (8)
I-Q 16-QAM	2	2	3	8.19	8	7.2
8-PSK	2	3	2	32	16	9.2 (9.2)
I-Q 16-QAM	2	3	4	6.55	16	8
8-PSK	2	4	3	37.44	32	10.3 (10.3)
I-Q 16-QAM	2	4	5	5.24	32	8.8
8-PSK	2	5	3	128	64	10.3 (11.5)
I-Q 16-QAM	2	5	6	4.19	64	9.6
8-PSK	2	6	4	128	128	8.68 (12.7)
I-Q 16-QAM	2	6	7	3.36	128	11.2
16-QAM	3	3	2	7.2	64/3	4.8 (6)
I-Q 16-QAM	3	3	2	23.04	32/3	3.6
16-QAM	3	4	2	14.4	128/3	6 (7.2)
I-Q 16-QAM	3	4	3	6.91	128/3	4.8
16-QAM	3	5	2	46.08	256/3	7.2 (7.2)
I-Q 16-QAM	3	5	4	8.29	256/3	6

Table 3.2: A comparison between the proposed codes and conventional TCM schemes. The numbers inside parentheses correspond to Ungerboeck codes.

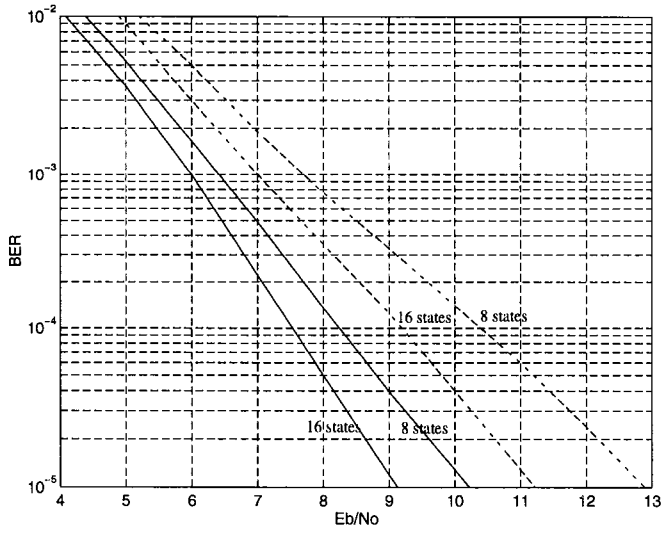


Figure 3.11: A comparison between the proposed 1 bit/sec/Hz code and Gray-mapped rate 1/2 convolutional coding. The solid lines indicate I-Q TCM; dashed lines indicate Gray-mapped convolutional codes.

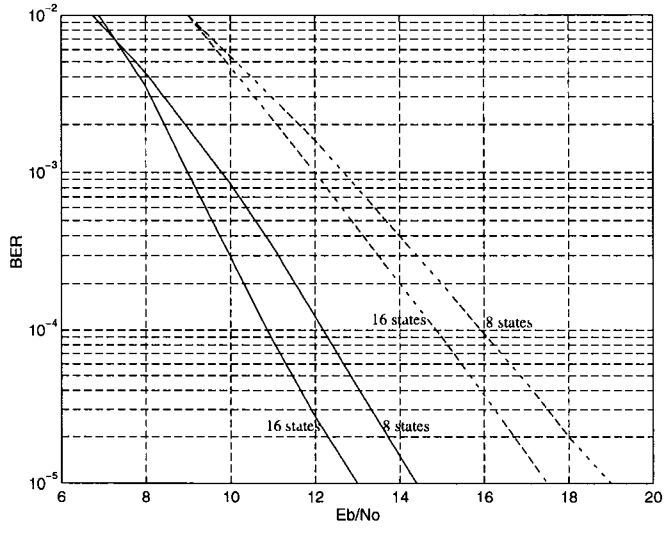


Figure 3.12: A comparison between the proposed 2 bit/sec/Hz codes and codes from [46]. The solid lines indicate I-Q TCM; dashed lines indicate [46].

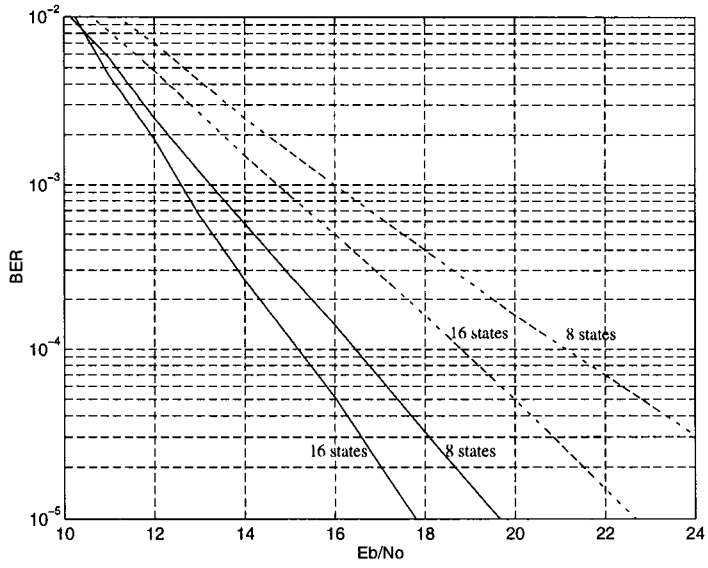


Figure 3.13: A comparison between the proposed 3 bit/sec/Hz code and codes from [49]. The solid lines indicate I-Q TCM; dashed lines indicate [49].

3.4 Performance over Correlated Rayleigh Fading

Often, successive fading amplitudes are correlated – as in, for example, narrowband cellular mobile radio [31]. In this section we briefly describe the performance of the proposed codes when correlated fading is assumed and appropriate interleaving is employed.

We assume an omnidirectional antenna at the mobile and that the received plane waves are uniformly distributed in arrival angle. In this case, the power spectral density of the faded amplitude due to Doppler shift is expressed as [31]

$$V(f) = \begin{cases} \frac{1}{2\pi\sqrt{{f_D}^2 - f^2}} & \text{if } |f| < |f_D| \\ 0 & \text{otherwise.} \end{cases} \tag{3.10}$$

Thus the autocorrelation function $R_H(0;\Delta t)$ can be written as

$$R_H(0;\Delta t) = J_0(2\pi f_D \Delta t) \tag{3.11}$$

where $J_0(x) = \sum_{n=0}^{\infty}(-1)^2(\frac{x^n}{2^n n!})^2$ is the zero-order Bessel function of the first kind. In the simulations, the fade amplitudes are generated by applying two independent white noise sources to two filters with a spectrum shape of $\sqrt{V(f)}$ and scaling them so that $E(a_i^2) = 1$; an FIR filter with 200 coefficients is used to approximate the power spectrum. The filter bandwidth is specified by the vehicle speed, the carrier frequency and the symbol rate; i.e., $f_D = \frac{vf_c}{c}$. In our simulations, we have assumed a vehicle speed of 60 mi/hr, a carrier frequency $f_c = 900$ MHz and a symbol rate of 8000 symbols/sec. This results in a fade rate $f_D T_s = 0.01$.

A block interleaver of 400 symbols was used, corresponding to a delay of 50 msec. The interleaver is implemented as a buffer with d rows (depth) and s columns (span); the encoded symbols are written in successive rows and transmitted over the channel in columns. It has been shown [64] that an interleaver depth of at least one-quarter of the fade cycle provides near-optimal performance – i.e., $d \geq 1/(4f_D T_s)$. For our purposes, this meant we kept the interleaver depth to at least 25 symbols; accordingly, the span s must satisfy $s \leq 16$.

Computer simulations were performed to determine a suitable interleaver span. Ideally, the interleaver span should be on the order of the decoder buffer size so that correlated amplitudes will not affect the decoding of the current symbol. Figure 3.14 shows BER results for the 4-state I-Q 16-QAM encoded system with independent fading and different decoder buffer sizes. Note that there is (relatively) little performance degradation when a buffer of size 4ν is used, compared with the case 8ν. A similar conclusion has been observed for

other codes. Simulation results for a fixed interleaver size ($s \cdot d = 400$) but different values of d and s are shown in Figure 3.15; it is observed that $s = 4\nu$ yields the best results.

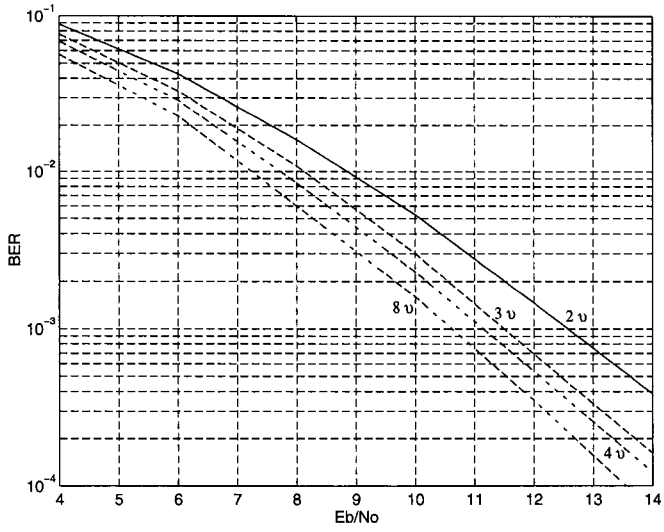


Figure 3.14: BER for the 4-state I-Q 16-QAM encoded system with different decoder buffer sizes.

Figure 3.16 shows the BER results for 8-state codes with different throughputs. For the 2 bits/sec/Hz code an interleaver of dimension $d \times 4\nu = 33 \times 12$ was used. For the other two codes – with their multiplicity of two baud per branch – the guideline of 4ν stages in the decoder buffer would suggest an interleaver with span $s = 24$; however, that would imply an interleaver depth of $d = 16$, which is less than one-quarter of the fade cycle. As a compromise, we used a 20×20 interleaver table for the 1 bit/sec/Hz code and the 3 bits/sec/Hz code. A loss of about 1-2 dB is observed when these results are compared with the independent Rayleigh fading results in Figures 3.3, 3.5, and 3.10. This loss is due to the effects of residual correlated fading.

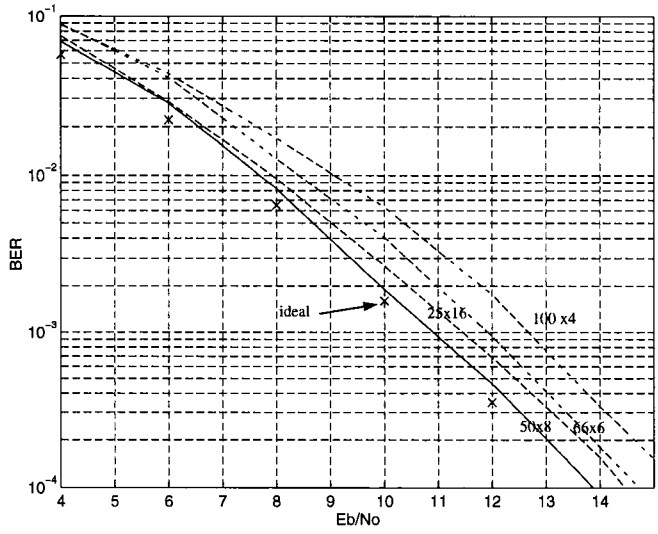


Figure 3.15: BER for the 4-state encoded I-Q 16-QAM with different interleaver dimensions.

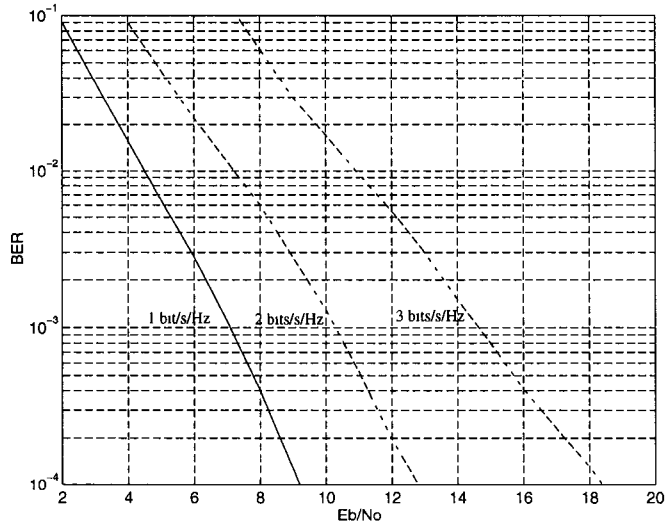


Figure 3.16: BER for the interleaved 8-state encoded systems with different throughputs.

3.5 Other Considerations

In this section we consider the performance of some I-Q codes over the Rician channel where a line-of-sight is assumed to exist. Also, the effect of channel state information at the decoder is addressed.

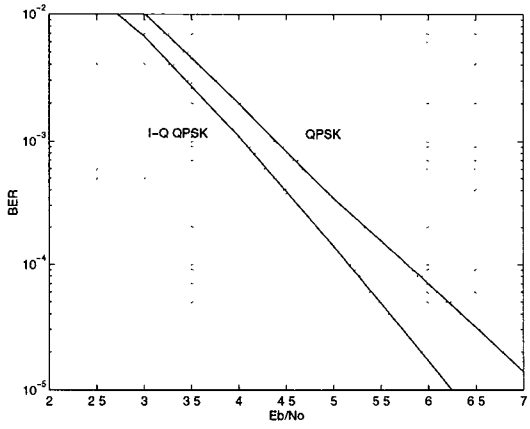
3.5.1 Performance over Rician Channels

Previously, we assumed a Rayleigh distribution for the channel fade amplitudes. This model assumes no effective line-of-sight, and the resulting primary design criterion is to maximize the minimum time diversity of the code.

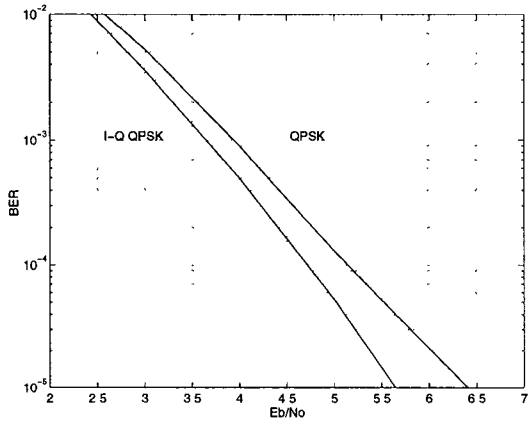
In some environments a line-of-sight exists. A typical example is the mobile satellite channel. As mentioned in Chapter 2, in this case the fade amplitude will have a Rician distribution, and the minimum time diversity of the code will not be the dominant design criterion. In this section we compare the performances of I-Q TCM and traditionally designed codes when the channel is modeled to have the Rician distribution.

We simulated two 16-state schemes: I-Q QPSK and Gray mapped QPSK. Both schemes assume full interleaving and ideal channel state information. Rician parameter values of $K = 5$ and $K = 9$ and are used - i.e., the line-of-sight component is 7 and 9.5 dB above the diffuse component, respectively. The I-Q QPSK code and the "traditional" convolutionally-encoded Gray-mapped QPSK have the same minimum squared Euclidean distance. Thus their performance in AWGN channels are very similar. Figure 3.17 shows the BER performance of the two codes in both channel environments. For the $K = 9$ case, the I-Q code provided about 0.8 dB coding gain over the conventional code at an output

BER of 10^{-5}. At an output BER of 10^{-3} the coding gain was about 0.4 dB. For the $K = 5$ case, there was a very slight increase in coding gain – the gains were about 1.0 dB and 0.4 dB at 10^{-5} and 10^{-3}, respectively. Even though the I-Q code is designed for the Rayleigh channel, slight improvements can be obtained if a Rician distribution exists.



(a)



(b)

Figure 3.17: A comparison between the I-Q and the Gray-mapped QPSK 16-state codes with different Rician factors (a) $K = 5$ (b) $K = 9$

3.5.2 Effect of Channel State Information

Although there exists powerful methods to estimate the channel state information very accurately, here we provide some results where no channel state information is used at the decoder. In this case the decoding metric will be

$$m(\mathbf{y}_N, \mathbf{x}_N) = \sum_{l=1}^{N} \ln P(y_l|x_l) = \sum_{l=1}^{N} -|y_l - x_l|^2. \tag{3.12}$$

For constant amplitude signals, such as BPSK, QPSK or 8-PSK, the absence of channel state information results in a loss of 1.5-2.5 dB. However, for multi-amplitude signals such as 16-QAM, a significant loss occurs if no channel state information is used at the receiver. In fact, our simulations show that a high error floor $\sim 10^{-2}$ occurs if no form of CSI is used at the receiver. (By comparison, Ho *et. al* used pilot-symbol techniques to decode 16-QAM signals with a loss of only 1-1.5 dB from ideal CSI [60].)

Figure 3.18 show the performance of the 16-state I-Q QPSK schemes over the Rayleigh and Rician ($K = 9$) channels. In the Rayleigh channel case, about 1.3 dB (at BER $=10^{-5}$) is lost if no CSI is used. For the Rician case, the use of CSI at the decoder gives only about .2 dB additional coding gain. This is because the line-of-sight component is 9.5 dB above the diffuse component and hence the channel is close to the AWGN non-fading channel. Finally, a comparison between the 16-state I-Q QPSK code and the Gray mapped QPSK with a 16-state rate 1/2 convolutional code over both the Rayleigh and Rician ($K = 9$) channels is shown in Figure 3.19. In both systems *no CSI* is used at the decoder. It is clear that still 2 dB of extra coding gain (at BER $= 10^{-5}$) is achieved for the Rayleigh channel, while about 0.7 dB of additional coding gain is achieved over the Rician channel with $K = 9$.

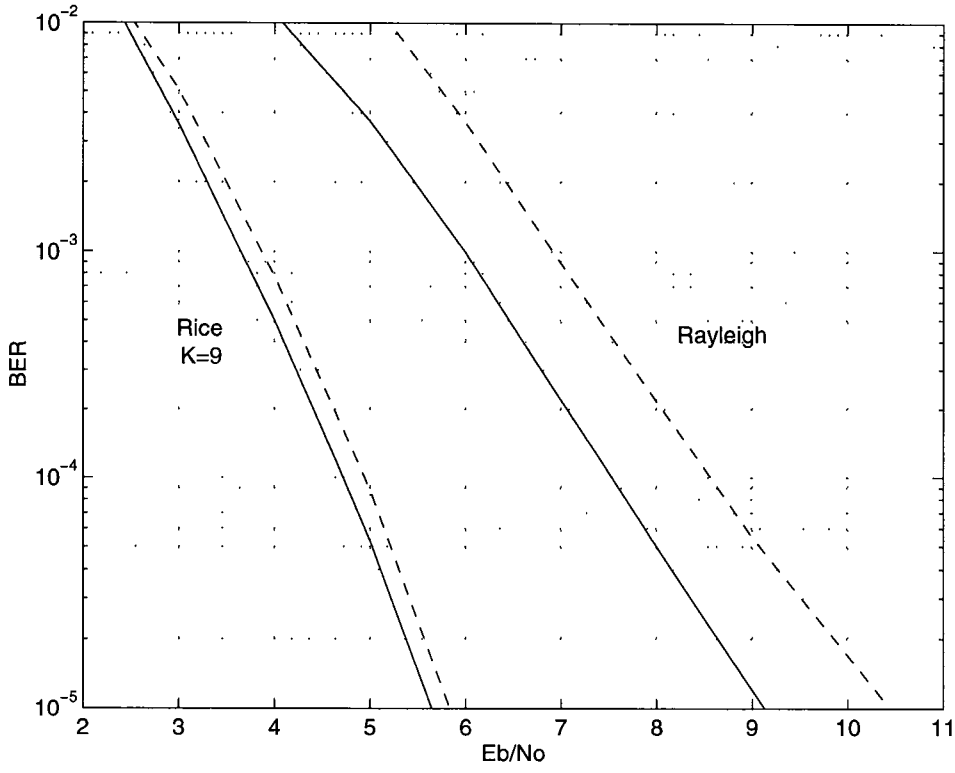


Figure 3.18: BER for the 16-state I-Q QPSK system over both the Rayleigh and Rician ($K = 9$) channels. Solid = Ideal CSI, dashed = NO CSI.

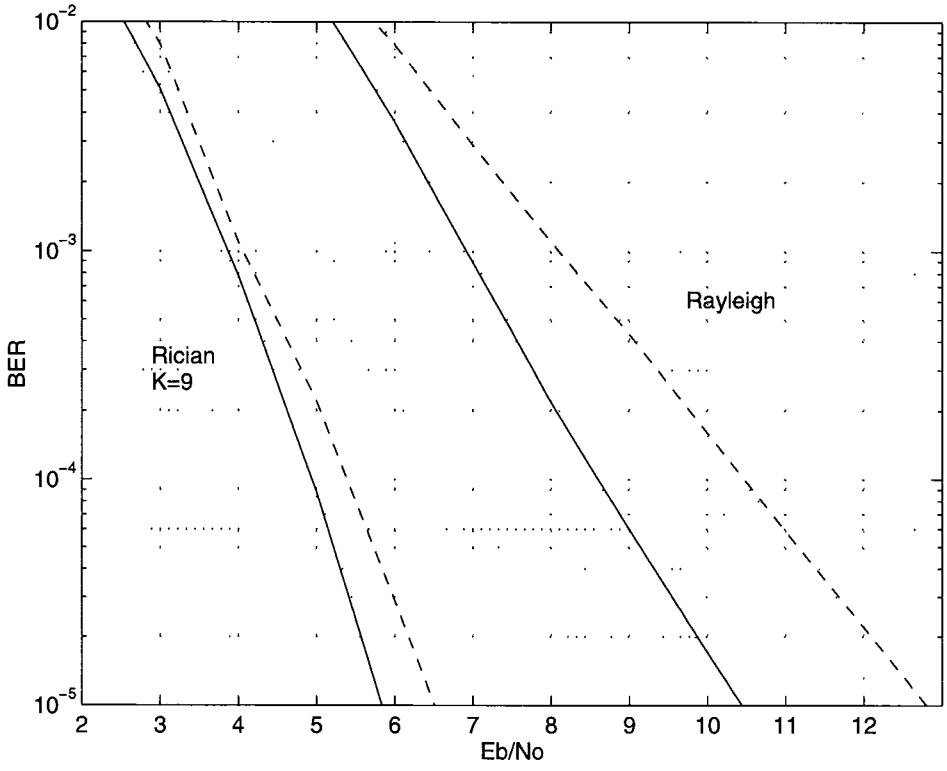


Figure 3.19: BER comparison between the 16-state I-Q QPSK code and the Gray mapped QPSK with a 16-state rate 1/2 convolutional code over both the Rayleigh and Rician ($K = 9$) channels. All curves assume NO CSI. Solid = I-Q code, dashed = rate 1/2 conv. encoded QPSK.

3.6 Summary

I-Q TCM schemes with bandwidth efficiencies of 1,2, and 3 bits/sec/Hz were presented. Computer simulations and tight upper bounds show a significant improvement in bit error rate for I-Q TCM schemes relative to conventional trellis codes over the Rayleigh fading channel; coding gains close to what is expected from the cutoff rate limits are achieved for codes with moderate complexity. Simulations of interleaved systems confirm that an interleaver depth of $1/(4f_DT_s)$ yields good performance, and it has been shown that a suitable value of the interleaver span is 4ν. Finally, it was shown that I-Q TCM yields (reduced) coding gains over conventional codes over Rician channels; also, the effect of no CSI was evaluated. In the next chapter we will investigate the use of combined coding and diversity to further improve the error performance.

Chapter 4

Diversity Reception

4.1 Introduction

Diversity combining is a well-known and effective method for improving the performance of digital communication systems over fading channels [7, 31, 65, 66]. The basic principle of M-fold diversity is to use M independent channels so that the probability of a "deep fade" on *all* channels is low. (These independent channels can be created in a number of ways, including frequency, time, and/or polarization diversity; if multiple antennas are used to receive multiple independent versions of the received signal, the approach is called *spatial* diversity.) A combining circuit is used to form a single resultant signal from the M different "branch" signals. There are (at least) three different methods for combining.

- The optimal combining scheme is called *maximal ratio combining* (MRC) [31]. In such a scheme the M received K-dimensional signals are treated as a single MK-dimensional signal; the resultant SNR at the output of the combiner is the sum of the SNR's of the M branches.

- In *equal gain combining* (EGC) the resultant signal is simply the sum of the signals from the M branches.

- In *selection combining* (SC) the resultant signal is the one with highest SNR among the M received signals; in practice, the signal with the strongest received signal – i.e., signal plus noise – is selected.

Error probability expressions for uncoded systems with different combining schemes in Rayleigh fading are presented in [67, 68]

As shown in the previous chapter, trellis-coded modulation (TCM) also provides a form of diversity – *time diversity.* The performance of TCM schemes may be evaluated using the pairwise error probability and the transfer function approach [13, 61]. Upper bounds on the bit error probability for TCM with different combining schemes were presented in [69, 70]. These expressions use the Chernoff bound to establish an upper limit on the pairwise error probability and are loose. A tighter upper bound on the bit error probability for maximal ratio combining (only) has recently been developed by Ventura-Traveset, Caire, Biglieri, and Taricco [71]; the upper bound in [71] requires numerical evaluation of the pairwise error probability and the use of a truncated transfer function which does not account for all paths in the trellis.

In uncoded systems, the diversity combining schemes are compared based on the probability distribution of the signal-to-noise power ratio (SNR) at the output of the combiner [31, 66, 68]. On the other hand, the typical performance criterion for systems employing coded modulation is the cutoff rate parameter (R_o). In this Chapter, the use of trellis-coded systems with diversity receivers is investigated. Cutoff rate expressions for the Rayleigh distributed channel with diversity reception and the three different combining schemes are presented.

Also, tight upper bounds on the pairwise error probability for transmission over the Rayleigh distributed channel with diversity reception and the three combining schemes are derived. The same system configurations in [69, 70] are used; the new bounds are shown to be tighter than those presented in [69, 70].

The next section describes the system model and the combining metrics. In Section 4.3, expressions for the cutoff rates for the three combining schemes are derived and compared. Tight upper bounds on the bit error probability for trellis coded systems with the three combining schemes are derived and analyzed in Section 4.4. Moreover, the application of these bounds on uncoded systems is discussed in section 4.5. The effect of branch correlation on MRC schemes is presented in Section 4.6. Comparisons complexity tradeoffs are presented in Section 4.7. Finally, the Chapter is concluded with a summary.

4.2 System Model

The underlying system can be described as follows. Suppose the complex signal x_i is transmitted at time i and M corresponding signals $\underline{y}_i = \{y_{i,1}, y_{i,2}, \ldots, y_{i,M}\}$ are received; i.e.,

$$\begin{aligned} y_{i,1} &= a_{i,1}x_i + n_{i,1} \\ y_{i,2} &= a_{i,2}x_i + n_{i,2} \\ &\vdots \\ y_{i,M} &= a_{i,M}x_i + n_{i,M} \end{aligned} \tag{4.1}$$

where $\underline{a}_i = \{a_{i,1}, a_{i,2}, \ldots, a_{i,M}\}$ are the fading amplitudes, assumed to be Rayleigh-distributed with $E(a_i^2) = 1$. Here also, $\underline{n}_i = \{n_{i,1}, n_{i,2}, \ldots, n_{i,M}\}$ are the two-

dimensional additive white Gaussian noise samples with one-sided power spectral density N_o. We assume full interleaving. We also assume the branches are affected by independent fading. Thus, $\{a_{i,j}\}$ are i.i.d. Rayleigh.

The transmitter produces a sequence of signals $\mathbf{x}_N = \{x_1, x_2, \ldots, x_N\}$. At the receiver, the sequence of received M-tuples $\mathbf{y}_N = \{\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_N\}$ and the channel fade amplitudes $\mathbf{a}_N = \{\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_N\}$, are the inputs to the TCM decoder which performs maximum likelihood (ML) decoding assuming ideal channel state information. The decoding metric is

$$m(\mathbf{x}_N, \mathbf{y}_N; \mathbf{a}_N) = \sum_{i=1}^{N} m(x_i, \underline{y}_i; \underline{a}_i) \tag{4.2}$$

Here the symbol metric $m(x_i, \underline{y}_i; \underline{a}_i)$ depends on which form of signal combining is being used.

- For maximal ratio combining,

$$m(x_i, \underline{y}_i; \underline{a}_i) = -\sum_{l=1}^{M} |y_{i,l} - a_{i,l} x_i|^2.$$

- For equal gain combining,

$$m(x_i, \underline{y}_i; \underline{a}_i) = -\left| \sum_{l=1}^{M} (y_{i,l} - a_{i,l} x_i) \right|^2.$$

- For selection combining,

$$m(x_i, \underline{y}_i; \underline{a}_i) = -|y_{i,j^\star} - a_{i,j^\star} x_i|^2$$

where

$$j^\star = \arg\max\{a_{i,j}, j = 1, \ldots, M\}.$$

4.3 Cutoff Rate for Diversity Reception

Recall that the pairwise error probability, $P(\mathbf{x}_N \rightarrow \hat{\mathbf{x}}_N)$, is the conditional probability that the metric associated with the coded sequence $\hat{\mathbf{x}}_N$ exceeds that of $\mathbf{x}_N$, given $\mathbf{x}_N$ was in fact transmitted. It can be upper bounded using the Chernoff bound as follows

$$\begin{aligned} P(\mathbf{x}_N \rightarrow \hat{\mathbf{x}}_N) &= P(m(\hat{\mathbf{x}}_N, \mathbf{y}_N; \mathbf{a}_N) - m(\mathbf{x}_N, \mathbf{y}_N; \mathbf{a}_N) \geq 0) \\ &\leq E[\exp(\lambda\{m(\hat{\mathbf{x}}_N, \mathbf{y}_N; \mathbf{a}_N) - m(\mathbf{x}_N, \mathbf{y}_N; \mathbf{a}_N)\})] \\ &= C(\mathbf{x}_N, \hat{\mathbf{x}}_N, \lambda) = \prod_{i=1}^{N} C(x_i, \hat{x}_i, \lambda) \end{aligned} \tag{4.3}$$

where

$$C(x_i, \hat{x}_i, \lambda) = E[\exp(\lambda\{m(\hat{x}_i, \underline{y}_i; \underline{a}_i) - m(x_i, \underline{y}_i; \underline{a}_i)\})], \tag{4.4}$$

and the expectation is taken with respect to the noise $\underline{n}_i$ and the fading $\underline{a}_i$. As mentioned previously the cutoff rate, R_o, in bits/transmitted signal can be expressed as [46]

$$R_o = 2\log_2(|A|) - \log_2\left(\sum_{x_i \in A}\sum_{\hat{x}_i \in A} C(x_i, \hat{x}_i)\right) \tag{4.5}$$

where A is the signal set and

$$C(x_i, \hat{x}_i) = \min_{\lambda} C(x_i, \hat{x}_i, \lambda). \tag{4.6}$$

In the case of single channel reception ($M = 1$) the Chernoff factor can be written in a simplified form as

$$C(x_i, \hat{x}_i) = \frac{1}{1 + \frac{|x_i - \hat{x}_i|^2}{4N_o}}. \tag{4.7}$$

In the next subsections the Chernoff bound is derived for M-diversity systems with the three different types of metrics. The cutoff rate expressions are also presented and compared.

4.3.1 Maximal Ratio Combiner

For maximal ratio combining, the conditional Chernoff bound can be expressed as

$$\begin{aligned} C(x_i, \hat{x}_i, \lambda|\underline{a}_i) &= E_{\underline{n}_i}[\exp(\lambda\{m(\hat{x}_i, \underline{y}_i; \underline{a}_i) - m(x_i, \underline{y}_i; \underline{a}_i)\})] \\ &= E_{\underline{n}_i}[\exp(\lambda\{\textstyle\sum_{l=1}^{M}(|y_{i,l} - a_{i,l}x_i|^2 - |y_{i,l} - a_{i,l}\hat{x}_i|^2)\})] \end{aligned} \tag{4.8}$$

which can be simplified as follows

$$\begin{aligned} C(x_i, \hat{x}_i, \lambda|\underline{a}_i) &= \textstyle\prod_{l=1}^{M} E_{n_{il}}[\exp(\lambda\{|y_{i,l} - a_{i,l}x_i|^2 - |y_{i,l} - a_{i,l}\hat{x}_i|^2\})] \\ &= \textstyle\prod_{l=1}^{M} E_{n_{i,l}}[\exp(\lambda\{-a_{i,l}^2|x_i - \hat{x}_i|^2 - 2a_{i,l}\Re(n_{i,l} \cdot (x_i - \hat{x}_i)^*)\})] \\ &= \exp(-\lambda \textstyle\sum_{l=1}^{M} a_{i,l}^2|x_i - \hat{x}_i|^2)\times \\ &\quad \textstyle\prod_{l=1}^{M} E_{n_{i,l}}[\exp(\lambda\{-2a_{i,l}\Re(n_{i,l} \cdot (x_i - \hat{x}_i)^*)\})], \end{aligned} \tag{4.9}$$

where $\Re(\cdot)$ denotes the real part and $(\cdot)^*$ denotes complex conjugate. However,

$$E_{n_{i,l}}[\exp(\lambda\{-2a_{i,l}\Re(n_{i,l} \cdot (x_i - \hat{x}_i)^*)\})] = \exp(\lambda^2 a_{i,l}^2 N_o|x_i - \hat{x}_i|^2) \tag{4.10}$$

Therefore, the conditional Chernoff factor can be reduced to

$$C(x_i, \hat{x}_i, \lambda|\underline{a}_i) = \exp(-\lambda \sum_{l=1}^{M} a_{i,l}^2(1 - N_o\lambda)|x_i - \hat{x}_i|^2). \tag{4.11}$$

It is easy to see that $C(x_i, \hat{x}_i, \lambda|\underline{a}_i)$ can be minimized by choosing $\lambda = \frac{1}{2N_o}$. Therefore, the tightest conditional Chernoff bound can be written as

$$C(x_i\hat{x}_i|\underline{a}_i) = \min_{\lambda}\{C(x_i, \hat{x}_i, \lambda|\underline{a}_i)\} = \exp(-\beta\frac{|x_i - \hat{x}_i|^2}{4N_o}) \tag{4.12}$$

where

$$\beta = \sum_{l=1}^{M} a_{i,l}^2. \tag{4.13}$$

Since the $\{a_{i,1}, a_{i,2} \ldots a_{i,M}\}$ are independent Rayleigh distributed random variables, then $\{a_{i,1}^2, a_{i,2}^2 \ldots a_{i,M}^2\}$ will be independent exponentially distributed random variables. So, β will have an M-Erlang distribution with parameter one;

i.e., its distribution is given by

$$f(\beta) = \frac{1}{(M-1)!}\beta^{M-1}e^{-\beta}, \quad \beta \geq 0. \tag{4.14}$$

The last step is to perform the following intergration

$$C(x_i, \hat{x}_i) = \int_0^\infty \frac{1}{(M-1)!}\beta^{M-1}\exp(-\beta)\exp(-\beta\frac{|x_i - \hat{x}_i|^2}{4N_o})d\beta. \tag{4.15}$$

Using the following equality

$$\int_0^\infty x^{M-1}e^{-x}dx = (M-1)! \tag{4.16}$$

the integral reduces to

$$C(x_i, \hat{x}_i) = \frac{1}{\left(1 + \frac{|x_i - \hat{x}_i|^2}{4N_o}\right)^M}. \tag{4.17}$$

Therefore, the cutoff rate for maximal ratio combining receivers is given by

$$R_o = 2\log_2(|A|) - \log_2\left(\sum_{x_i \in A}\sum_{x_j \in A}\frac{1}{\left(1 + \frac{|x_i - \hat{x}_i|^2}{4N_o}\right)^M}\right). \tag{4.18}$$

Figure 4.1 shows the cutoff rate values of the 16-QAM signal constellation and maximal ratio combining with diversity orders of $M = 1, 2, 3, 4$. It is clear that the largest incremental gain is obtained in going from single to double diversity. The coding gains diminish as the order of diversity increases. For example, the curves show that reliable communication at a rate of 2 bits/symbol can be achieved at $E_s/N_o = 11$ dB (or $E_b/N_o = 8$ dB) for single channel reception. However, the required SNR can be reduced to $E_b/N_o = 6.1$ dB (or $E_b/N_o = 3.1$ dB) if double diversity with maximal ratio combining is used.

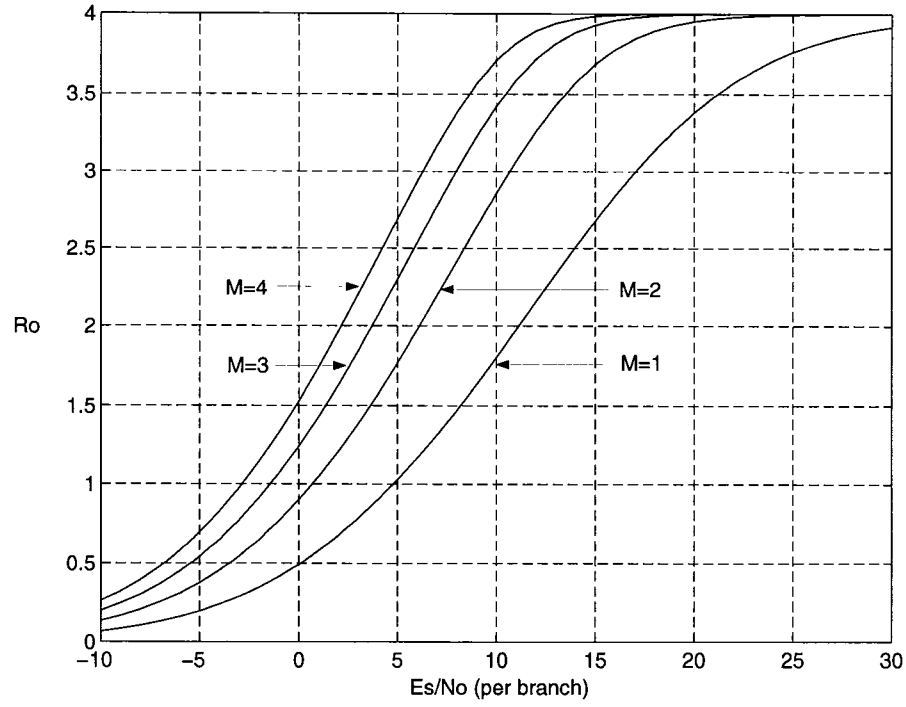


Figure 4.1: The cutoff rate of 16-QAM with maximal ratio combining and different diversity orders

4.3.2 Equal Gain Combining

In the equal gain combining scheme the Chernoff bound can be written as

$$\begin{aligned} C(x_i, \hat{x}_i, \lambda|\underline{a}_i) &= E_{\underline{n}_i}[\exp(\lambda\{m(\hat{x}_i, \underline{y}_i; \underline{a}_i) - m(x_i, \underline{y}_i; \underline{a}_i)\})] \\ &= E_{\underline{n}_i}[\exp(\lambda\{|\sum_{l=1}^{M}(y_{i,l} - a_{i,l}x_i)|^2 - |\sum_{l=1}^{M}(y_{i,l} - a_{i,l}\hat{x}_i)|^2\})] \end{aligned} \tag{4.19}$$

which can be simplified to

$$\begin{aligned} C(x_i, \hat{x}_i, \lambda|\underline{a}_i) &= E_{n_{il}}[\exp(\lambda\{-(\sum_{l=1}^{M} a_{i,l})^2|x_i - \hat{x}_i|^2 - \\ &\quad -2(\sum_{l=1}^{M} a_{i,l})\Re(\sum_{l=1}^{M} n_{i,l} \cdot (x_i - \hat{x}_i)^*)\})] \\ &= \exp(-\lambda(\sum_{l=1}^{M} a_{i,l})^2|x_i - \hat{x}_i|^2)\times \\ &\quad E_{n_{i,l}}[\exp\{-2\lambda\sum_{l=1}^{M} a_{i,l}\Re(\sum_{l=1}^{M} n_{il} \cdot (x_i - \hat{x}_i)^*)\}]. \end{aligned} \tag{4.20}$$

However,

$$E_{n_{il}}[\exp\{-2\lambda\sum_{l=1}^{M} a_{i,l}\Re(\sum_{l=1}^{M} n_{i,l}\cdot(x_i-\hat{x}_i)^*)\}] = \exp\{\lambda^2(\sum_{l=1}^{M} a_{i,l})^2 MN_o|x_i-\hat{x}_i|^2\}. \tag{4.21}$$

Therefore,

$$C(x_i,\hat{x}_i,\lambda|\underline{a}_i) = \exp(-\lambda(\sum_{l=1}^{M} a_{i,l})^2(1-MN_o\lambda)|x_i-\hat{x}_i|^2). \tag{4.22}$$

In this case $C(x_i,\hat{x}_i,\lambda|\underline{a}_i)$ can be minimized by choosing $\lambda = \frac{1}{2MN_o}$. Therefore, the tightest conditional Chernoff bound can be written as

$$C(x_i\hat{x}_i|\underline{a}_i) = \min\{C(x_i,\hat{x}_i,\lambda|\underline{a}_i)\} = \exp(-\mu\frac{|x_i-\hat{x}_i|^2}{4MN_o}) \tag{4.23}$$

where

$$\mu = t^2 = (\sum_{l=1}^{M} a_{i,l})^2. \tag{4.24}$$

However, no closed form expression for the sum of Rayleigh distributed random variables is available for the case of $M > 2$, so an approximate expression is used. This expression is based on the small argument approximation [65, 31]. (Beaulieu [72] showed that this expression is very accurate for $M \leq 8$). The approximation to the pdf of t is given by

$$f(t) = \frac{t^{(2M-1)}\exp(-\frac{t^2}{2b_o})}{2^{(M-1)}b_o^M(M-1)!} \tag{4.25}$$

for $t \geq 0$, where

$$b_o = [(2M-1)!!]^{1/M} = [(2M-1)\cdot(2M-3)\cdots 3\cdot 1]^{1/M}. \tag{4.26}$$

Using the transformation $f_\mu(\mu) = f_T(\sqrt{\mu})/2\sqrt{\mu}$, $f(\mu)$ can be written as

$$f(\mu) = \frac{\mu^{(M-1)}\exp(-\frac{\mu}{b_o})}{b_o^M(M-1)!} \tag{4.27}$$

for $\mu \geq 0$. In other words, μ has an M-Erlang distribution with parameter $1/b_o$. The last step is to perform the following intergration

$$C(x_i, \hat{x}_i) = \frac{1}{(2M-1)!! \cdot (M-1)!} \int_0^\infty \mu^{M-1} \exp(-\frac{\mu}{b_o}) \exp(-\mu \frac{|x_i - \hat{x}_i|^2}{4MN_o}) d\mu. \tag{4.28}$$

Using variable transformation and Eqn. 4.16 the previous integral reduces to

$$C(x_i, \hat{x}_i) = \frac{1}{\left(1 + \frac{[(2M-1)!!]^{1/M}}{M} \frac{|x_i - \hat{x}_i|^2}{4N_o}\right)^M}. \tag{4.29}$$

Therefore, the cutoff rate for equal gain combining receivers is expressed as

$$R_o = 2\log_2(|A|) - \log_2 \left(\sum_{x_i \in A} \sum_{x_j \in A} \frac{1}{\left(1 + \frac{[(2M-1)!!]^{1/M}}{M} \frac{|x_i - \hat{x}_i|^2}{4N_o}\right)^M} \right). \tag{4.30}$$

Clearly, maximal ratio combining, at high SNR, is better than equal gain combining by $\frac{M}{[(2M-1)!!]^{1/M}}$ which is always > 1 and monotonically increases with increasing M as shown in Table 4.1.

M	$\frac{M^M}{(2M-1)!!}$	$\frac{M^M}{(2M-1)!!}$ in dB
1	1	0
2	1.155	0.625
3	1.216	0.861
4	1.250	0.968
8	1.302	1.147

Table 4.1: Improvement factor for MRC over EGC

Figure 4.2 shows the cutoff rate values of the 16-QAM signal constellation and equal gain combining with a diversity order of $M = 1, 2, 3, 4$. Similar to

the maximal ratio combining case, the largest incremental gain is obtained in going from single to double diversity and coding gains diminish as the order of diversity increases.

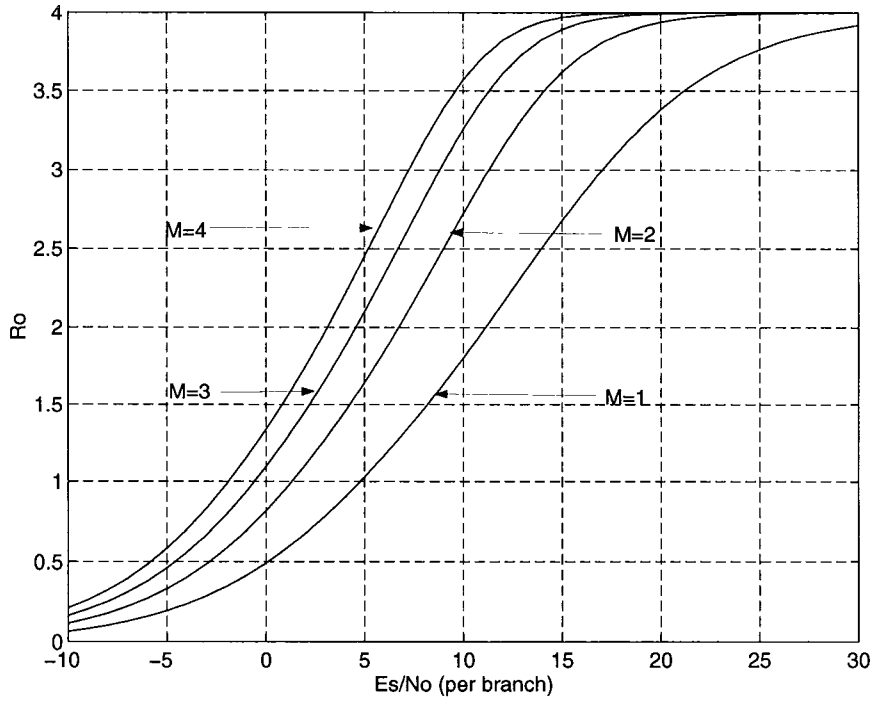


Figure 4.2: The cutoff rate of 16-QAM with equal gain combining and different diversity orders

4.3.3 Selection Combining

In this case the Chernoff bound can be written as

$$\begin{aligned} C(x_i, \hat{x}_i, \lambda|\underline{a}_i) &= E_{\underline{n}_i}[\exp(\lambda\{m(\hat{x}_i, \underline{y}_i; \underline{a}_i) - m(x_i, \underline{y}_i; \underline{a}_i)\})] \\ &= E_{n_{i,j^\star}}[\exp(\lambda\{|y_{i,j^\star} - a_{i,j^\star}x_i|^2 - |y_{i,j^\star} - a_{i,j^\star}\hat{x}_i|^2\})]. \end{aligned} \tag{4.31}$$

Again, this can be simplified to

$$\begin{aligned} C(x_i, \hat{x}_i, \lambda|\underline{a}_i) &= E_{n_{i,j^\star}}[\exp(\lambda\{-a_{i,j^\star}^2|x_i - \hat{x}_i|^2 - -2(a_{i,j^\star})\Re(n_{i,j^\star} \cdot (x_i - \hat{x}_i)^*)\})] \\ &= \exp(-\lambda a_{i,j^\star}^2|x_i - \hat{x}_i|^2) E_{n_{i,j^\star}}[\exp\{-2\lambda a_{i,j^\star}\Re(n_{i,j^\star} \cdot (x_i - \hat{x}_i)^*)\}] \\ &= \exp(-\lambda a_{i,j^\star}^2|x_i - \hat{x}_i|^2) \exp\{\lambda^2 a_{i,j^\star}^2 N_o|x_i - \hat{x}_i|^2\}. \end{aligned} \tag{4.32}$$

Therefore,

$$C(x_i, \hat{x}_i, \lambda|\underline{a}_i) = \exp(-\lambda(a_{i,j^\star})^2(1 - N_o\lambda)|x_i - \hat{x}_i|^2). \tag{4.33}$$

$C(x_i, \hat{x}_i, \lambda|\underline{a}_i)$ can be minimized by choosing $\lambda = \frac{1}{2N_o}$. Therefore, the tightest conditional Chernoff bound can be written as

$$C(x_i\hat{x}_i|a_{i,j^\star}) = \min_{\lambda}\{C(x_i, \hat{x}_i, \lambda|a_{i,j^\star})\} = \exp(-\nu\frac{|x_i - \hat{x}_i|^2}{4N_o}) \tag{4.34}$$

where

$$\nu = a_{i,j^\star} = \max\{a_{i,1}^2, a_{i,2}^2, \cdots a_{i,M}^2\}. \tag{4.35}$$

Since $\{a_{i,1}, a_{i,2} \ldots a_{i,M}\}$ are independent, ν is just the maximum of M independent exponential random variables, so

$$f(\nu) = M[1 - \exp(-\nu)]^{(M-1)} \exp(-\nu) \quad \nu \geq 0 \tag{4.36}$$

which can be rewritten using the binomial expansion as

$$f(\nu) = \sum_{k=1}^{M}(-1)^{k+1} M \begin{pmatrix} M-1 \\ k-1 \end{pmatrix} \exp(-k\nu). \tag{4.37}$$

The last step is to perform the following intergration

$$C(x_i, \hat{x}_i) = \sum_{k=1}^{M}(-1)^{k+1} M \begin{pmatrix} M-1 \\ k-1 \end{pmatrix} \int_0^\infty \exp(-k\nu) \exp(-\nu\frac{|x_i - \hat{x}_i|^2}{4N_o}) d\nu. \tag{4.38}$$

Evaluating the integral, the Chernoff factor reduces to

$$C(x_i, \hat{x}_i) = \sum_{k=1}^{M} (-1)^{k+1} M \begin{pmatrix} M-1 \\ k-1 \end{pmatrix} \frac{1}{k + \frac{|x_i - \hat{x}_i|^2}{4N_o}}. \tag{4.39}$$

Therefore, the cutoff rate for maximal ratio combining receivers is given by

$$R_o = 2\log_2(|A|) - \log_2 \left(\sum_{x_i \in A} \sum_{x_j \in A} \left[\sum_{k=1}^{M} (-1)^{k+1} M \begin{pmatrix} M-1 \\ k-1 \end{pmatrix} \frac{1}{k + \frac{|x_i - \hat{x}_i|^2}{4N_o}} \right] \right). \tag{4.40}$$

Figure 4.3 shows the cutoff rate values of the 16-QAM signal constellation and selection combining with diversity orders of $M = 1, 2, 3, 4$. Once again, the largest incremental gain is obtained in going from single to double diversity and the coding gains diminish as the order of diversity increases. Moreover, when more diversity is added the "diminishing returns" effect is more obvious in this case.

4.3.4 Comparisons and Numerical Results

Comparing the expressions for the three combining schemes, it is clear that the MRC scheme achieves the best performance. In fact, it is the optimal combining scheme for equal energy branches [65]. Moreover, SC gives the lowest performance among all other schemes. Figure 4.4 and Figure 4.5 show the cutoff rate for the 16-QAM constellation with the three combining schemes and diversity orders of $M = 2$ and $M = 4$, respectively. Finally, Tables 4.2- 4.5 give the values of E_s/N_o corresponding to different rates and different constellations.

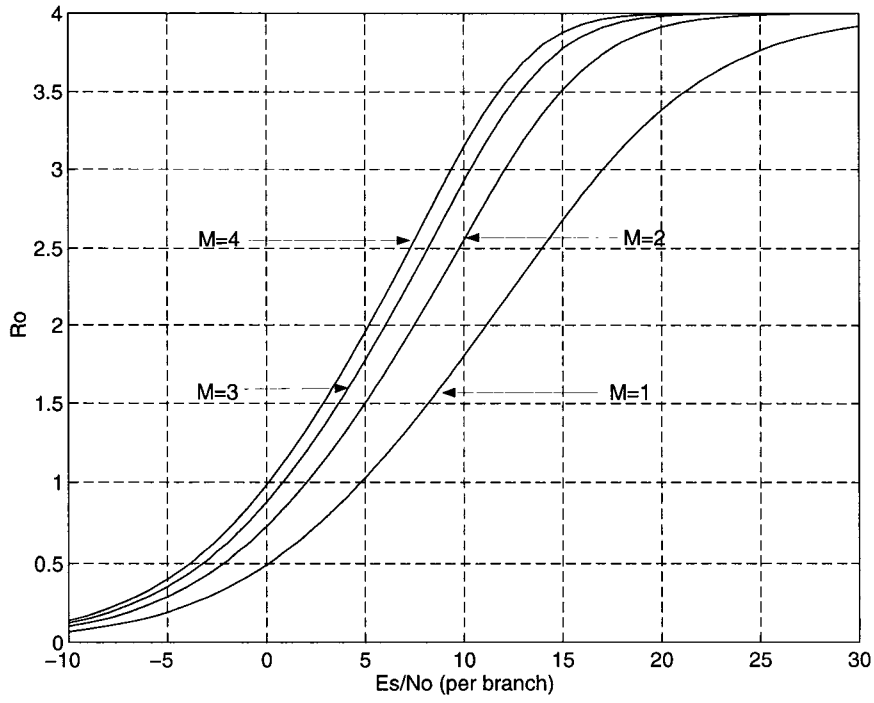


Figure 4.3: The cutoff rate of 16-QAM with selection combining and different diversity orders

M	MRC	EGC	SC
2	0.7	1.4	2.1
3	-1.5	-.7	0.8
4	-2.9	-2.0	0.1
8	-6.3	-5.2	-1.4

Table 4.2: The E_s/N_o per branch values (in dB) to achieve $R_o = 1$ for the QPSK constellation

M	MRC	EGC	SC
2	6.7	7.3	8.1
3	4.2	5.0	6.5
4	2.6	3.5	5.6
8	-1.1	-.2	4.0

Table 4.3: The E_s/N_o per branch values (in dB) to achieve $R_o = 2$ for the 8-PSK constellation

M	MRC	EGC	SC
2	6.1	6.7	7.5
3	3.7	4.6	6.0
4	2.1	3.1	5.2
8	-1.3	-0.2	3.6

Table 4.4: The E_s/N_o per branch values (in dB) to achieve $R_o = 2$ for the 16-QAM constellation

M	MRC	EGC	SC
2	10.7	11.3	12.1
3	8.0	8.8	10.3
4	6.3	7.2	9.4
8	2.7	3.8	7.6

Table 4.5: The E_s/N_o per branch values (in dB) to achieve $R_o = 3$ for the 16-QAM constellation

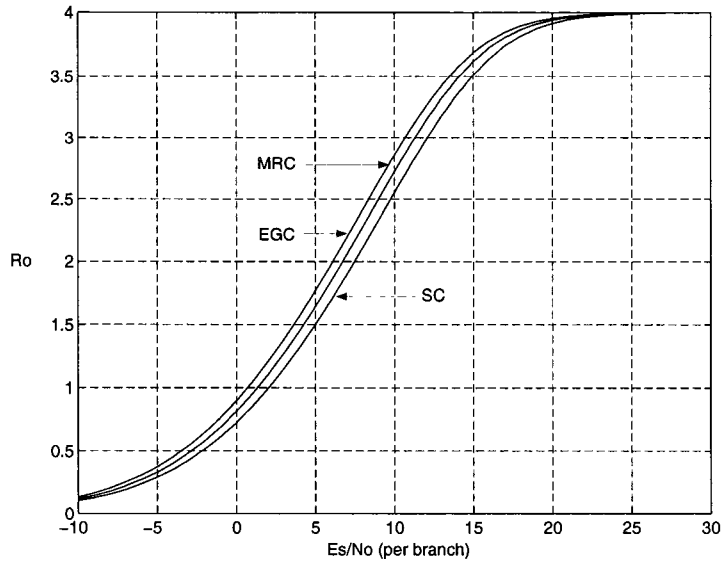


Figure 4.4: The cutoff rate of 16-QAM with $M = 2$ and different combining schemes.

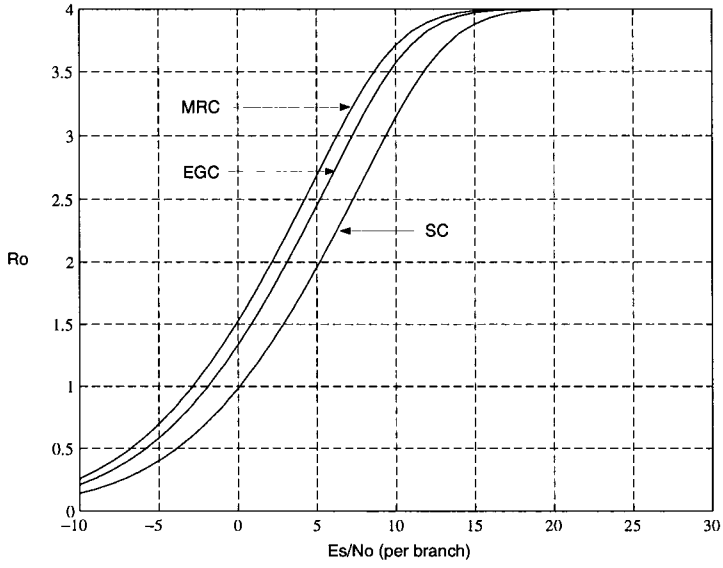


Figure 4.5: The cutoff rate of 16-QAM with $M = 4$ and different combining schemes.

4.4 Pairwise Error Probability

In this section, tight upper bounds on the pairwise error probability are derived for the three combining schemes. Moreover, the pairwise error probability expressions are expressed in product form – i.e.,

$$P(\mathbf{x}_N \rightarrow \hat{\mathbf{x}}_N | \mathbf{a}_N) = K_c \times \prod_{l=1}^{N} W(x_l, \hat{x}_l), \tag{4.41}$$

where K_c is a constant that does not depend on the length of the error sequence, and $W(x_l, \hat{x}_l)$ is the error weight profile between x_l and $\hat{x}_l$. Expressing the pairwise error probability in this form allows the use of the transfer function technique of trellis codes to be used. In contrast with other methods of finding the bit error probability, the transfer function method accounts for all the paths of the trellis.

Using the union bound approach, the average bit error probability, P_b, of trellis coded systems is upper bounded as [61]

$$P_b \leq \sum_{n=1}^{\infty} \sum_{\mathbf{x}_n} \sum_{\mathbf{x}_n \neq \hat{\mathbf{x}}_n} N_n(\mathbf{x}_n, \hat{\mathbf{x}}_n) \Pr(\mathbf{x}_n) P(\mathbf{x}_n \rightarrow \hat{\mathbf{x}}_n) \tag{4.42}$$

where $\Pr(\mathbf{x}_n)$ is the a priori probability of transmitting the sequence $\mathbf{x}_n$, and $N_n(\mathbf{x}_n, \hat{\mathbf{x}}_n)$ is the number of bit errors when $\mathbf{x}_n$ is the transmitted sequence and $\hat{\mathbf{x}}_n$ is chosen by the decoder. Note that [61]

$$T(\bar{D}, I) = \sum_{n=1}^{\infty} \sum_{\mathbf{x}_n} \sum_{\mathbf{x}_n \neq \hat{\mathbf{x}}_n} \Pr(\mathbf{x}_n) I^{N_n(\mathbf{x}_n, \hat{\mathbf{x}}_n)} \prod_{i=1}^{N} W(x_i, \hat{x}_i) \tag{4.43}$$

where $T(\bar{D}, I)$ is the transfer function of the code with D averaged over the fading distribution.

4.4.1 Maximal Ratio Combiner

The conditional pairwise error probability for maximal ratio combining can be expressed as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) &= P(m(\hat{\mathbf{x}}_N, \mathbf{y}_N; \mathbf{a}_N) - m(\mathbf{x}_N, \mathbf{y}_N; \mathbf{a}_N) \geq 0|\mathbf{a}_N) \\ &= P(\textstyle\sum_{i=1}^{N}\sum_{l=1}^{M}(|y_{i,l} - a_{i,l}x_i|^2 - |y_{i,l} - a_{i,l}\hat{x}_i|^2) \geq 0|\mathbf{a}_N) \end{aligned} \tag{4.44}$$

which can be simplified to

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) &= P(\textstyle\sum_{i=1}^{N}\sum_{l=1}^{M}[-a_{i,l}^2|x_i - \hat{x}_i|^2 - 2a_{i,l}\Re\{n_{i,l}(x_i - \hat{x}_i)^*\}] \leq 0|\mathbf{a}_N) \\ &= P(z_N \geq \textstyle\sum_{i=1}^{N}\sum_{l=1}^{M} a_{i,l}^2|x_i - \hat{x}_i|^2|\mathbf{a}_N), \end{aligned} \tag{4.45}$$

where z_N is a Gaussian random variable with zero mean and a variance of $2N_o\sum_{i=1}^{N}\sum_{l=1}^{M} a_{i,l}^2|x_i - \hat{x}_i|^2$. This probability can be expressed as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) = \frac{1}{2}\text{erfc}\left(\sqrt{\sum_{i=1}^{N}\gamma_i d_i}\right) \tag{4.46}$$

where $\text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt$, $d_i = \frac{|x_i - \hat{x}_i|^2}{4N_o}$ and $\gamma_i = \sum_{l=1}^{M} a_{i,l}^2$. Since the $a_{i,l}$'s are i.i.d. Rayleigh distributed random variables with $E(a_{i,l}^2) = 1$, their squares are i.i.d. exponentially distributed random variables with a mean equal to one. Hence, γ_i will have an M-Erlang distribution with parameter one – i.e., its pdf is

$$f(\gamma_i) = \frac{1}{(M-1)!}\gamma_i^{(M-1)}e^{-\gamma_i}. \tag{4.47}$$

The unconditional pairwise error probability is thus

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) = \frac{1}{2}\int_0^\infty \cdots \int_0^\infty \text{erfc}\left(\sqrt{\sum_{i=1}^{N}\gamma_i d_i}\right) f(\gamma_1)\cdots f(\gamma_N)d\gamma_1 \cdots d\gamma_N. \tag{4.48}$$

Define

$$\delta_i = \frac{d_i}{1 + d_i} \qquad \text{and} \qquad \omega_i = \gamma_i(1 + d_i). \tag{4.49}$$

So $f(\gamma_i)d\gamma_i$ can be expressed as

$$\begin{aligned} f(\gamma_i)d\gamma_i &= \tfrac{1}{(1+d_i)^M} e^{\delta_i\omega_i} \times \tfrac{1}{(M-1)!}\omega_i^{M-1}e^{-\omega_i}d\omega_i \\ &= \tfrac{1}{(1+d_i)^M} e^{\delta_i\omega_i} \times f(\omega_i)d\omega_i. \end{aligned} \tag{4.50}$$

where $f(\omega_i)$ is the M-Erlang distribution of ω_i with parameter one. Therefore, the unconditional pairwise error probability can be represented as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) &= \tfrac{1}{2}\textstyle\prod_{i\in\eta} \frac{1}{(1+d_i)^M} \int_0^\infty \cdots \int_0^\infty \operatorname{erfc}\left(\sqrt{\textstyle\sum_{i=1}^N \delta_i\omega_i}\right) \\ &\quad \times \exp\left[\textstyle\sum_{i=1}^N \delta_i\omega_i\right] f(\omega_1)\cdots f(\omega_N)d\omega_1\cdots d\omega_N \end{aligned} \tag{4.51}$$

where $\eta = \{i : x_i \neq \hat{x}_i\}$ and $L_\eta = |\eta|$. Note that

$$\sum_{i\in\eta} \delta_i\omega_i \geq \sum_{i\in\eta} \delta_m\omega_i \tag{4.52}$$

where $\delta_m = min\{\delta_i, i \in \eta\}$, and since $\operatorname{erfc}(x)e^{x^2}$ is monotonically decreasing for $x \geq 0$, then the pairwise error probability can be upper bounded as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \frac{1}{2}\prod_{i\in\eta} \frac{1}{(1+d_i)^M} \int_0^\infty \operatorname{erfc}\left(\sqrt{\delta_m\Omega}\right) \times e^{(\delta_m\Omega)} f(\Omega)d\Omega, \tag{4.53}$$

where $\Omega = \sum_{i\in\eta}\omega_i$. Since the ω_i's are independent M-Erlang distributed random variables each with parameter one, Ω will have an (ML_η)-Erlang distribution with parameter one; i.e.,

$$f(\Omega) = \tfrac{1}{(ML_\eta-1)!}\Omega^{(ML_\eta-1)}e^{-\Omega}, \quad \Omega \geq 0. \tag{4.54}$$

To evaluate the integral, we use the following equality [62]

$$\frac{1}{2(K-1)!}\int_0^\infty \operatorname{erfc}(\sqrt{xy})e^{-y(1-x)}y^{(K-1)}dy = \frac{1}{2^{2K}}\sum_{j=1}^{K}\binom{2K-j-1}{K-1}\left(\frac{2}{1+\sqrt{x}}\right)^j \tag{4.55}$$

which is valid for $x < 1$. Integration yields

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \left[\frac{1}{2^{2ML_\eta}}\sum_{j=1}^{ML_\eta}\binom{2ML_\eta-j-1}{ML_\eta-1}\left(\frac{2}{1+\sqrt{\delta_m}}\right)^j\right] \times \prod_{i\in\eta}\frac{1}{(1+d_i)^M}. \tag{4.56}$$

Let L be the minimum time diversity of the code – i.e., the minimum Hamming distance, in signal symbols, between any two signal symbols. Then, $L \leq L_\eta$ and we can further upper bound the pairwise error probability by

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \left[\frac{1}{2^{2ML}} \sum_{j=1}^{ML} \begin{pmatrix} 2ML-j-1 \\ ML-1 \end{pmatrix} \left(\frac{2}{1+\sqrt{\delta_m}} \right)^j \right] \times \prod_{i \in \eta} \frac{1}{(1+d_i)^M}. \tag{4.57}$$

Note that the upper bound in [62] is a special case of this bound ($M = 1$). The upper bound in Eqn. 4.57 is written in a product form. This allows us to use the transfer function of the trellis code to account for all trellis paths. Therefore, the bit error probability can be expressed as

$$P_b \leq \frac{1}{k} \left[\frac{1}{2^{2ML}} \sum_{j=1}^{ML} \begin{pmatrix} 2ML-j-1 \\ ML-1 \end{pmatrix} \left(\frac{2}{1+\sqrt{\delta_m}} \right)^j \right] \frac{\partial T(\bar{D}, I)}{\partial I} \Big|_{I=1, D=e^{-E_s/4N_o}} \tag{4.58}$$

where

$$\bar{D} \Big|_{D=e^{-E_s/4N_0}} = \frac{1}{(1+\frac{E_s}{4N_o})^M}. \tag{4.59}$$

This upper bound has the same complexity as the Chernoff bound. The extra term on the left of the transfer function is always smaller than or equal to one. So this bound is at least as tight as the Chernoff bound. To see the tightness of the bound, the 8-state 16-QAM code with bandwidth efficiency of 2 bits/sec/Hz is used. For this specific coding/modulation

$$\delta_m = \frac{0.8E_s/N_o}{1+0.8E_s/N_o} \tag{4.60}$$

Figure 4.6 compares the newly derived bound for dual diversity (M=2) and maximal ratio combining with the Chernoff bound for the same code; the new

bound is slightly more than 1 dB tighter than the Chernoff bound at a bit error rate of 10^{-5}.

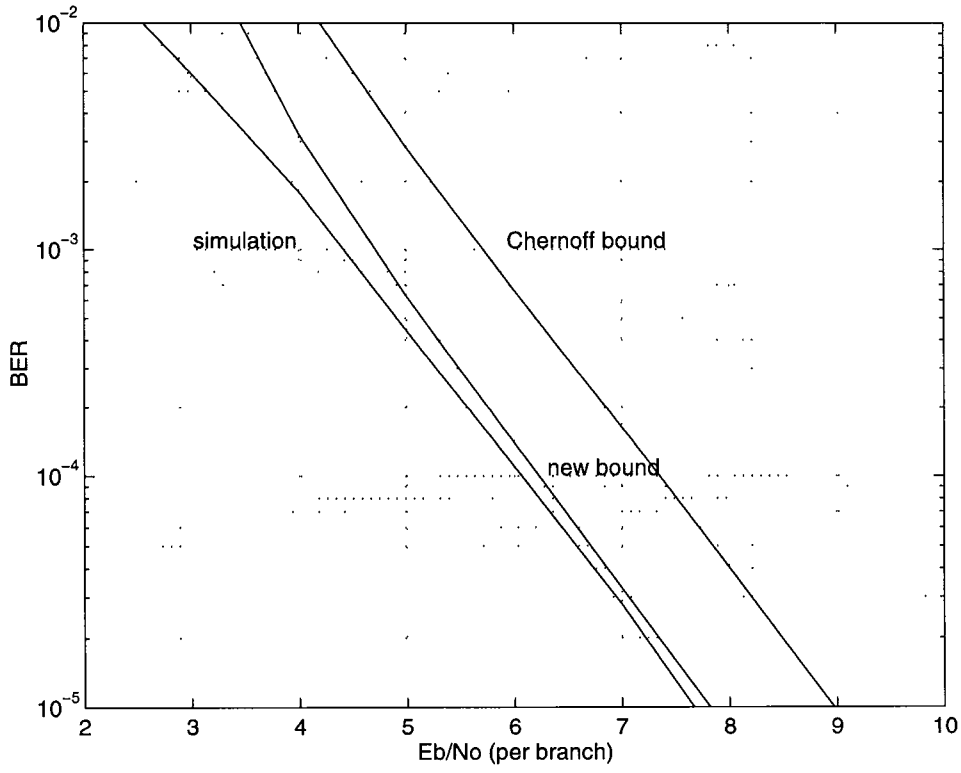


Figure 4.6: A comparison between the new bound and the Chernoff bound for the 16-QAM I-Q TCM 8-state code with maximal ratio combining and double diversity.

The performance of the 16-QAM with MRC and different orders of diversity is shown in Figure 4.7, where 4, 8, and 16-state codes are shown. It is clear that the largest gain is obtained in going from single to double diversity; i.e., the coding gains diminish as the order of diversity increases. This confirms the conclusions obtained from the cutoff rate curves.

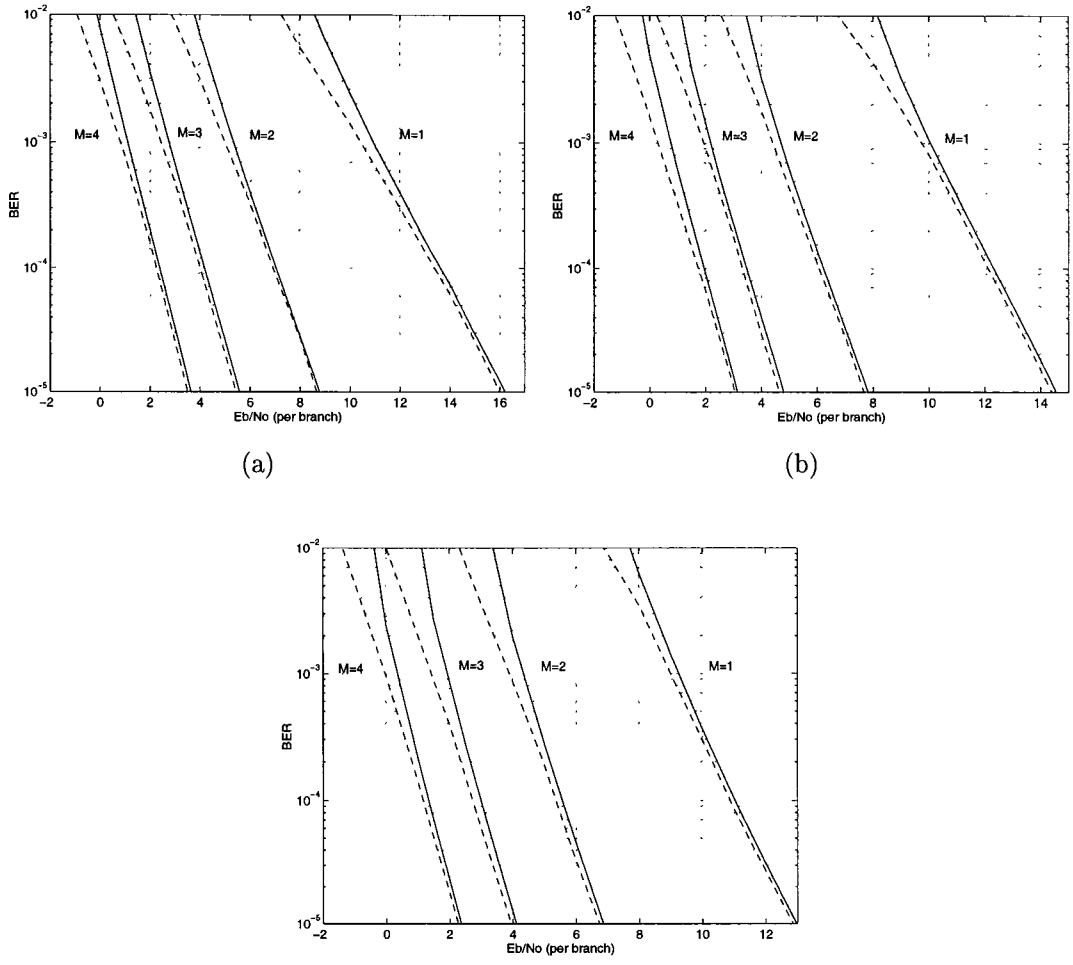


Figure 4.7: BER of 16-QAM I-Q TCM 4,8, and 16-state codes with maximal ratio combining and different diversity orders. solid(bound), dashed(simulation), (a) $\nu = 2$ (b) $\nu = 3$ (c) $\nu = 4$

4.4.2 Equal Gain Combiner

In the case of equal gain combining the conditional pairwise error probability can be expressed as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) &= P(m(\hat{\mathbf{x}}_N, \mathbf{y}_N; \mathbf{a}_N) - m(\mathbf{x}_N, \mathbf{y}_N; \mathbf{a}_N) \geq 0|\mathbf{a}_N) \\ &= P(\textstyle\sum_{i=1}^{N}(|\sum_{l=1}^{M}(y_{i,l} - a_{i,l}x_i)|^2 \\ &\quad -|\textstyle\sum_{l=1}^{M}(y_{i,l} - a_{i,l}\hat{x}_i)|^2) \geq 0|\mathbf{A}_N). \end{aligned} \tag{4.61}$$

It can be simplified to

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) &= P(\textstyle\sum_{i=1}^{N}\{-(\sum_{l=1}^{M} a_{i,l})^2|x_i - \hat{x}_i|^2 \\ &\quad -2(\textstyle\sum_{l=1}^{M} a_{i,l})\Re[\sum_{l=1}^{M} n_{i,l}(x_i - \hat{x}_i)^*]\} \leq 0|\mathbf{a}_N) \\ &= P(z_N \geq \textstyle\sum_{i=1}^{N}(\sum_{l=1}^{M} a_{i,l})^2|x_i - \hat{x}_i|^2|\mathbf{a}_N), \end{aligned} \tag{4.62}$$

where z_N is a Gaussian random variable with zero mean and a variance of $2N_oM\sum_{i=1}^{N}(\sum_{l=1}^{M} a_{i,l})^2|x_i - \hat{x}_i|^2$. This probability can be expressed as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) = \frac{1}{2}\text{erfc}\left(\sqrt{\sum_{i=1}^{N}\mu_i\tilde{d}_i/b_o}\right), \tag{4.63}$$

where $\tilde{d}_i = \frac{b_o|x_i - \hat{x}_i|^2}{4MN_o}$ and $\mu_i = (\sum_{l=1}^{M} a_{i,l})^2$. Here, μ_i has pdf

$$f(\mu_i) = \frac{\mu_i^{(M-1)}\exp(-\frac{\mu_i}{b_o})}{b_o^M(M-1)!} \quad \mu_i \geq 0. \tag{4.64}$$

Define $\gamma_i = \mu_i/b_o$. Then

$$f(\gamma_i) = \frac{\gamma_i^{(M-1)}\exp(-\gamma_i)}{(M-1)!} \quad \gamma_i \geq 0. \tag{4.65}$$

Therefore, the unconditional pairwise error probability can be expressed as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) &= \tfrac{1}{2}\textstyle\int_0^\infty \cdots \int_0^\infty \text{erfc}\left(\sqrt{\sum_{i=1}^{N}\gamma_i\tilde{d}_i}\right) \\ &\quad \times f(\gamma_1)\cdots f(\gamma_N)d\gamma_1\cdots d\gamma_N. \end{aligned} \tag{4.66}$$

Similarly, define

$$\tilde{\delta}_i = \frac{\tilde{d}_i}{1+\tilde{d}_i} \qquad \text{and} \qquad \tilde{\omega}_i = \gamma_i(1+\tilde{d}_i). \tag{4.67}$$

Therefore,

$$\begin{aligned} f(\gamma_i)d\gamma_i &= \tfrac{1}{(1+\tilde{d}_i)^M} e^{\tilde{\delta}_i\tilde{\omega}_i} \times \tfrac{1}{(M-1)!}\tilde{\omega}_i^{M-1} e^{-\tilde{\omega}_i} d\tilde{\omega}_i \\ &= \tfrac{1}{(1+\tilde{d}_i)^M} e^{\tilde{\delta}_i\tilde{\omega}_i} \times f(\tilde{\omega}_i)d\tilde{\omega}_i \end{aligned} \tag{4.68}$$

where $f(\tilde{\omega}_i)$ is M-Erlang distribution of $\tilde{\omega}_i$ with parameter one. Hence the unconditional pairwise error probability can be represented as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) &= \tfrac{1}{2}\textstyle\prod_{i\in\eta} \tfrac{1}{(1+\tilde{d}_i)^M} \int_0^\infty \cdots \int_0^\infty \operatorname{erfc}\left(\sqrt{\textstyle\sum_{i=1}^N \tilde{\delta}_i\tilde{\omega}_i}\right) \\ &\quad \times e^{(\sum_{i=1}^N \tilde{\delta}_i\tilde{\omega}_i)} f(\tilde{\omega}_1)\cdots f(\tilde{\omega}_N) d\tilde{\omega}_1 \cdots d\tilde{\omega}_N \end{aligned} \tag{4.69}$$

where $\eta = \{i : x_i \neq \hat{x}_i\}$ and $L_\eta = |\eta|$. Note that

$$\sum_{i\in\eta} \tilde{\delta}_i\tilde{\omega}_i \geq \sum_{i\in\eta} \delta_e\tilde{\omega}_i \tag{4.70}$$

where $\delta_e = min\{\tilde{\delta}_i, i \in \eta\}$. Hence, the pairwise error probability can be upper bounded as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \frac{1}{2}\prod_{i\in\eta} \frac{1}{(1+\tilde{d}_i)^M} \int_0^\infty \operatorname{erfc}\left(\sqrt{\delta_e\tilde{\Omega}}\right) \times e^{(\delta_e\tilde{\Omega})} f(\tilde{\Omega}) d\tilde{\Omega} \tag{4.71}$$

where $\tilde{\Omega} = \sum_{i\in\eta} \tilde{\omega}_i$ and $\tilde{\Omega}$ is distributed as

$$f(\tilde{\Omega}) = \tfrac{1}{(ML_\eta - 1)!}\tilde{\Omega}^{(ML_\eta-1)} e^{-\tilde{\Omega}} \quad \tilde{\Omega} \geq 0. \tag{4.72}$$

Finally performing the integration yields

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \left[\frac{1}{2^{2ML_\eta}} \sum_{j=1}^{ML_\eta} \binom{2ML_\eta - j - 1}{ML_\eta - 1} \left(\frac{2}{1+\sqrt{\delta_e}}\right)^j\right] \times \prod_{i\in\eta} \frac{1}{(1+\tilde{d}_i)^M}. \tag{4.73}$$

Again, since $L \leq L_\eta$ we can further upper bound $P(\mathbf{x}_N \rightarrow \hat{\mathbf{x}}_N)$ by

$$P(\mathbf{x}_N \rightarrow \hat{\mathbf{x}}_N) \leq \left[\frac{1}{2^{2ML}} \sum_{j=1}^{ML} \begin{pmatrix} 2ML-j-1 \\ ML-1 \end{pmatrix} \left(\frac{2}{1+\sqrt{\delta_e}} \right)^j \right] \times \prod_{i \in \eta} \frac{1}{(1+\tilde{d}_i)^M}. \tag{4.74}$$

Therefore, the bit error probability can be tightly upper bounded by

$$P_b \leq \frac{1}{k} \left[\frac{1}{2^{2ML}} \sum_{j=1}^{ML} \begin{pmatrix} 2ML-j-1 \\ ML-1 \end{pmatrix} \left(\frac{2}{1+\sqrt{\delta_e}} \right)^j \right] \frac{\partial T(\bar{D}, I)}{\partial I} \Big|_{I=1, D=e^{-E_s/4N_o}}. \tag{4.75}$$

where

$$\bar{D} \Big|_{D=e^{-E_s/4N_0}} = \frac{1}{(1+\frac{b_o E_s}{4MN_o})^M}. \tag{4.76}$$

For the same previous coding/modulation schemes, δ_e is expressed as

$$\delta_e = \frac{0.8 b_o E_s/MN_o}{1+(0.8 b_o E_s/MN_o)}. \tag{4.77}$$

The performance of the previous codes with equal gain combining and different orders of diversity is shown in Figure 4.8. Similar to the MRC case, the largest gain is obtained in going from single to double diversity.

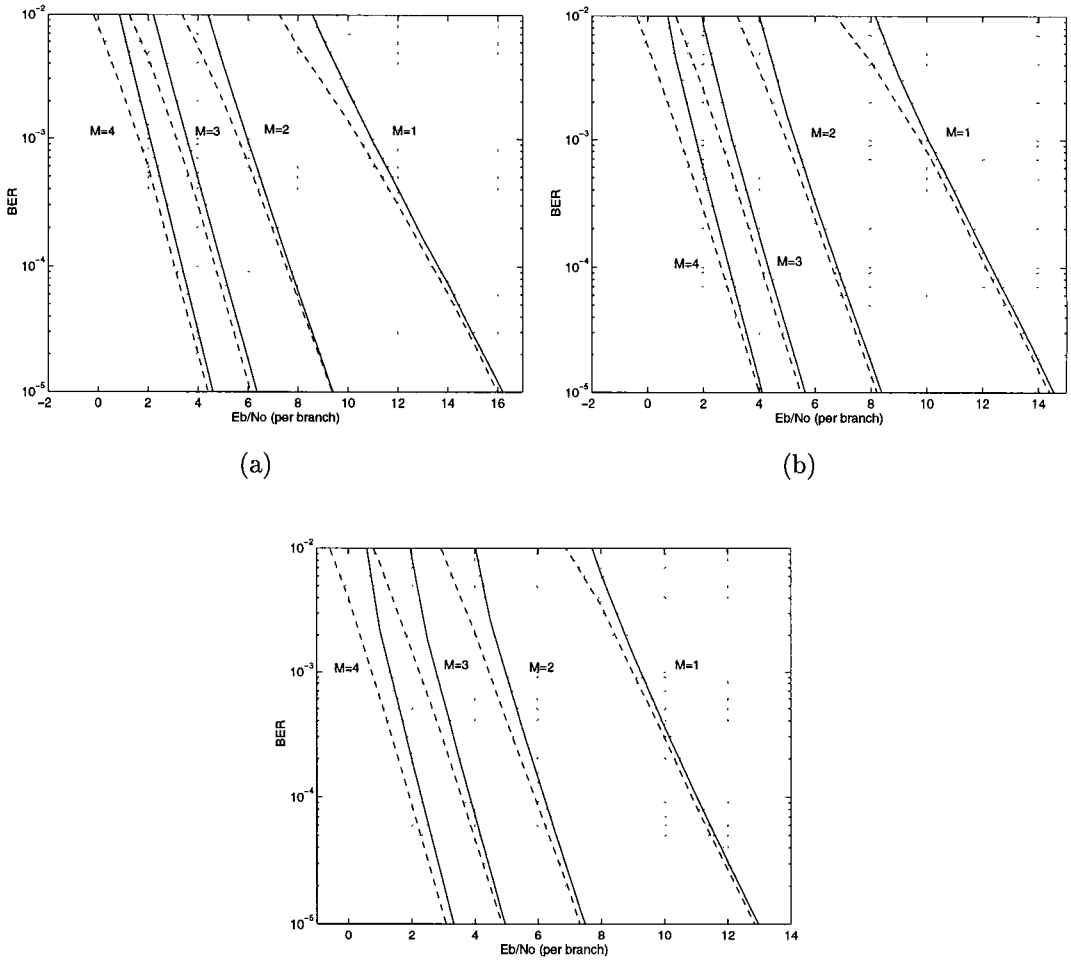


Figure 4.8: BER of 16-QAM I-Q TCM 4,8, and 16-state codes with equal gain combining and different diversity orders. solid(bound), dashed(simulation), (a) $\nu = 2$ (b) $\nu = 3$ (c) $\nu = 4$

4.4.3 Selection Combining

With selection combining, the conditional pairwise error probability can be expressed as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) &= P(m(\hat{\mathbf{x}}_N, \mathbf{y}_N; \mathbf{a}_N) - m(\mathbf{x}_N, \mathbf{y}_N; \mathbf{a}_N) \geq 0|\mathbf{a}_N) \\ &= P(\textstyle\sum_{i=1}^{N}(|y_{i,l} - a_{i,j^\star}x_i|^2 - |y_{i,l} - a_{i,j^\star}\hat{x}_i|^2) \geq 0|\mathbf{a}_N). \end{aligned} \tag{4.78}$$

This expression can be simplified to

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) &= P(\textstyle\sum_{i=1}^{N}(-a_{i,j^\star}^2|x_i - \hat{x}_i|^2 - 2a_{i,j^\star}\Re\{n_{i,l}(x_i - \hat{x}_i)^*\}) \leq 0|\mathbf{a}_N) \\ &= P(z_N \geq \textstyle\sum_{i=1}^{N} \nu_i|x_i - \hat{x}_i|^2|\mathbf{a}_N). \end{aligned} \tag{4.79}$$

where $\nu_i = a_{i,j^\star}^2$, z_N is a Gaussian random variable with zero mean and a variance of $2N_o \sum_{i=1}^{N} \nu_i|x_i - \hat{x}_i|^2$. This probability can be expressed as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N|\mathbf{a}_N) = \frac{1}{2}\text{erfc}\left(\sqrt{\sum_{i=1}^{N}\nu_i d_i}\right) \tag{4.80}$$

where $d_i = \frac{|x_i - \hat{x}_i|^2}{4N_o}$. The pdf of ν_i is

$$\begin{aligned} f(\nu_i) &= M(1 - e^{-\nu_i})^{(M-1)}e^{-\nu_i} \\ &= \textstyle\sum_{k_i=1}^{M} M(-1)^{k_i+1}\begin{pmatrix} M-1 \\ k_i - 1 \end{pmatrix} e^{-k_i\nu_i}. \end{aligned} \tag{4.81}$$

The unconditional pairwise error probability can be expressed as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) &= \tfrac{1}{2}\textstyle\sum_{k_1=1}^{M}\cdots\sum_{K_N=1}^{M}\prod_{i=1}^{N}\left\{M(-1)^{k_i+1}\begin{pmatrix} M-1 \\ k_i - 1 \end{pmatrix}\right\} \\ &\quad \times \textstyle\int_0^\infty \cdots \int_0^\infty \text{erfc}\left(\sqrt{\sum_{i=1}^{N}\nu_i d_i}\right) e^{-k_1\nu_1}\cdots e^{-k_N\nu_N} d\nu_1 \cdots d\nu_N. \end{aligned} \tag{4.82}$$

Define

$$\delta_{i,k_i} = \frac{d_i}{k_i + d_i} \qquad \text{and} \qquad \omega_{i,k_i} = \nu_i(k_i + d_i). \tag{4.83}$$

Hence,

$$f(\nu_i)d\nu_i = \sum_{k_i=1}^{M} M(-1)^{k_i+1} \begin{pmatrix} M-1 \\ k_i-1 \end{pmatrix} \left(\frac{1}{k_i+d_i}\right) e^{\delta_{i,k_i}\omega_{i,k_i}} e^{-\omega_{i,k_i}} d\omega_{i,k_i}. \quad (4.84)$$

So, the pairwise error probability can be expressed again as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) &= \tfrac{1}{2}\textstyle\sum_{k_1=1}^{M}\cdots\sum_{K_N=1}^{M}\prod_{i=1}^{N}\{M(-1)^{k_i+1}\begin{pmatrix} M-1 \\ k_i-1 \end{pmatrix}\left(\tfrac{1}{k_i+d_i}\right)\} \\ &\quad \times \textstyle\int_0^\infty\cdots\int_0^\infty \operatorname{erfc}\left(\sqrt{\sum_{i=1}^{N}\delta_{i,k_i}\omega_{i,k_i}}\right) e^{(\sum_{i=1}^{N}\delta_{i,k_i}\omega_{i,k_i})} \\ &\quad \times e^{-\omega_{1,k_1}}\cdots e^{-\omega_{N,k_N}} d\omega_{1,k_1}\cdots d\omega_{N,k_N}. \end{aligned} \quad (4.85)$$

Define

$$\Gamma = \sum_{i\in\eta}\omega_{i,k_i}. \quad (4.86)$$

Therefore,

$$f(\Gamma) = \tfrac{1}{(L_\eta-1)!}\Gamma^{(L_\eta-1)}e^{-\Gamma} \quad \Gamma \geq 0. \quad (4.87)$$

Also, observe that

$$\sum_{i\in\eta}\delta_{i,k_i}\omega_{i,k_i} \geq \sum_{i\in\eta}\delta_s\omega_{i,k_i} \quad (4.88)$$

where

$$\delta_s = \min\{\delta_{i,k_i}, i\in\eta, k_i\in\{1,\ldots M\}\} = \frac{\min\{d_i\}}{M+\min\{d_i\}}. \quad (4.89)$$

Using the above expressions, the pairwise error probability can be expressed as

$$\begin{aligned} P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) &\leq \tfrac{1}{2}\textstyle\prod_{i\in\eta}\left[\sum_{k=1}^{M} M(-1)^{k+1}\begin{pmatrix} M-1 \\ k-1 \end{pmatrix}\tfrac{1}{(k+d_i)}\right] \times \\ &\quad \textstyle\int_0^\infty \operatorname{erfc}\left(\sqrt{\delta_s\Gamma}\right) e^{(\delta_s\Gamma)} f(\Gamma)d\Gamma. \end{aligned} \quad (4.90)$$

The final step is to evaluate the integral and replace L_η by L. Hence, the pairwise error probability can now be expressed as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \left[\frac{1}{2^{2L}} \sum_{j=1}^{L} \binom{2L-j-1}{L-1} \frac{2^j}{(1+\sqrt{\delta_s})^j} \right] \times \prod_{i \in \eta} \left[\sum_{k=1}^{M} M(-1)^{k+1} \binom{M-1}{k-1} \frac{1}{(k+d_i)} \right]. \tag{4.91}$$

Therefore, the bit error probability can now be expressed as

$$P_b \leq \frac{1}{k} \left[\frac{1}{2^{2L}} \sum_{j=1}^{L} \binom{2L-j-1}{L-1} \left(\frac{2}{1+\sqrt{\delta_s}} \right)^j \right] \frac{\partial T(\bar{D}, I)}{\partial I} \Big|_{I=1, D=e^{-E_s/4N_o}}, \tag{4.92}$$

where

$$\bar{D} \Big|_{D=e^{-E_s/4N_0}} = \sum_{k=1}^{M} M(-1)^{k+1} \binom{M-1}{k-1} \frac{1}{k + \frac{E_s}{4N_o}}. \tag{4.93}$$

For the same previous coding/modulation schemes, δ_s is expressed as

$$\delta_s = \frac{0.8E_s/N_o}{M + 0.8E_s/N_o}. \tag{4.94}$$

The performance of the previous codes with selection combining and different orders of diversity is shown in Figure 4.9. Similar to the two combining schemes, the largest gain is obtained in going from single to double diversity. Moreover, with increasing the diversity order the coding gains diminish faster than in the other combining cases.

Comparisons between the three combining schemes are shown in Figure 4.10 for $M = 2$ and $M = 4$. It is clear that MRC achieves the best performance. EGC error performance is within 1 dB from MRC. As M increases, the difference between the three schemes increases. This is clear between MRC and SC. Also, it is obvious that the upper bound is very tight and gives very accurate values for the bit error probability, especially, at bit error rates less than 10^{-3}.

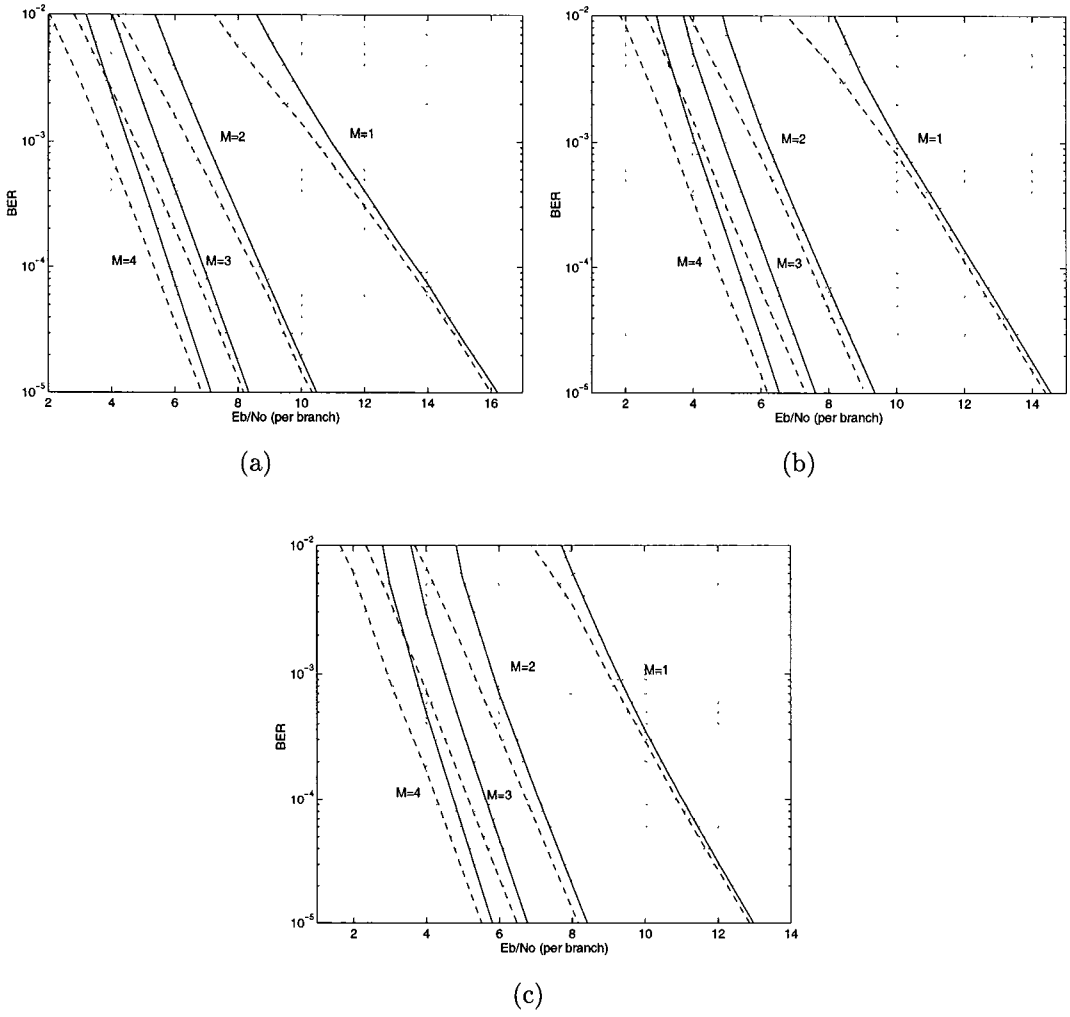


Figure 4.9: BER of 16-QAM I-Q TCM 4,8, and 16-state codes with selection combining and different diversity orders. solid(bound), dashed(simulation), (a) $\nu = 2$ (b) $\nu = 3$ (c) $\nu = 4$

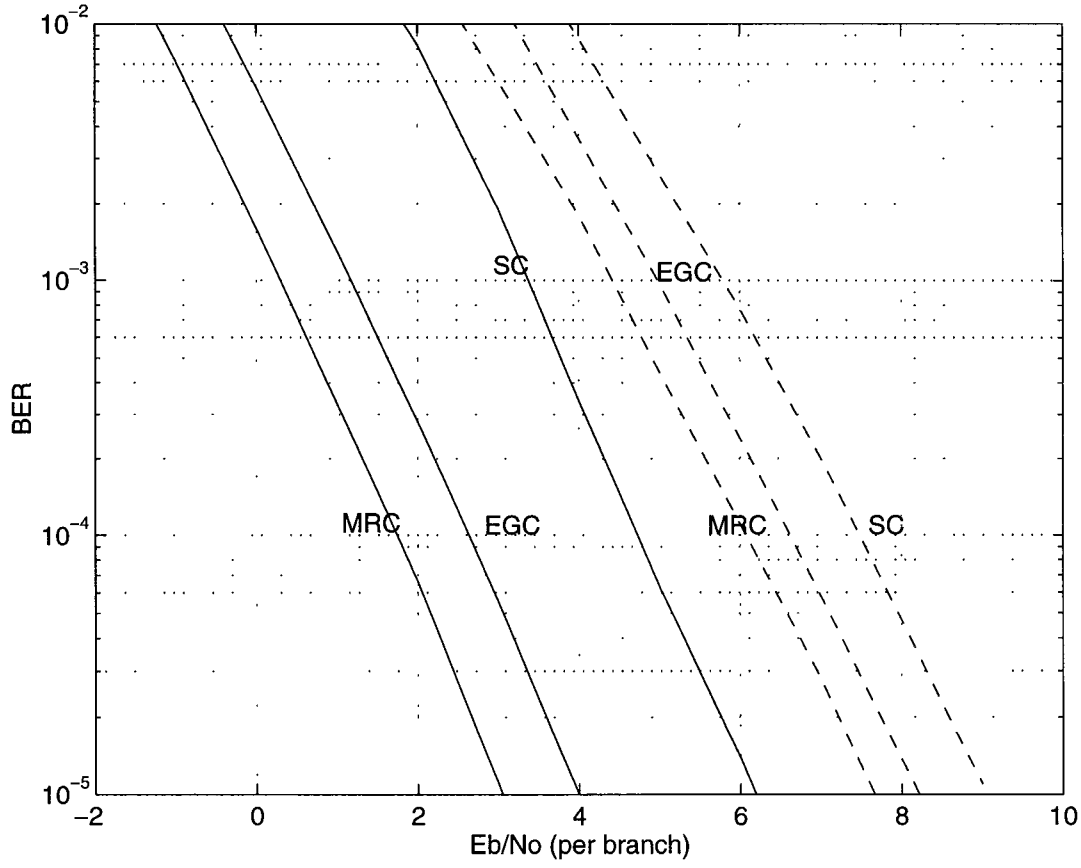


Figure 4.10: Simulated BER of 16-QAM I-Q TCM 8-state code with different combining schemes. dashed ($M = 2$), solid ($M = 4$).

4.5 Performance of Uncoded Systems

The derived bounds can also be used to evaluate the performance of uncoded systems. In uncoded BPSK, the error event length is $L_\eta = 1$. This means that the pairwise error probability is equal to the bit error probability, P_b.

- For MRC receivers, $d_i = E_s/N_o$. Therefore,

$$\delta_m = \delta_i = \frac{E_s/N_o}{1 + E_s/N_o} \tag{4.95}$$

and

$$\frac{1}{1+d_i} = 1 - \delta_i = (1 - \sqrt{\delta_i})(1 + \sqrt{\delta_i}). \tag{4.96}$$

Since $\delta_m = \delta_i$, the upper bound is satisfied with equality in this case. Therefore, the bit error probability can be expressed as

$$\begin{aligned} P_b &= \frac{1}{2^{2M}} \sum_{j=1}^{M} \begin{pmatrix} 2M-j-1 \\ M-1 \end{pmatrix} \left(\frac{2}{1+\sqrt{\delta_i}}\right)^j \left(\frac{1}{1+d_i}\right)^M \\ &= \left(\frac{1-\sqrt{\delta_i}}{2}\right)^M \sum_{j=1}^{M} \begin{pmatrix} 2M-j-1 \\ M-j \end{pmatrix} \left(\frac{1+\sqrt{\delta_i}}{2}\right)^{M-j}. \end{aligned} \tag{4.97}$$

If we define $k = M - j$, then P_b can be written as

$$P_b = \left(\frac{1-\sqrt{\delta_i}}{2}\right)^M \sum_{k=0}^{M-1} \begin{pmatrix} M+k-1 \\ k \end{pmatrix} \left(\frac{1+\sqrt{\delta_i}}{2}\right)^k. \tag{4.98}$$

This is exactly the same expression in [1].

- For EGC receivers, $\tilde{d}_i = \frac{b_o E_s}{M N_o}$. Therefore,

$$\tilde{\delta}_e = \tilde{\delta}_i = \frac{\tilde{d}_i}{1 + \tilde{d}_i} \tag{4.99}$$

and

$$\frac{1}{1+\tilde{d}_i} = 1 - \tilde{\delta}_i = (1-\sqrt{\tilde{\delta}_i})(1+\sqrt{\tilde{\delta}_i}). \tag{4.100}$$

Therefore, the bit error probability can be expressed as

$$P_b = \frac{1}{2^{2M}} \sum_{j=1}^{M} \binom{2M-j-1}{M-1} \left(\frac{2}{1+\sqrt{\tilde{\delta}_{min}}}\right)^j \left(\frac{1}{(1+\tilde{d}_i)}\right)^M, \tag{4.101}$$

which again can be written as

$$P_b = \left(\frac{1-\sqrt{\tilde{\delta}_i}}{2}\right)^M \sum_{k=0}^{M-1} \binom{M+k-1}{k} \left(\frac{1+\sqrt{\tilde{\delta}_i}}{2}\right)^k. \tag{4.102}$$

- For selection combining receivers, $d_i = E_s/N_o$.

$$\delta_{i,k_i} = \frac{d_i}{k_i+d_i} \qquad \text{and} \qquad \frac{k_i}{k_i+d_i} = (1-\sqrt{\delta_{i,k}})(1+\sqrt{\delta_{i,k}}) \tag{4.103}$$

Therefore,

$$P_b = \frac{1}{2} \sum_{k_i=1}^{M} M(-1)^{k_i+1} \binom{M-1}{k_i-1} \frac{1}{(k_i+d_i)} \frac{1}{(1+\sqrt{\delta_{i,k_i}})}, \tag{4.104}$$

which can be finally written as

$$P_b = \frac{1}{2} \sum_{k=1}^{M} (-1)^{k+1} \binom{M}{k} (1-\sqrt{\frac{d_i}{k+d_i}}). \tag{4.105}$$

Note that in all the previous cases when $M = 1$ the formula reduces to

$$P_b = \frac{1}{2}\left(1 - \sqrt{\frac{E_b/N_o}{1+E_b/N_o}}\right) \tag{4.106}$$

which is exactly as in [1]. Figure 4.11 shows P_b values for the three combining schemes with diversity orders of $M = 1, 2, 3, 4$.

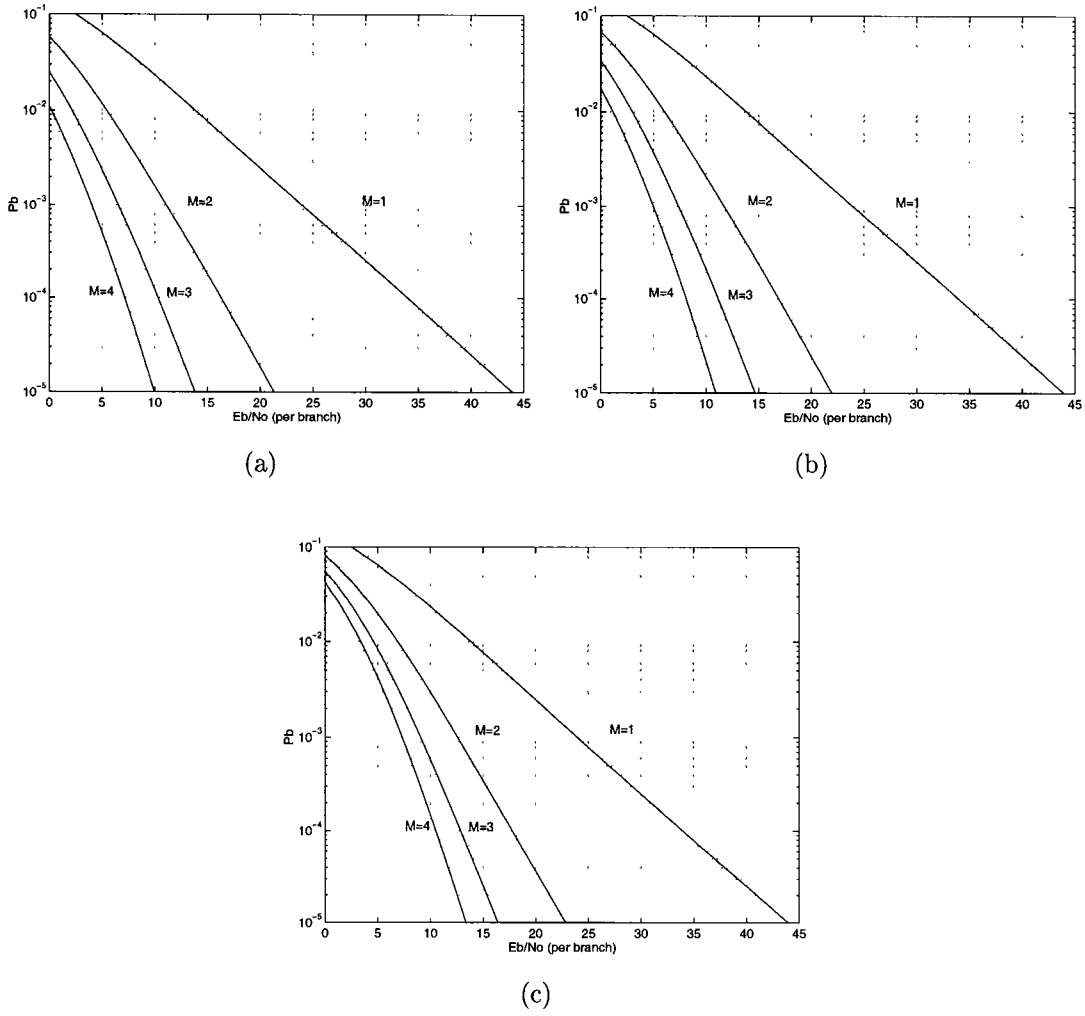


Figure 4.11: Bit error probability of uncoded BPSK with different diversity orders and combining schemes (a) MRC (b) EGC (c) SC

4.6 The Effect of Branch Correlation

Up to now, we have assumed that the fading in the different diversity branches are independent. In some cases, this is difficult to achieve due to improper antenna positioning or receiver space limitations. Therefore, it is important to examine the possible degradation in the system performance when the branch signals are correlated. The effect of branch correlation on the distribution of the received signal was studied by Schwartz et. al. in 1966 [65]. Recently, the effect of correlation on non-coherent orthogonal digital modulation was studied [73]. They derived upper bounds for binary convolutional codes and non-coherent orthogonal digital modulation.

In this section the pairwise error probability for maximal ratio combining with correlated branch signals is derived. Recall from Eqn. 4.46 that the conditional pairwise error probability may be expressed as

$$P(\mathbf{x}_N \rightarrow \hat{\mathbf{x}}_N|\mathbf{a}_N) = \frac{1}{2}\text{erfc}\left(\sqrt{\sum_{i=1}^{N}\sum_{l=1}^{M} a_{il}^2 d_i}\right) \tag{4.107}$$

where $d_i = \frac{|x_i - \hat{x}_i|^2}{4N_o}$. Also, $a_{il}^2 = |h_{il}|^2$ where h_{il} is a complex Gaussian random variable with zero mean and variance of $1/2$ for both the real and imaginary parts, and $|\cdot|$ denotes its envelope. Observe that

$$\sum_{l=1}^{M} a_{il}^2 = \underline{h}_i \underline{h}_i^* \tag{4.108}$$

where $\underline{h}_i = \{h_{i1}, \ldots, h_{iM}\}$ and $(\cdot)^*$ denotes the Hermitian transpose. The probability density function of $\underline{h}_i$ is expressed as

$$f(\underline{h}_i) = \frac{1}{\pi^N \det K_{\underline{h}_i}} \exp(-\underline{h}_i K_{\underline{h}_i}^{-1} \underline{h}_i^*), \tag{4.109}$$

where $K_{\underline{h}_i}$ is an $N \times N$ covariance matrix with entries $(K_{\underline{h}_i})_{lk} = E(h_{il} h_{ik}^*)$. It is

also assumed that the real parts of h_{il}'s are independent of the imaginary parts; i.e., the cross covariances are zero.

The first step is to uncorrelate the random variables using a linear transformation. We are interested in generating a new vector $\underline{g}_i = \{g_{i1}, \ldots, g_{iM}\}$ with a diagonal covariance matrix. Let U be the transformation - i.e.,

$$\underline{g}_i = U\underline{h}_i. \tag{4.110}$$

Using this transformation, the covariance matrix of $\underline{g}_i$, denoted by $K_{\underline{g}_i}$ is expressed as follows

$$K_{\underline{g}_i} = UK_{\underline{h}_i}U^t, \tag{4.111}$$

where U^t is the transpose of U. Since $K_{\underline{h}_i}$ is a symmetric matrix, it can be represented in the following form

$$K_{\underline{h}_i} = Q\Lambda Q^t, \tag{4.112}$$

where Λ is a diagonal matrix that consists of the eigenvalues of $K_{\underline{h}_i}$ and Q is a matrix whose columns are the orthonormal set of the eigenvectors of $K_{\underline{h}_i}$. The last equation can be rewritten as

$$\Lambda = Q^t K_{\underline{h}_i} Q. \tag{4.113}$$

Therefore, if we let $U = Q^t$ then the Gaussian random variables $\underline{g}_i$ are independent and $K_{\underline{g}_i} = \Lambda$; i.e.,

$$K_{\underline{g}_i} = \begin{cases} \lambda_j & \text{i=j} \\ 0 & \text{otherwise} \end{cases} \tag{4.114}$$

The second step is to make another transformation so that the covariance matrix becomes the identity matrix. This is achieved via the following transformation

$$\underline{g}_i = \sqrt{K_{\underline{g}_i}}\,\underline{p}_i. \tag{4.115}$$

This makes the Gaussian random variables $\underline{p}_i$ independent with same variance. Using this transformation, the unconditional pairwise error probability can be expressed as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) = \tfrac{1}{2}\int_0^\infty \cdots \int_0^\infty \operatorname{erfc}\left(\sqrt{\textstyle\sum_{i=1}^N \sum_{l=1}^M \lambda_l q_{il} d_i}\right) \times \\ f(q_{11}) \cdots f(q_{1M}) \cdots f(q_{NM}) dq_{11} \cdots dq_{1M} \cdots dq_{NM}, \tag{4.116}$$

where $q_{il} = |p_{il}|^2$. Define

$$\tilde{\delta}_{il} = \frac{\lambda_l d_i}{1 + \lambda_l d_i} \qquad \text{and} \qquad \tilde{\omega}_{il} = q_{il}(1 + \lambda_l d_i). \tag{4.117}$$

Hence,

$$f(\tilde{\omega}_{il}) d\tilde{\omega}_{il} = \frac{1}{1 + \lambda_l d_i} e^{\tilde{\delta}_{il}\tilde{\omega}_{il}} e^{-\tilde{\omega}_{il}} d\tilde{\omega}_{il} = \frac{1}{1 + \lambda_l d_i} e^{\tilde{\delta}_{il}\tilde{\omega}_{il}} f(\tilde{\omega}_{il}) d\tilde{\omega}_{il}, \tag{4.118}$$

where $f(\tilde{\omega}_{il})$ is the exponential distribution of $\tilde{\omega}_{il}$ with parameter one. Hence, the unconditional pairwise error probability can be expressed as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) = \tfrac{1}{2} \textstyle\prod_{i\in\eta} \prod_{l=1}^M \left(\frac{1}{1+\lambda_l d_i}\right) \int_0^\infty \cdots \int_0^\infty \operatorname{erfc}\left(\sqrt{\textstyle\sum_{i=1}^N \sum_{l=1}^M \tilde{\delta}_{il}\tilde{\omega}_{il}}\right) \times \\ f(\tilde{\omega}_{11}) \cdots f(\tilde{\omega}_{1M}) \cdots f(\tilde{\omega}_{NM}) d\tilde{\omega}_{11} \cdots d\tilde{\omega}_{1M} \cdots d\tilde{\omega}_{NM} \tag{4.119}$$

where $\eta = \{i : x_i \neq \hat{x}_i\}$ and L_η is its cardinality. However,

$$\sum_{i\in\eta} \sum_{l=1}^M \tilde{\delta}_{il}\tilde{\omega}_{il} \geq \sum_{i\in\eta} \sum_{l=1}^M \delta_c \tilde{\omega}_{il}, \tag{4.120}$$

where $\delta_c = \min\{\tilde{\delta}_{il}, i \in \eta, l = 1, \cdots M\}$ Hence, the pairwise error probability can be upper bounded as

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \frac{1}{2} \prod_{i\in\eta} \prod_{l=1}^M \frac{1}{(1 + \lambda_l d_i)} \int_0^\infty \operatorname{erfc}\left(\sqrt{\delta_c \Phi}\right) \times e^{(\delta_c \Phi)} f(\Phi) d\Phi, \tag{4.121}$$

where $\Phi = \sum_{i\in\eta} \sum_{l=1}^M \tilde{\omega}_{il}$. Since the $\tilde{\omega}_{il}$'s are independent exponentially distributed random variables each with parameter one, Φ will have an (ML_η)-Erlang distribution with parameter one – i.e.,

$$f(\Phi) = \tfrac{1}{(ML_\eta - 1)!} \Phi^{(ML_\eta - 1)} e^{-\Phi} \quad \Phi \geq 0. \tag{4.122}$$

Finally performing the integration yields

$$P(\mathbf{x}_N \to \hat{\mathbf{x}}_N) \leq \left[\frac{1}{2^{2ML_\eta}} \sum_{j=1}^{ML_\eta} \begin{pmatrix} 2ML_\eta - j - 1 \\ ML_\eta - 1 \end{pmatrix} \left(\frac{2}{1+\sqrt{\delta_c}} \right)^j \right] \times \prod_{i \in \eta} \prod_{l=1}^{M} \frac{1}{(1+\lambda_l d_i)}. \tag{4.123}$$

Therefore, the bit error probability can be expressed as

$$P_b \leq \frac{1}{k} \left[\frac{1}{2^{2ML}} \sum_{j=1}^{ML} \begin{pmatrix} 2ML - j - 1 \\ ML - 1 \end{pmatrix} \left(\frac{2}{1+\sqrt{\delta_c}} \right)^j \right] \frac{\partial T(\bar{D}, I)}{\partial I} \mid_{I=1, D=e^{-E_s/4N_o}}, \tag{4.124}$$

where

$$\bar{D} \mid_{D=e^{-E_s/4N_0}} = \prod_{j=1}^{M} \frac{1}{(1+\lambda_j \frac{E_s}{4N_o})}. \tag{4.125}$$

Similarly, the cutoff rate can be expressed as

$$R_o = 2\log_2(|A|) - \log_2 \left(\sum_{x_i \in A} \sum_{x_j \in A} \left[\prod_{l=1}^{M} \frac{1}{\left(1+\lambda_l \frac{|x_i - \hat{x}_i|^2}{4N_o}\right)} \right] \right). \tag{4.126}$$

4.6.1 Example

In many practical systems, dual diversity is used. The covariance matrix $K_{\underline{h}_i}$ is represented as

$$K_{\underline{h}_i} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \tag{4.127}$$

In other words, ρ is the correlation coefficient between the two antenna elements. It is straightforward to show that the eigenvalues of $K_{\underline{h}_i}$ are $(1-\rho)$ and $(1+\rho)$. The 4-state I-Q TCM 16-QAM scheme (2 bits/s/Hz) is used as an example. For this configuration, δ_c will be

$$\delta_c = \frac{(1-\rho)\frac{0.8E_s}{N_o}}{1+(1-\rho)\frac{0.8E_s}{N_o}}. \tag{4.128}$$

A comparison between the bound and simulations are shown in Figure 4.12 for the case of $\rho = 0.5$. Clearly, the bound is very tight. Figure 4.13 shows the bit error probability upper bound curves for the mentioned code with different values of ρ. It should be pointed out that values as large as $\rho = 0.5$ degrade the performance slightly. The effect of space correlation is not as severe as time correlation which is minimized via interleaving.

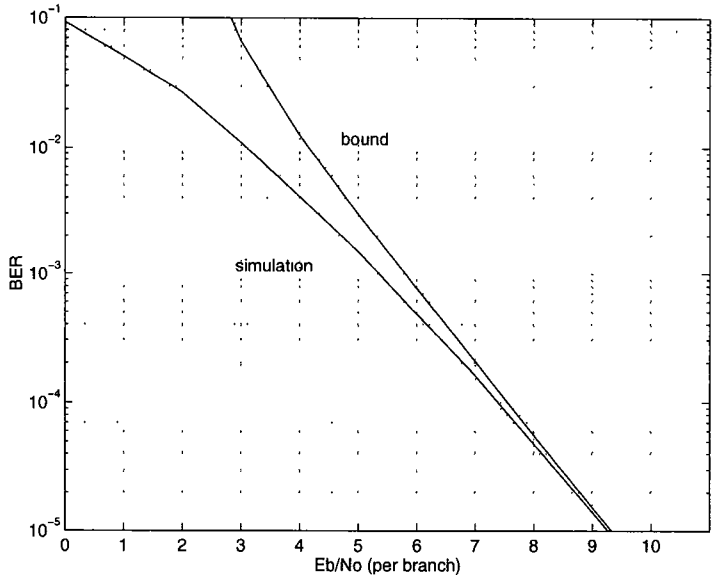


Figure 4.12: A comparison between the upper bound and simulated BER of 16-QAM I-Q TCM 4-state code with MRC dual diversity and $\rho = 0.5$.

4.7 Comparisons and Complexity Tradeoffs

In this section, we compare the performance of an I-Q code with a traditionally designed trellis code. Figure 4.14 shows a comparison between the I-Q 16-QAM and the 8-PSK TCM 8-state codes. MRC is used and the diversity order is $M = 2$. Even in double diversity environments, the I-Q code outperforms the

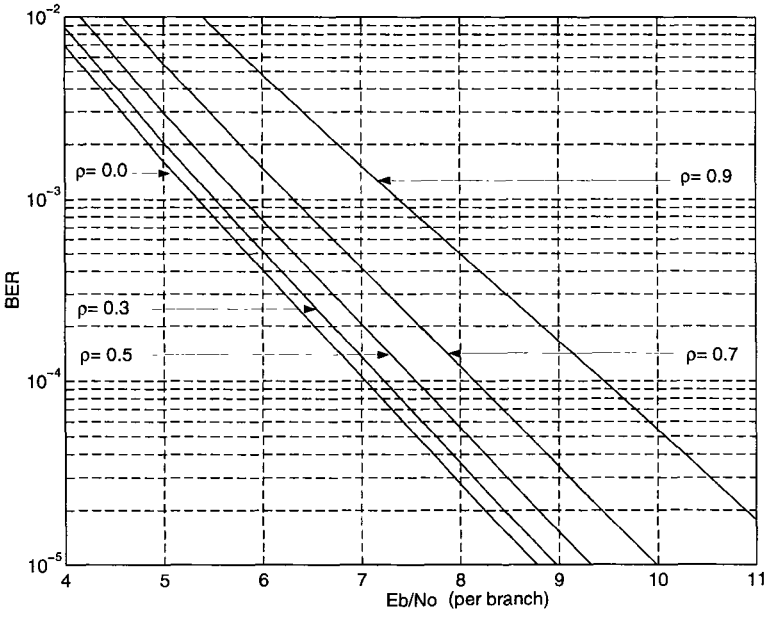


Figure 4.13: Analytical BER of 16-QAM I-Q TCM 4-state code with MRC dual diversity and different correlation values.

8-PSK code by about 1.5 dB at a BER=10^{-5}.

A comparison between three schemes is shown in Figure 4.15. The first scheme uses a 4-state code and maximal ratio combining of two branches. The branch correlation ρ is assumed to be 50%. The second scheme uses a 64-state code but no diversity combining is employed. Both schemes are 16-QAM. The third scheme employs diversity only. It uses uncoded QPSK with Gray mapping. The diversity order (M) is 3. All systems have a bandwidth efficiency of 2 b/s/Hz. Simulation results are plotted for the first two systems and analytical values are shown for the third system. Clearly, the first scheme outperforms the other schemes at BER $< 4 \times 10^{-3}$ even though moderate branch correlation (50%) exists. This suggests that a combination of simple channel coding and

space diversity might yield in general better performance than using complex channel coding schemes or several diversity receivers. Moreover, increased delay and interleaving for complicated channel codes is avoided.

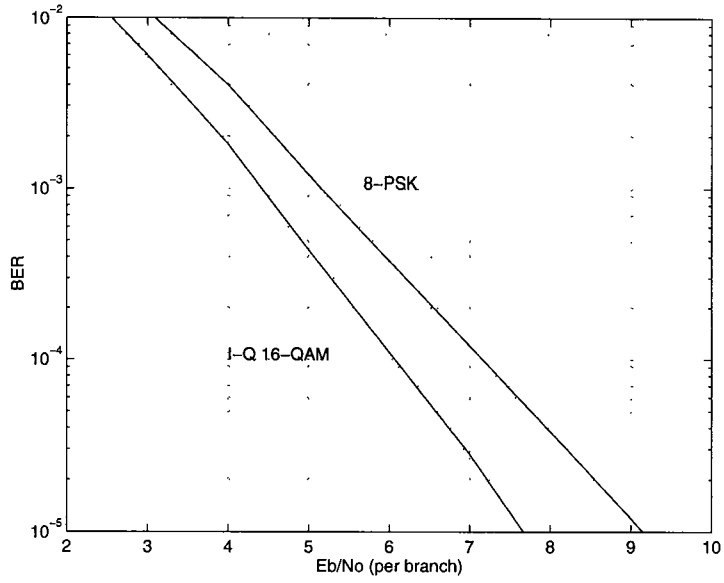


Figure 4.14: Simulated BER of 16-QAM I-Q TCM 8-state code and the 8-PSK TCM 8-state code with MRC and double diversity.

4.8 Summary

In this chapter, cutoff rate expressions of coherent systems with maximal ratio, equal gain, and selection combining schemes have been evaluated using Chernoff bounds. Moreover, tight upper bounds on the pairwise error probability have been derived. These upper bounds were used to evaluate a variety of system configurations, including uncoded and coded systems. The upper bounds were expressed in product form to allow the use of the transfer function approach for evaluating the performance of trellis coded systems. Simulations of different

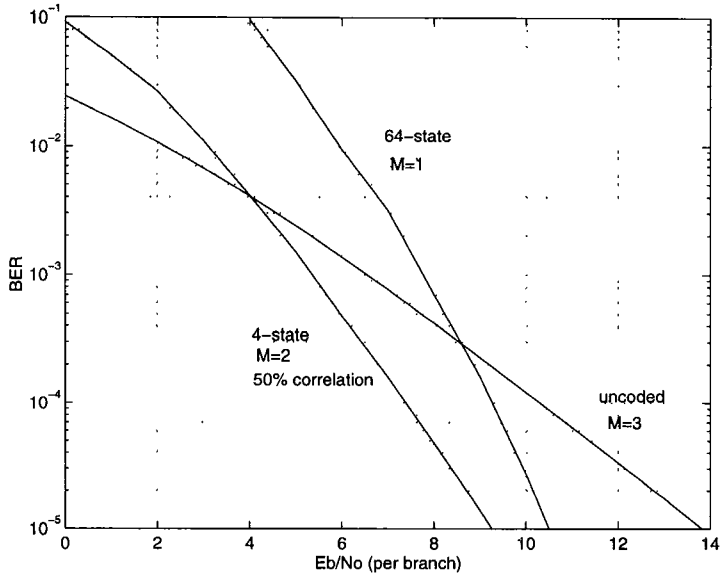


Figure 4.15: Copmarison between the 16-QAM I-Q TCM 4-state code (with double diversity, MRC and 50% branch correlation) and the 64-state code (with no diversity).

systems show that the derived bounds are very tight.

For the case of branch correlation, the cutoff rate and a tight upper bound on the pairwise error probability were derived for maximal ratio combining. Also, the pairwise error probability is expressed in product form and the transfer function approach is used. Branch correlation with correlation coefficients less than 0.5 result in a slight performance loss. The results indicate that the joint use of coding and diversity results in a substantial improvement in the Rayleigh fading channel.

Chapter 5

Sequence MAP Decoding

5.1 Introduction

Traditionally, source coding (data compression) and channel coding (error control) are designed independently of one another. The reason for this is based on the separation principle presented by Shannon [33], which states the no performance degradation is suffered if the two functions are thus partitioned. However, this principle is an asymptotic result with unlimited delay and complexity. Given a limited decoding complexity/delay, joint source-channel coding could outperform separately designed pairs.

Some recent work in this area has investigated the design of source codes that are robust to channel errors [74, 75, 76]. On the other hand, the work in [77] and [78] has concentrated on the design of channel decoders that exploit the known characteristics of the source code. The work here focuses on the performance of trellis codes with sequence maximum a posteriori (MAP) decoding of correlated signals transmitted over very noisy AWGN and Rayleigh channels.

An ideal source encoder would compress an audio, image or a data signal and produce an independent, identically distributed (i.i.d.) sequence of bits at the output. Most source coding algorithms are not ideal; the output bit stream contains residual redundancy in the form of memory and/or non-uniform distribution. Hence, the number of bits produced per sample (or unit time) is significantly greater than the entropy rate of the source encoder's output.

This residual redundancy may be exploited at the receiver by adjusting the Viterbi algorithm decoding metric to use the source's a priori probabilities. Sequence MAP Viterbi decoders take advantage of this residual redundancy to enhance the performance of the system under very noisy channel conditions. This decoding method used in conjunction with other means (i.e., soft-decision decoding and channel state information estimation) results in a very robust system under bad channel conditions.

The Viterbi algorithm selects the maximum likelihood sequence as its estimate; it does not minimize the probability of error of the data bits [79, 80, 81]. However, its performance is very close to that of the optimal symbol-by-symbol decoding algorithm [79]. Furthermore, although the existing symbol-by-symbol decoding algorithms apply only to linear codes [79], the Viterbi algorithm is applicable to linear and non-linear codes. Sequence MAP decoding does not require substantial modifications of the existing decoders for trellis coded system, and as indicated in [77, 78], may be used if needed where a bad channel environment exists. In [77], Hagenauer showed that 2-3 dB gains could be obtained for PCM transmission and the full rate *group special mobile* (GSM) speech codec. A 16-state rate 1/2 convolutional code with BPSK modulation and a dynamic two state Markov correlation estimator were used. In [78], Alajaji, Phamdo and

Fuja used both block and convolutional codes to exploit the residual redundancy in the CELP speech encoder. A 32-state rate 3/4 convolutional code with BPSK modulation and sequence MAP decoding was used. The receiver was supplied with the source a priori information which was obtained using first and second order Markov models on a large training sequence from the TIMIT [82] speech database. Decoding gains of 2-5 dB were obtained.

The work in this chapter is divided into two main parts. The first part assumes a simple ideal first order two state Markov source model. This simple model is chosen because its a priori information may be easily estimated at the receiver during periods of good channel conditions. A variety of different systems with different sources, modulation schemes and trellis code complexities are simulated. Extensive simulations of these system configurations are performed to assess their effect on the sequence MAP decoding gains* (simply called MAP gains) over ML decoding. In performing such simulations, we try to address the following questions

- What is the effect of increasing the trellis encoder memory?
- What is the effect of increasing the encoder rate?
- What is the effect of increasing the signal constellation?
- What is the effect of redundancy type?
- What is the difference in MAP decoding gains for coded as well as uncoded systems?

We use the term "decoding gains" instead of "coding gains" since we are comparing two systems with the same coding schemes but the decoders are different

- What is the effect of the channel type (AWGN or Rayleigh) on the MAP decoding gains?

The first part is concluded with a comparative example between a 2-D and a 4-D QPSK trellis coded modulation schemes with the same spectral efficiencies and ML asymptotic coding gains. It is shown that increasing the dimensionality (i.e., multiplicity) of the code results in increased decoding gains.

In the second part, we consider a practical example for coding the CELP line spectral parameters (LSP's) using trellis codes with 4-D QPSK modulation. Two source models are used. One is based on the intra-frame correlation while the second one models both intra-frame and inter-frame correlations. Both AWGN and Rayleigh channels are considered.

5.2 MAP decoding for Ideal Sources

5.2.1 System Model

The basic model block diagram is shown in Figure 5.1. The information bits from the source , u_i's, are modeled by a stationary first order two-state Markov process as shown in Figure 5.2. The sequence $\{u_i\}$ represents the output of a source encoder, or (if the source is not compressed) the output of the source itself. We denote the transition probabilities $\Pr(u_i = 0|u_{i-1} = 0)$ and $\Pr(u_i = 1|u_{i-1} = 1)$ by $\Pr(0|0)$ and $\Pr(1|1)$, respectively.

The entropy of the Markov chain can be computed to arrive at an estimate of the source redundancy. Let $H_\infty(U)$ be the source entropy rate and let $H(U)$ be the entropy of a memoryless source with the same marginal distribution as the source. Define [83]

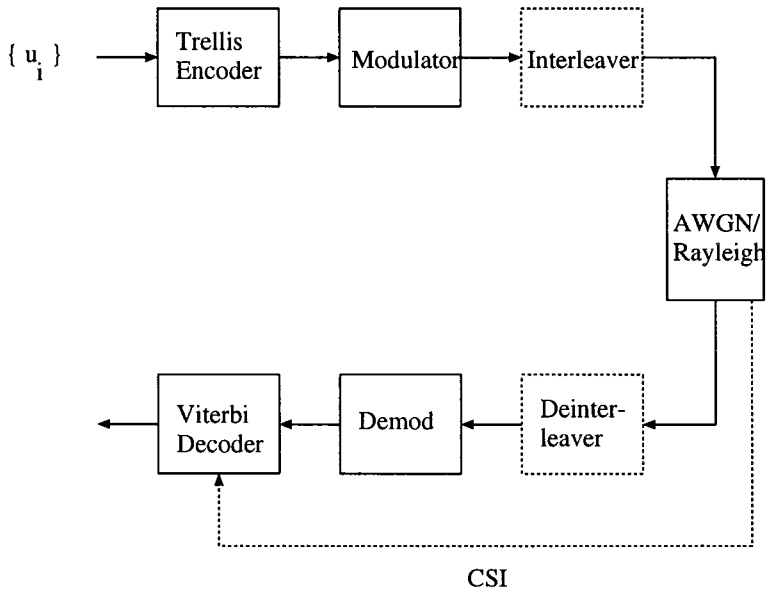


Figure 5.1: Block diagram of the system model.

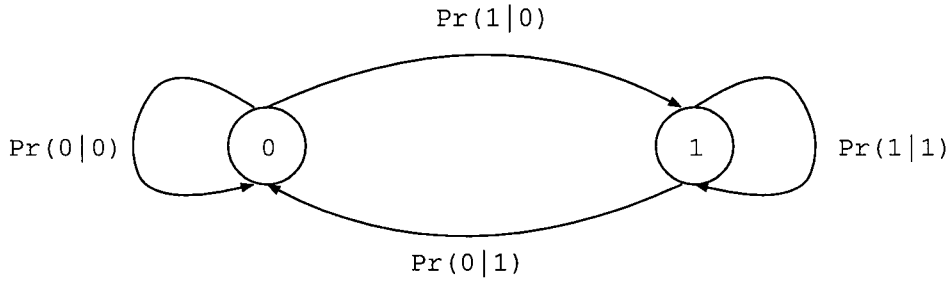


Figure 5.2: Two-state binary Markov source model.

$$\begin{aligned} \rho_D &\triangleq 1 - H(U) \\ \rho_M &\triangleq H(U) - H_\infty(U) \\ \rho_T &\triangleq \rho_D + \rho_M = 1 - H_\infty(U). \end{aligned} \tag{5.1}$$

In other words, ρ_D denotes the redundancy due to the non-uniform distribution of the source, ρ_M denotes the redundancy due to the source memory and ρ_T denotes the total redundancy of the source. We assume that the source has $\rho_T > 0$. The two forms of redundancy are utilized by the decoder to combat channel errors.

The source bits are arranged in a sequence of binary k-tuples $\{\underline{u}_1, \underline{u}_2, \ldots\}$. At time i, $\underline{u}_i$ is an input to a trellis encoder which produces a $(k+1)$-tuple of binary bits $\underline{c}_i$. The trellis encoder output $\underline{c}_i$ is mapped to $\underline{x}_i = \{x_i^1, \ldots, x_i^m\}$, a complex vector in general, where m is the number of transmitted signals per trellis branch (baud);i.e., the multiplicity of the code. For example for rate 1/2 trellis code with BPSK modulation $m = 2$. The sequence $\mathbf{x}_N = \{\underline{x}_1, \ldots \underline{x}_N\}$ is transmitted over the channel. This can be described by the discrete time relation as

$$\begin{aligned} y_i^1 &= a_i^1 x_i^1 + n_i^1 \\ y_i^2 &= a_i^2 x_i^2 + n_i^2 \\ &\cdot \\ &\cdot \\ &\cdot \\ y_i^m &= a_i^m x_i^m + n_i^m \end{aligned} \tag{5.2}$$

where n_i^j is a two-dimensional zero mean additive Gaussian noise with a single-sided power spectral density of N_o. The distribution of a_i^j depends on the channel assumption

- For a purely AWGN non-fading channel, $a_i^j = 1$.

- For a fully interleaved Rayleigh fading channel we assume $\{a_i^j\}$ is a sequence of i.i.d. Rayleigh distributed random variables with $E[(a_i^j)^2] = 1$. Note that the assumption of full interleaving is fully justified since under very noisy channel conditions (the region of interest in this work) interleaving requirements are much less than at high SNR's. In other words, at low SNR the additive noise plays a more dominant role than the possible correlation in the fading.

5.2.2 MAP decoding rule

The sequence MAP decision rule is to choose $\hat{\mathbf{x}}_N$ which maximizes

$$f(\mathbf{y}_N|\hat{\mathbf{x}}_N)\Pr(\hat{\mathbf{x}}_N).$$

For the AWGN channel, this reduces to choosing $\hat{\mathbf{x}}_N$ which minimizes

$$\sum_{i=1}^{N}\sum_{l=1}^{m}|y_i^l - \hat{x}_i^l|^2 - N_o \ln \Pr(\hat{\mathbf{x}}_N).$$

Using the chain rule $\ln \Pr(\mathbf{x}_N)$ can be expressed as

$$\ln \Pr(\mathbf{x}_N) = \sum_{i=1}^{N} \ln \Pr(\underline{x}_i|\underline{x}_{i-1}, \underline{x}_{i-2}, \ldots). \tag{5.3}$$

However, because of the first order Markov property, we can further simplify this to

$$\ln \Pr(\underline{x}_i|\underline{x}_{i-1}, \underline{x}_{i-2}, \ldots) = \ln \Pr(\underline{u}_i|\underline{u}_{i-1}). \tag{5.4}$$

Therefore, the MAP decoding rule is to choose $\hat{\mathbf{x}}_N$ which minimizes

$$\sum_{i=1}^{N}\left(\sum_{l=1}^{m}|y_i^l - \hat{x}_i^l|^2 - N_o \Pr(\underline{u}_i|\underline{u}_{i-1})\right).$$

For the Rayleigh distributed fading channel, CSI is incorporated in the decoding metric to choose $\hat{\mathbf{x}}_N$ which minimizes

$$\sum_{i=1}^{N}\left(\sum_{l=1}^{m}|y_i^l - a_i^l\hat{x}_i^l|^2 - N_o\Pr(\underline{u}_i|\underline{u}_{i-1})\right).$$

5.2.3 Bit Error Probability Upper Bound

Using the union bound approach, the average bit error probability, P_b, of trellis coded systems is upper bounded as [61]

$$P_b \leq \sum_{n=1}^{\infty}\sum_{\mathbf{x}_n}\sum_{\mathbf{x}_n\neq\hat{\mathbf{x}}_n} N_n(\mathbf{x}_n,\hat{\mathbf{x}}_n)\Pr(\mathbf{x}_n)P(\mathbf{x}_n\rightarrow\hat{\mathbf{x}}_n) \tag{5.5}$$

where $P(\mathbf{x}_n \rightarrow \hat{\mathbf{x}}_n)$ is the pairwise error probability between the sequences $\mathbf{x}_n$ and $\hat{\mathbf{x}}_n$, $\Pr(\mathbf{x}_n)$ is the a priori probability of transmitting the sequence $\mathbf{x}_n$, and $N_n(\mathbf{x}_n,\hat{\mathbf{x}}_n)$ is the number of bit errors when $\mathbf{x}_n$ is the transmitted sequence and $\hat{\mathbf{x}}_n$ is chosen by the decoder.

In enumerating the trellis paths, the super-state diagram with $2^{2\nu}$ states is used, where ν is the number of memory elements of the minimal encoder [61]. It is not possible to use the state diagram of the code for path enumeration even if the code satisfies the uniformity properties. This is because not all sequences are equiprobable and hence the paths will contribute *unequally* to the upper bound of the bit error probability.

However, the union bound with either maximum likelihood (ML) [84] or MAP [77] decoding is loose at high decoding error rates ($\sim 10^{-2}$). It does not provide useful quantitative values for the MAP decoding gains. Recently, an improved upper bound on the error event probability, P_e, for the Viterbi decoder over the binary symmetric channel (BSC) was presented [85]. At high error rates, this upper bound does not give trivial values for P_e, as in the case of the standard

Source	Pr(0|0)	Pr(1|1)	ρ_D	ρ_M	ρ_T
I	0.8	0.2	0.278	0.0	0.278
II	0.8	0.8	0.0	0.278	0.278
III	0.88	0.65	0.180	0.187	0.367
IV	0.9	0.8	0.082	0.365	0.447
V	0.2	0.9	0.496	0.007	0.503

Table 5.1: The five source models and their redundancies

union bound. Also, it may be deduced from [85] that a higher length of an error event path contributes to the reduction of P_e. In this work we resort to computer simulations to accurately asses the performance of the MAP decoder.

5.3 Simulated Systems

Different configurations of systems were simulated. The source is assumed to have one of five distributions. Table 5.1 shows these distributions with their redundancies. The first source is a non-uniformly distributed i.i.d. process. The second is a symmetric binary Markov source. These two sources have the same total redundancy but in completely different forms; the i.i.d. source exhibits all its redundancy in its non-uniform distribution, while the symmetric Markov source exhibits its redundancy in the form of memory. The last three distributions are asymmetric Markov sources. Source III has its total redundancy almost split equally in the forms of memory and distribution, while sources IV and V are highly skewed to one form of redundancy.

BPSK, QPSK and 8-PSK modulation schemes were used. Trellis codes with

different encoder rates and complexities were implemented. These different configurations are tested on two channel models – namely, the pure non-fading AWGN model and the Rayleigh-distributed fading channel model with ideal interleaving. These two channel models describe the extreme cases of channels encountered in practice. Hence, the results obtained over other channels (e.g., Rician) will lie in between our results. The simulated systems are described as follows

For comparison purposes, uncoded BPSK modulated systems were simulated. Five different sources were used over both the Gaussian and Rayleigh channels. A two-state Viterbi decoder with sequence MAP detection was used. MAP detection gains, compared to symbol-by-symbol ML detection, were highly affected by the type of redundancy. Uncoded systems with redundancy in the form of memory have higher MAP decoding gains than uncoded systems having all or most of their redundancy in the form of non-uniform distribution.

In coded systems, the chosen codes are optimized for the channel model assumed. The code design criterion for the AWGN channel is to maximize the minimum Euclidean distance of the code. On the other hand, for the Rayleigh channel, the design criterion is to maximize the minimum time diversity of the code; i.e., the code's minimum Hamming distance in signal symbols.

5.3.1 BPSK Modulated Systems

For systems with BPSK modulation, two families of trellis codes were used, namely rate 1/2 and rate 2/3 codes. Both families of codes were obtained from the tables of best convolutional codes (in terms of the Hamming metric) [63]. For trellis codes with BPSK modulation, there is a one-to-one correspondence be-

tween the code's minimum Euclidean distance, d_{free}, and its minimum Hamming distance, d_H – i.e.,

$$d_{free} = 4E_s d_H \tag{5.6}$$

Therefore, the codes are optimized for both channels. Rate 1/2 codes with 4, 8, and 16-state encoders were used for the five source models. Their minimum Hamming distances are 5, 6, and 7, respectively. Their minimum normalized Euclidean distances (i.e., d_free/E_s) are accordingly 20, 24 and 28, respectively. The bandwidth efficiency of these codes are 0.5 bit/s/Hz. A decoding buffer length of 8ν was used for these codes, where ν is the number of memory elements in the encoder. At a decoding bit error rate (BER) of 0.02, MAP decoding gains (over ML decoding) are as high as 1.4 dB for the Gaussian channel and 2 dB for the Rayleigh channel for a correlated source with about 50% source redundancy.

Rate 2/3 codes with 4, 8, and 16-state encoders were also used to show the effect of increasing the encoder rate on the MAP decoding gains. Higher MAP decoding gains were obtained. For example, at a decoding BER of 0.02 gains (MAP vs. ML) as high as 2.5 dB for the Gaussian channel and 4 dB for the Rayleigh channel for a correlated source with about 50% source redundancy.

5.3.2 QPSK Modulated Systems

For QPSK modulated systems, the same rate 1/2 trellis codes mentioned above were used. However, each 2-tuple of bits is mapped to two-dimensional (2-D) QPSK signal points. Gray mapping was used so that maximizing the minimum Hamming distance corresponds to maximizing the minimum Euclidean distance. Moreover, using this approach the codes are optimized for both channels. The spectral efficiency in this case is 1 b/s/Hz. Encoders with 4, 8, and 16 states

were used. Their minimum Hamming distances (in symbols) are 3, 4, and 5, respectively. Moreover, the codes' normalized minimum Euclidean distances are 10, 12 and 14, respectively. A decoding buffer length of 8ν was used for these codes. The gains are very close to that of rate 1/2 codes with BPSK modulation. The reason for this similarity is because the two systems have the same encoder rates and QPSK signals have the same error performance as BPSK signals.

5.3.3 8-PSK Modulated Systems

Octal PSK modulation was also simulated to see the effect of increasing the signal constellation. This system was constructed using rate 2/3 codes with natural 8-PSK mapping. Ungerboeck's 8 and 16-state codes were used [10]. These code are fortunately optimized for both the Gaussian and the Rayleigh channels [46]. The minimum Hamming distances (in symbols) are 2 and 3, respectively. Moreover, the codes' minimum Euclidean distances are 4.586 and 5.172, respectively [46]. A decoding buffer length of 8ν was used for these codes. Higher MAP decoding gains than that of rate 2/3 codes with BPSK are obtained. For example, at a decoding BER of 0.02 gains (MAP vs. ML) as high as 3.3 dB for the Gaussian channel and 4.6 dB for the Rayleigh channel for a correlated source with about 50% source redundancy.

5.3.4 Observations and Discussions

To determine the performance of the different codes, extensive simulations were performed. Tables 5.2-5.11 show the MAP decoding gains (vs. ML decoding) for all different systems at BER of 10, 5, 2 and 1%. We first notice that the gains diminish as the BER decreases. The diminishing rate is slower for uncoded sys-

tems. At lower error rates ML decoding will have only a very slight degradation in error performance compared to sequence MAP decoding. This suggests that to reduce the decoding computations MAP decoding should be used only in bad channel environments.

We notice also that the gains for the Rayleigh channel are significantly higher than the corresponding gains for the AWGN channel. The reason for that is because the slope of the decoding BER curves for the Rayleigh channel is more gradual. This is very clear for uncoded systems where coding substantially improves the performance over Rayleigh channels.

Comparisons for the uncoded systems reveal that the type of redundancy significantly affects the decoding gains, especially for the Rayleigh channel. Redundancy in the form of memory results in larger MAP decoding gains. Another remark is that the gains does not diminish quickly as the BER decreases. This suggests that even at low error rates ($\sim 10^{-3}$) appreciable gains over the Rayleigh channels may be obtained.

For the coded systems, it can be seen that the rate 1/2 codes with BPSK and QPSK modulations give comparable gains. This is due to the fact that the same encoder is used and Gray mapped QPSK has the same error performance as BPSK. It is clear also that the gains monotonically increase with increasing the total redundancy; i.e., the effect of redundancy type is not apparent as in uncoded systems. This is attributed to the fact that convolutional codes are constructed using finite state machines which introduce memory to the system. Even for sources with no memory, the coded bits will have some memory correlation.

Regarding the number of encoder memory elements, it can be seen that

the major portion of decoding gains can be obtained using 4-state codes with low complexity; using more complex codes does not significantly improve the performance at relatively high error rates. This is because codes with higher number of states suffer from error propagation at very high channel error rates.

The effect of increasing the encoder rate is clear from the simulation results for rate 2/3 codes with BPSK modulation. More decoding gains are obtained. For example, for source V a decoding gain of 2.2(resp. 3.6) dB at a BER of 2% is obtained over the AWGN(resp. Rayleigh) channel using a 4-state rate 2/3 code, while the 4-state rate 1/2 codes achieves a decoding gains of only 1.1(resp. 1.6) dB at the same BER and source model over the AWGN(resp. Rayleigh) channel. Increasing the encoder rate further will also yield higher decoding gains. This may be noticed from the results of the rate 3/4 code with BPSK modulation reported in [78]. Another (perhaps surprising) observation is that the performance of the 8-state code is slightly worse than the performance of the 4-state code at high error rates. This is because three memory elements are used and one of the bits has only one delay element in the encoder structure. As a result, the 4-state code will outperform the 8-state code at high error rates. This fact has an effect on the decoding gains for the two codes and in some instances the decoding gains of the 4-state code are slightly higher than the gains for the 8-state code.

Finally, the effect of increasing the signal constellation is shown in the results for the 8-PSK coded systems. For example, at an encoder rate of 2/3, the decoding gains for the 8-PSK systems are higher than the gains for the same encoder rate systems with BPSK modulation. This is because the signal constellation has more points and hence it is more sensitive to large noise values. This fact

of increased MAP decoding gains is desirable since as the demand increases for systems with higher spectral efficiencies, and hence higher signal constellations, more MAP decoding gains could be obtained. Decoding gains as much as 3.8 dB at a BER of only 1% are obtained (see Table 5.11).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	0.9	0.5	0.4	0.3
BPSK	1/2	1/2	2	1.2	0.9	0.6	0.5
BPSK	1/2	1/2	3	1.2	0.9	0.7	0.5
BPSK	1/2	1/2	4	1.4	1.0	0.8	0.5
BPSK	2/3	2/3	2	2.4	1.7	1.1	.8
BPSK	2/3	2/3	3	2.5	1.7	1.1	1.0
BPSK	2/3	2/3	4	2.4	1.7	1.3	1.0
QPSK	1/2	1	2	1.2	0.8	0.6	0.5
QPSK	1/2	1	3	1.2	0.8	0.7	0.5
QPSK	1/2	1	4	1.3	1.0	0.8	0.5
8PSK	2/3	2	3	3.2	2.4	1.7	1.4
8PSK	2/3	2	4	3.2	2.4	1.9	1.5

Table 5.2: MAP decoding gains for different trellis codes over AWGN channels. Source I (with distribution $\Pr(0|0) = .8, \Pr(1|1) = .2$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	2.1	1.6	1.3	1.2
BPSK	1/2	1/2	2	1.6	1.1	0.8	0.7
BPSK	1/2	1/2	3	1.7	1.2	0.9	0.8
BPSK	1/2	1/2	4	1.8	1.4	1.1	0.9
BPSK	2/3	2/3	2	3.5	2.6	1.9	1.6
BPSK	2/3	2/3	3	3.6	2.6	1.9	1.6
BPSK	2/3	2/3	4	3.6	2.7	2.2	1.9
QPSK	1/2	1	2	1.5	1.1	0.8	0.6
QPSK	1/2	1	3	1.5	1.1	0.9	0.7
QPSK	1/2	1	4	1.8	1.4	1.1	1.0
8PSK	2/3	2	3	3.8	2.8	2.1	1.9
8PSK	2/3	2	4	3.9	3.0	2.5	2.1

Table 5.3: MAP decoding gains for different trellis codes over Rayleigh channels. Source I (with distribution $\Pr(0|0) = .8, \Pr(1|1) = .2$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	1.3	0.9	0.6	0.5
BPSK	1/2	1/2	2	1.3	0.9	0.7	0.5
BPSK	1/2	1/2	3	1.1	0.9	0.7	0.5
BPSK	1/2	1/2	4	1.2	1.0	0.8	0.5
BPSK	2/3	2/3	2	2.0	1.5	1.1	.9
BPSK	2/3	2/3	3	1.6	1.4	1.1	.9
BPSK	2/3	2/3	4	2.1	1.8	1.3	1.1
QPSK	1/2	1	2	1.4	0.9	0.7	0.6
QPSK	1/2	1	3	1.1	0.8	0.7	0.5
QPSK	1/2	1	4	1.3	1.0	0.8	0.6
8PSK	2/3	2	3	1.9	1.5	1.1	1.0
8PSK	2/3	2	4	2.8	2.2	1.8	1.5

Table 5.4: MAP decoding gains for different trellis codes over AWGN channels. Source II (with distribution $\Pr(0|0) = .8, \Pr(1|1) = .8$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	2.6	2.5	2.3	2.2
BPSK	1/2	1/2	2	1.6	1.2	0.9	0.8
BPSK	1/2	1/2	3	1.5	1.1	0.9	0.8
BPSK	1/2	1/2	4	1.7	1.4	1.1	1.0
BPSK	2/3	2/3	2	3.0	2.4	2.1	1.8
BPSK	2/3	2/3	3	2.4	2.1	1.8	1.5
BPSK	2/3	2/3	4	3.2	2.7	2.3	1.9
QPSK	1/2	1	2	1.6	1.2	1.0	0.9
QPSK	1/2	1	3	1.5	1.1	0.9	0.8
QPSK	1/2	1	4	1.6	1.3	1.1	1.0
8PSK	2/3	2	3	2.3	2.1	1.8	1.7
8PSK	2/3	2	4	3.5	2.8	2.4	2.1

Table 5.5: MAP decoding gains for different trellis codes over Rayleigh channels. Source II (with distribution $\Pr(0|0) = .8, \Pr(1|1) = .8$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	1.7	1.1	0.7	0.6
BPSK	1/2	1/2	2	1.6	1.1	0.8	0.7
BPSK	1/2	1/2	3	1.5	1.1	0.8	0.7
BPSK	1/2	1/2	4	1.7	1.2	0.9	0.8
BPSK	2/3	2/3	2	2.5	1.8	1.3	1.1
BPSK	2/3	2/3	3	2.3	1.7	1.3	1.1
BPSK	2/3	2/3	4	2.7	2.1	1.8	1.3
QPSK	1/2	1	2	1.5	1.1	0.8	0.7
QPSK	1/2	1	3	1.5	1.1	0.8	0.6
QPSK	1/2	1	4	1.6	1.2	0.9	0.8
8PSK	2/3	2	3	2.8	2.1	1.6	1.3
8PSK	2/3	2	4	3.5	2.6	2.0	1.6

Table 5.6: MAP decoding gains for different trellis codes over AWGN channels. Source III (with distribution $\Pr(0|0) = .88, \Pr(1|1) = .65$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	3.3	2.9	2.7	2.6
BPSK	1/2	1/2	2	2.1	1.6	1.1	1.0
BPSK	1/2	1/2	3	2.0	1.5	1.1	0.9
BPSK	1/2	1/2	4	2.2	1.7	1.3	1.2
BPSK	2/3	2/3	2	3.8	2.9	2.3	2.1
BPSK	2/3	2/3	3	3.5	2.6	2.1	1.9
BPSK	2/3	2/3	4	3.6	2.7	2.2	1.9
QPSK	1/2	1	2	2.0	1.5	1.2	1.0
QPSK	1/2	1	3	2.0	1.5	1.2	1.0
QPSK	1/2	1	4	2.2	1.6	1.3	1.2
8PSK	2/3	2	3	3.5	2.8	2.4	2.1
8PSK	2/3	2	4	4.4	3.4	2.9	2.6

Table 5.7: MAP decoding gains for different trellis codes over Rayleigh channels. Source III (with distribution $\Pr(0|0) = .88, \Pr(1|1) = .65$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	2.3	1.5	1.0	0.8
BPSK	1/2	1/2	2	1.9	1.4	1.0	0.9
BPSK	1/2	1/2	3	1.9	1.3	1.0	0.8
BPSK	1/2	1/2	4	2.1	1.5	1.2	1.0
BPSK	2/3	2/3	2	3.5	2.6	2.0	1.6
BPSK	2/3	2/3	3	2.9	2.2	1.8	1.6
BPSK	2/3	2/3	4	3.5	2.8	2.1	1.8
QPSK	1/2	1	2	1.9	1.4	1.0	0.8
QPSK	1/2	1	3	1.8	1.3	1.0	0.8
QPSK	1/2	1	4	2.1	1.5	1.2	0.9
8PSK	2/3	2	3	3.3	2.6	2.0	1.7
8PSK	2/3	2	4	4.6	3.5	2.7	2.3

Table 5.8: MAP decoding gains for different trellis codes over AWGN channels. Source IV (with distribution $\Pr(0|0) = .9, \Pr(1|1) = .8$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	4.2	3.8	3.6	3.5
BPSK	1/2	1/2	2	2.6	1.9	1.5	1.3
BPSK	1/2	1/2	3	2.5	1.9	1.4	1.3
BPSK	1/2	1/2	4	2.8	2.2	1.8	1.5
BPSK	2/3	2/3	2	4.9	4.1	3.4	3.0
BPSK	2/3	2/3	3	4.2	3.4	2.8	2.4
BPSK	2/3	2/3	4	4.9	4.1	3.5	3.0
QPSK	1/2	1	2	2.5	1.9	1.6	1.4
QPSK	1/2	1	3	2.4	1.8	1.4	1.2
QPSK	1/2	1	4	2.7	2.1	1.6	1.5
8PSK	2/3	2	3	4.1	3.5	3.1	2.9
8PSK	2/3	2	4	5.6	4.5	3.8	3.4

Table 5.9: MAP decoding gains for different trellis codes over Rayleigh channels. Source IV (with distribution $\Pr(0|0) = .9, \Pr(1|1) = .8$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	2.2	1.1	0.7	0.6
BPSK	1/2	1/2	2	2.6	1.6	1.1	0.9
BPSK	1/2	1/2	3	2.5	1.6	1.2	1.0
BPSK	1/2	1/2	4	2.9	1.9	1.4	1.2
BPSK	2/3	2/3	2	6.7	3.4	2.2	1.7
BPSK	2/3	2/3	3	6.5	3.4	2.3	1.8
BPSK	2/3	2/3	4	6.3	3.7	2.5	2.0
QPSK	1/2	1	2	2.9	1.6	1.1	0.9
QPSK	1/2	1	3	2.5	1.6	1.1	0.9
QPSK	1/2	1	4	2.9	1.9	1.4	1.1
8PSK	2/3	2	3	7.5	4.4	3.0	2.4
8PSK	2/3	2	4	7.7	4.7	3.3	2.7

Table 5.10: MAP decoding gains for different trellis codes over AWGN channels. Source V (with distribution $\Pr(0|0) = .2, \Pr(1|1) = .9$).

Modulation Type	Rate	Spectral Efficiency	ν	Gains (BER)			
				10%	5%	2%	1%
BPSK	uncoded	1	-	5.0	3.2	2.7	2.5
BPSK	1/2	1/2	2	3.6	2.3	1.6	1.3
BPSK	1/2	1/2	3	3.5	2.4	1.7	1.5
BPSK	1/2	1/2	4	3.9	2.7	2.0	1.7
BPSK	2/3	2/3	2	8.9	5.0	3.6	3.0
BPSK	2/3	2/3	3	8.7	5.0	3.6	2.9
BPSK	2/3	2/3	4	8.5	5.4	4.0	3.3
QPSK	1/2	1	2	3.7	2.2	1.5	1.3
QPSK	1/2	1	3	3.5	2.2	1.6	1.3
QPSK	1/2	1	4	3.9	2.6	2.0	1.7
8PSK	2/3	2	3	9.4	5.4	4.0	3.5
8PSK	2/3	2	4	9.4	5.9	4.6	3.8

Table 5.11: MAP decoding gains for different trellis codes over Rayleigh channels. Source V (with distribution $\Pr(0|0) = .2, \Pr(1|1) = .9$).

5.3.5 The effect of Increasing the Signal Dimensionality on MAP Gains

The previous results suggest that as the encoder rate is increased, the benefit of MAP decoding (relative to ML) increases. In trellis codes with multidimensional (greater than two) constellations (also referred to as multiple TCM) the information bits are grouped and sent over m bauds, where m here stands for the multiplicity of the TCM scheme. For example, in the previously mentioned two-dimensional Gray mapped QPSK (2-D QPSK) one information bit is delivered to a rate 1/2 encoder at every baud. The encoder 2-tuple output bits are mapped to one of the QPSK signals. The bandwidth efficiency is 1 bit/s/Hz. Suppose we have a 4-D QPSK with the same bandwidth efficiency and ML coding gains. What is the performance of the two codes with sequence MAP decoding? The previous results suggest that the 4-D QPSK coded system will perform better.

A comparative example is used here to show the effect of increasing the signal dimensionality on the sequence MAP decoding gains. The previously used 8-state Gray mapped QPSK code has a minimum time diversity of 4. Its normalized minimum Euclidean and product distances are 12 and 64, respectively. Comparing this code with the codes in [86] and [57], this code is optimized for both AWGN and Rayleigh channels. An 8-state 4-dimensional code is constructed. Its trellis diagram is shown in Figure 5.3. Two information bits are delivered to a rate 2/3 encoder every two bauds. The encoder output 3-tuple bits are used to select one of the 8 2-tuple QPSK signal points. The bandwidth efficiency of this code is also 1 bit/s/Hz. The code's minimum time diversity is 4. Its normalized minimum Euclidean and product distances are 12 and 64, respectively - i.e., its asymptotic coding gain with ML decoding is the same as that of

the 2-D QPSK scheme. Figure 5.4 show the performance of the two codes with ML decoding and source model V over the Rayleigh channel. If sequence MAP decoding is used, the performance of the two codes is different, especially at low SNR. Figure 5.5 show the performance of the two codes with MAP decoding. It is clear that the 4-D QPSK scheme outperforms the corresponding 2-D QPSK scheme. Increasing the signal dimensionality is also expected to yield more gains. However, since this will also increase the number of signals per trellis branch, more delay and interleaving are required which may not be suitable for real time applications.

5.4 CELP LSP's Coding

Codebook-excited linear predictive (CELP) coding is a frame-oriented technique that breaks a speech signal into blocks of samples that are processed as one unit. The particular implementation we consider is Federal Standard 1016 (FS 1016) 4.8 kbit/s CELP [87]. The CELP parameters that are transmitted over the noisy channel include the stochastic code book index and gain, the adaptive code book index (pitch delay) and gain, and 10 ordered line spectral parameters (LSP's). In the FS 1016 CELP, each LSP is quantized by either a three-bit or a four-bit scalar quantizer. The second through fifth LSP's are quantized by four-bit quantizers; the rest are quantized to three bits. The quantized LSP's refer to frequencies that must be ordered (LSP-1 $<$ LSP-2 $<$ $\cdots$ $<$ LSP-10). In this work, we consider only the three most significant bits of each LSP, ignoring the least significant bit in the second through fifth parameters.

The CELP encoder leaves some redundancy in the encoded bit stream in the

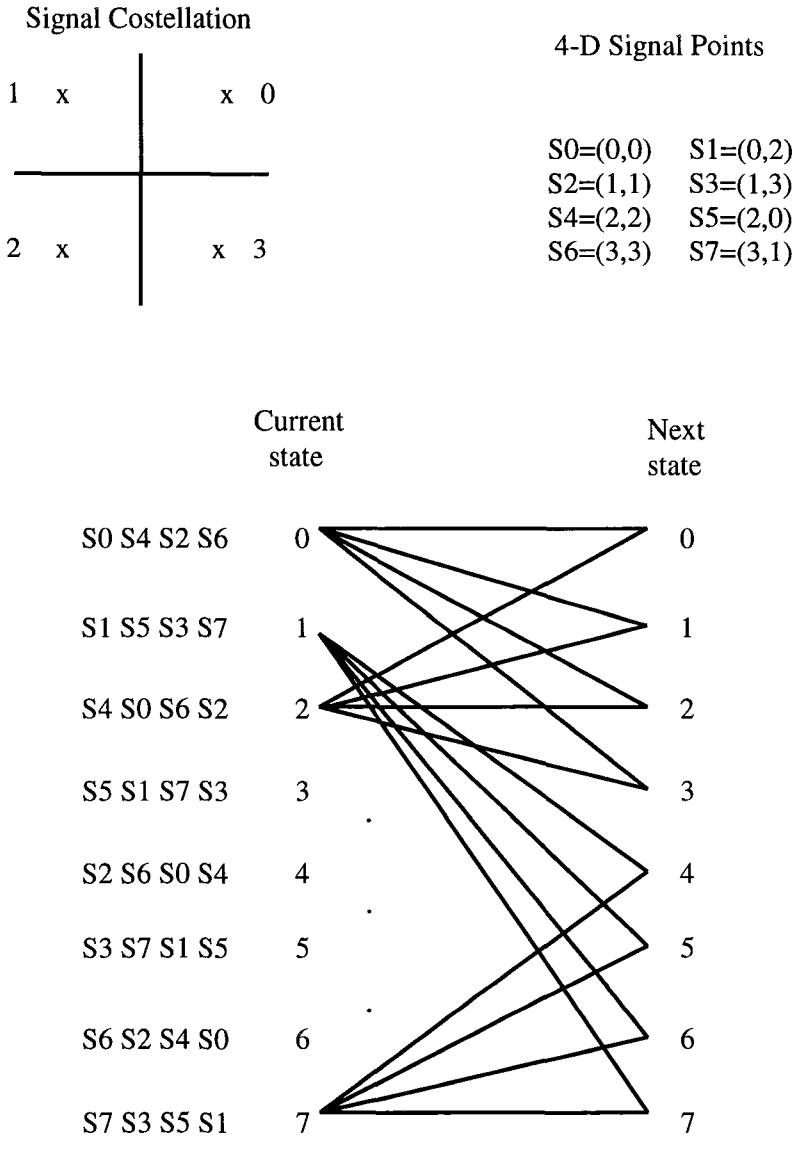


Figure 5.3: 4-D signal points and the trellis diagram of the 4-D QPSK 8-state encoder.

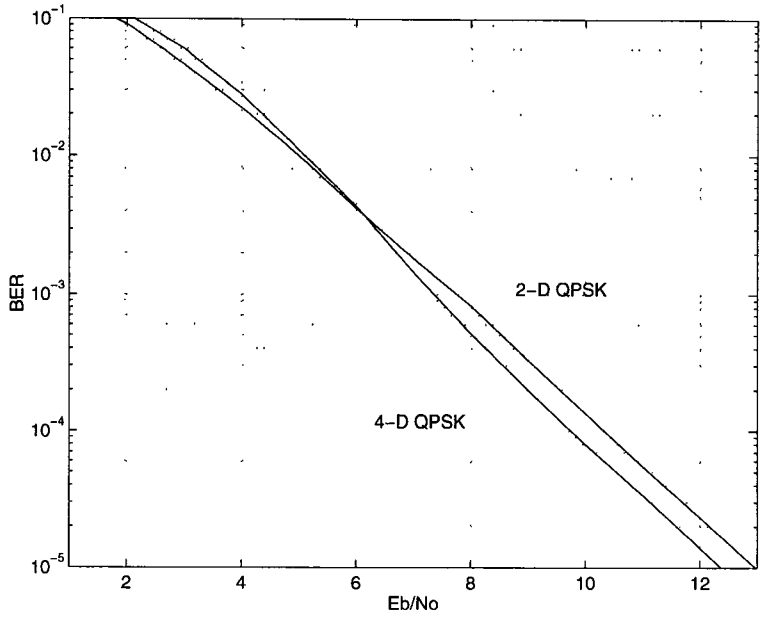


Figure 5.4: BER of the 2-D and 4-D QPSK 8-state schemes with ML decoding.

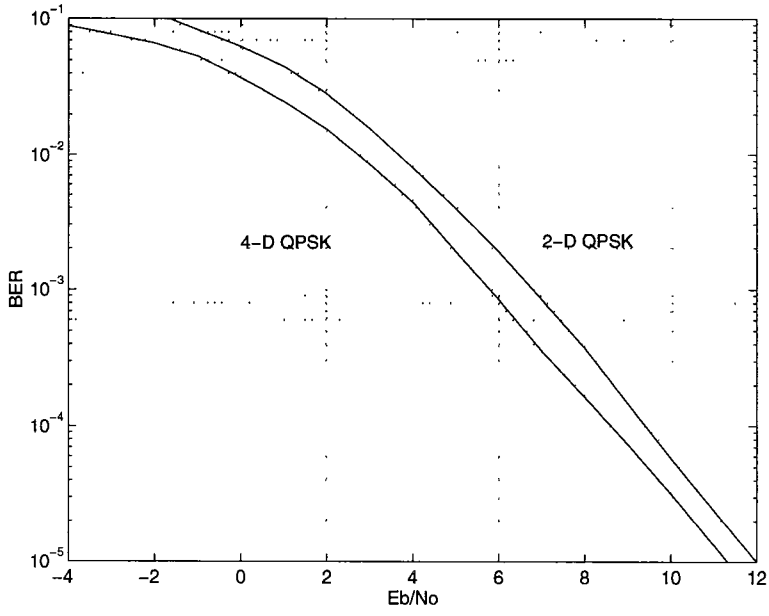


Figure 5.5: BER of the 2-D and 4-D QPSK 8-state schemes with MAP decoding.

form of memory and non-uniformity. The modeling of CELP encoded speech is described in detail in [78]. This modeling is briefly described below. A large training sequence consisting of 83,826 frames (about 42 minutes of speech) from the TIMIT speech database [82] was used to estimate the prior LSP distributions needed for the MAP decoding. For every 30 msec of speech an LPC analysis was performed according to FS 1016 standards to arrive at the 10 quantized LSP's. The relative frequency of transitions between the values of the three high-order bits of each LSP were compiled to extract Markov transition probabilities.

Suppose we encode a segment of speech using FS 1016 CELP, resulting in a sequence of CELP frames. The quantized LSP frames are *approximated* by a stationary Markov process [78]. Denote the process entropy rate (in bits/frame) by H_∞ which represents the minimum number of bits per frame required to describe three bits. The CELP encoder produces 30 bits/frame to describe the LSP's, so the *residual redundancy* – i.e., the total redundancy (per frame) in the CELP-encoded LSP's – is

$$\rho_T = 30 - H_\infty \text{ (bits/frame)}. \tag{5.7}$$

H_∞ (and so ρ_T) were estimated by observing a long training sequence and matching the observations to a particular model of a random process; the entropy rate of the model process was then computed and used as our estimate of H_∞.

Two models for the generation of LSP's distribution are proposed.

- Model 1 – which does not attempt to take into account *any* correlation between frames – indicates that $\rho_T = 9.867$ of the 30 high-order bits in the LSP's are redundant. Approximately $\rho_D = 5.275$ bits of redundancy were due to the non-uniform distribution of the LSP's, and approximately $\rho_M = 4.593$ bits of redundancy were due to the memory within a frame.

- Model 2 – which *does* take into account both inter-frame and intra-frame correlation – indicates that $\rho_T = 12.485$ of the 30 high-order bits in the LSP's are redundant. Once again, $\rho_D = 5.275$ bits of redundancy were due to the non-uniform distribution of the LSP's, while $\rho_M = 7.211$ bits of redundancy were due to the memory remaining both within a frame and between frames.

We employ three soft-decision decoding schemes based on the Viterbi decoding algorithm:

- ML – a maximum likelihood Viterbi decoding algorithm which *does not* utilize the LSP's a priori information.
- MAP 1 – a maximum a-posteriori (MAP) decoding algorithm that exploits only the redundancy due to the non-uniform distribution of the LSP's and their correlation within a frame – approximately 10 bits/frame.
- MAP 2 – which exploits the redundancy from the non-uniform distribution of the LSP's *and* their inter-frame and intra-frame correlation – approximately 12.5 bits/frame.

A decoding buffer length of 10 symbols is used to limit the decoding delay. All algorithms are implemented so as to yield a decoding delay of only one frame.

5.4.1 Coding of the LSP's

Since the quantized LSP's are represented by 3 bits, every binary 3-tuple representing one LSP must be encoded every encoder time unit. In [78], a rate 3/4 binary convolutional code with BPSK modulation was used. The resulting

system spectral efficiency was 0.75 bit/s/Hz. To increase the spectral efficiency we propose using QPSK modulation. The proposed codes are 4-state and 8-state rate 3/4 codes with 4-D QPSK. Their spectral efficiency is 1.5 bit/s/Hz. These codes were designed for Rayleigh fading channels [56]. Moreover, their minimum Euclidean distances are the same as the corresponding codes designed for the AWGN channel [86]. This means that both codes are optimum for both channels. The minimum Euclidean distances of the 4-state and 8-state codes are 6 and 8, respectively. The minimum time diversity of both codes are 2. However, the minimum product distance of the 8-state code is 16, which is twice that of the 4-state code. It should be pointed out that the number of branches in the minimum-length error event path for the 4-state code is 2 while that of the 8-state code is one (i.e., corresponds to parallel branches). The effect of this will be clear from the simulations.

5.4.2 Simulation Results

Simulations were used to determine the performance of the proposed decoding algorithms. The three high-order bits of each of the ten quantized LSP's were channel encoded using of the two codes described above. The outputs of the channel encoders were then mapped to a pair of QPSK signals and transmitted over either the AWGN channel or the fully interleaved Rayleigh channel. After appropriate demodulation, the signals were decoded with the proposed channel decoders and the decoded LSP's were fed into the CELP decoder for speech reconstruction. The decoder buffer length is set to 10 symbols [56]. This allows interleaving to be done within only one frame, and so only a delay of one frame exists. Note that sequence MAP decoding does not introduce any additional

delay.

The testing sequence consisted of 4753 frames (about 2.2 minutes of speech) – 48 sentences, half uttered by female speakers and half by male speakers from different dialect regions. No speaker appeared in both the training and testing sequences. Thus the approach used in this simulation was to use a single "universal" model – constructed from a very large training sequence – to decode *all* the speech samples.

In evaluating the performance of the various decoders we use two criteria. The first is the average spectral distortion (SD), the most commonly used distortion measure for the LSP's [88]. More specifically,

$$SD = \frac{1}{T} \sum_{j=1}^{T} \left[\int_{-\pi}^{\pi} \left(10 \log_{10} S_j(w) - 10 \log_{10} \hat{S}_j(w) \right)^2 \frac{dw}{2\pi} \right]^{\frac{1}{2}} \quad \text{dB}, \qquad (5.8)$$

where $S_j(w)$ and $\hat{S}_j(w)$ are the original and reconstructed spectra associated with frame j, and T is the total number of frames. Roughly speaking, an average spectral distortion of 1 dB or less is equivalent to perceptually transparent encoding of the LSP coefficients [88]. In addition to average spectral distortion, the percentage of outliers – i.e., the fraction of frames with distortion greater than 4 dB – were also compiled during the simulation. It should be noted that the spectral distortion introduced by CELP's scalar quantizer alone (when the channel is noiseless) is around 1.50 dB with 0.08 % of outliers $>$ 4 dB.

The second measure of the decoders' performance is symbol error rate (P_s) – i.e., the fraction of LSP's the decoder decoded incorrectly. Tables 5.12 and 5.13 indicate the MAP decoding gains for the two codes over both channels. The gains are shown for a P_s values of 1, 5, 10 and 15%, since error concealment and interpolation are made at P_s's as high as 15%.

It is clear that the gains over the Rayleigh channel are higher than the gains over the AWGN channel. Moreover, significant gains are achieved. For example, at an average spectral distortion of 2 dB, total MAP decoding gains of 1.5 and 3.1 dB were achieved over the AWGN and Rayleigh channels, respectively. Spectral distortion gains are shown in Figures 5.6 and 5.7. Also, Tables 5.12 and 5.13 summarize the decoding gains for various SD and P_s values. It is noticed that the large portion of the gains can be achieved using Model 1. This agrees with the calculation of the redundancies for both models, where the second model gives only 2.5 bits additional redundancy. A comparison between the two codes shows that even though the 8-state code outperforms the 4-state code at high SNR's, the MAP decoding gains for the 4-state code are slightly higher. The reason for that is attributed to the codes' minimum error event path length, which affects the performance at low SNR.

Decoding Gains	Channel Type	SD(dB)				P_s(%)			
		2.0	2.5	3.0	3.5	1%	5%	10%	15%
MAP 1 vs. ML	AWGN	1.2	1.5	1.8	2.1	0.8	1.1	1.3	1.6
	Rayleigh	2.4	2.7	2.9	3.2	2.0	2.1	2.2	2.3
MAP 2 vs. MAP 1	AWGN	0.3	0.4	0.4	0.4	0.2	0.3	0.3	0.3
	Rayleigh	0.7	0.8	0.8	0.8	0.3	0.6	0.6	0.6
MAP 2 vs. ML	AWGN	1.5	1.9	2.2	2.5	1.0	1.4	1.6	1.9
	Rayleigh	3.1	3.5	3.7	4.0	2.3	2.6	2.8	2.9

Table 5.12: Sequence MAP decoding gains for the the CELP encoded speech with 4-state 4-D QPSK TCM schemes over both AWGN and Rayleigh channels

Decoding Gains	Channel Type	SD(dB)				P_s(%)			
		2.0	2.5	3.0	3.5	1%	5%	10%	15%
MAP 1 vs. ML	AWGN	1.1	1.3	1.6	1.8	0.7	1.0	1.2	1.3
	Rayleigh	2.3	2.4	2.6	2.9	1.9	1.9	2.0	2.1
MAP 2 vs. MAP 1	AWGN	0.4	0.5	0.5	0.6	0.2	0.3	0.4	0.4
	Rayleigh	0.8	0.8	0.8	0.9	0.5	0.6	0.6	0.6
MAP 2 vs. ML	AWGN	1.5	1.8	2.1	2.4	0.9	1.3	1.6	1.7
	Rayleigh	3.1	3.2	3.4	3.8	2.4	2.5	2.6	2.7

Table 5.13: Sequence MAP decoding gains for the the CELP encoded speech with 8-state 4-D QPSK TCM schemes over both AWGN and Rayleigh channels

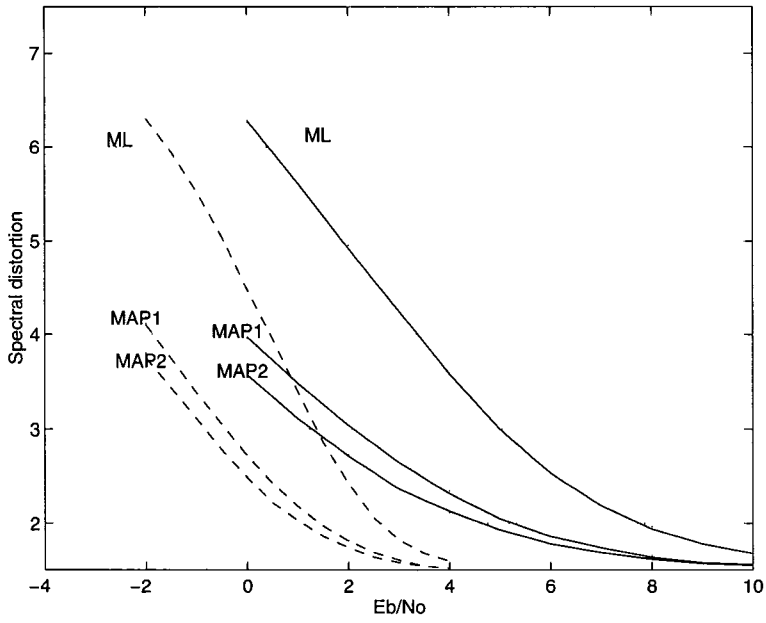


Figure 5.6: Spectral distortion vs. Eb/No for the 4-state 4-D QPSK scheme - solid(Rayleigh), dashed(AWGN).

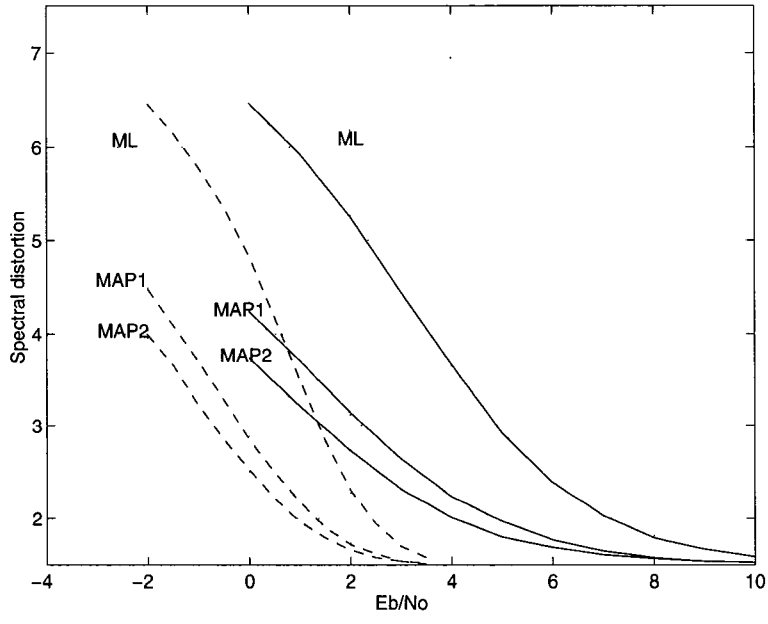


Figure 5.7: Spectral distortion vs. Eb/No for the 8-state 4-D QPSK scheme - solid(Rayleigh), dashed(AWGN).

5.4.3 Coding of CELP LSP's via I-Q QPSK

In this section, another coding scheme is proposed. In Chapter 3, it was shown that inphase-quadrature (I-Q) trellis codes offer the advantage of higher minimum time diversities. Also, the increased encoder rate gives higher decoding gains. These two features are used together to provide a robust system with a bandwidth efficiency of 1.5 bit/s/Hz. The proposed scheme is outlined as follows. *Two frames* are encoded together using two rate 3/4 encoders. Each encoder is mapped to a binary antipodal signal. The first frame is encoded using encoder-I and the second frame is encoded using encoder-Q. The transmitted signal is QPSK where its in-phase (quadrature) are specified by encoder-I (encoder-Q). Only Model 1 for the a priori LSP's information is used (i.e., correlation between frames is not exploited) and a delay of two frames is imposed. The same 32-state rate 3/4 code proposed in [78] is used. Using the approach, larger MAP decoding gains are obtained. Table 5.14 shows the decoding gains obtained using this scheme. Also, a comparison between this code and the previous ones in Figure 5.8 shows that the I-Q code is achieves a very robust performance at very noisy channel conditions.

Decoding	Channel	SD(dB)				P_s(%)			
Gains	Type	2.0	2.5	3.0	3.5	1%	5%	10%	15%
MAP 1	AWGN	2.1	2.4	2.5	1.8	0.7	1.0	1.2	1.3
vs. ML	Rayleigh	3.7	4.1	4.2	4.5	3.7	4.1	4.4	4.6

Table 5.14: Sequence MAP 1 decoding gains for the the CELP encoded speech with 32-state I-Q QPSK TCM scheme over both AWGN and Rayleigh channels

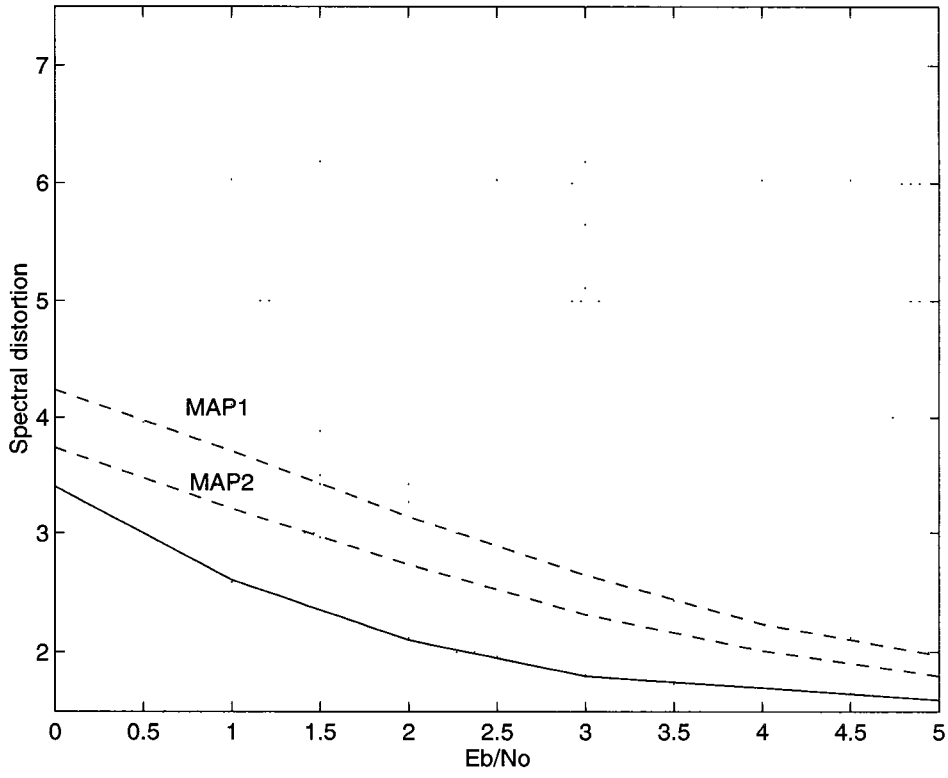


Figure 5.8: Spectral distortion vs. Eb/No for the 8-state 4-D QPSK scheme (dashed) - and the I-Q QPSK 32-state code (solid) over the Rayleigh channel.

5.5 Summary

In this chapter we considered sequence maximum a posteriori (MAP) decoding of correlated signals transmitted over very noisy AWGN and Rayleigh channels. In the first part of this work, a first order two-state Markov model is used for the source. A variety of different systems with different sources, BPSK, QPSK and 8-PSK modulation schemes and different encoder complexities were simulated. Sequence MAP decoding (compared to ML decoding) proves to substantially improve the performance at very noisy channel conditions, especially for systems with moderate redundancy. The MAP decoding gains when the channel is Rayleigh distributed are higher than the gains when the channel is AWGN. Most of the MAP decoding (relative to ML decoding) gains are achieved with low complexity encoders. Moreover, trellis coded systems with higher encoder rates have significantly more MAP decoding gains. Also, more decoding gains are obtained for encoders with larger signal constellations. A comparative example of a 4-D versus a 2-D constellations shows that the multidimensional constellation achieves more MAP gains than its corresponding trellis coded scheme with 2-dimensional (2-D) constellation.

In the second part, a practical example for coding the CELP line spectral parameters (LSP's) using trellis codes with 4-D QPSK modulation is presented. Two source models are used. One is based on the intra-frame correlation while the second one models both intra-frame and inter-frame correlations. Coding gains as much as 4 dB are achieved. Also, a comparison between the conventionally designed codes and an I-Q QPSK scheme shows that the I-Q scheme achieves better performance although only the correlation within a frame is exploited.

Chapter 6

Summary and Conclusions

6.1 Presentation Summary

In Chapter 2 the modeling of fading channels is discussed. Different statistical channel models and the techniques to generate them were described. In Chapter 3 we presented some TCM schemes that provide high coding gain over the frequency non-selective slowly Rayleigh distributed fading channel. It is shown that the use of I-Q TCM results in greater minimum time diversity than the conventional design in which a single encoder is used. Using this approach codes with bandwidth efficiencies of 1,2, and 3 bits/sec/Hz were described for various constraint lengths. The performance of these codes were examined analytically and via simulation; the results show a large improvement in the BER when compared with conventional TCM schemes.

In Chapter 4 the use of diversity with coding was investigated. Three different diversity combining schemes were compared. Expressions for the cutoff rate parameter were shown for the three different diversity combining schemes. Also,

tight upper bounds on the pairwise error probability were derived for the three combining schemes. The upper bounds were used to derive the bit error probabilities for uncoded as well as trellis coded systems. They were expressed in a product form to allow the use of the transfer function approach, which accounts for all trellis paths, for evaluating the performance of trellis coded systems.

In Chapter 5 sequence maximum a posteriori (MAP) decoding of correlated signals transmitted over very noisy AWGN and Rayleigh channels was presented. A variety of different systems with different sources, BPSK, QPSK and 8-PSK modulation schemes and different encoder complexities were simulated. Sequence MAP decoding (relative to ML decoding) proves to substantially improve the performance at very noisy channel conditions especially for systems with moderate redundancy. Moreover, it is shown that trellis coded systems with higher encoder rates have significantly more MAP decoding gains relative to ML decoding

6.2 Conclusions

The following is a synopsis of the major results and observations

1. In communicating over flat Rayleigh fading the code's minimum time diversity is the primary performance criterion. The use of I-Q TCM results in greater minimum time diversity of the code. Therefore, substantial improvements, with respect to conventionally designed codes, are obtained.

2. I-Q TCM schemes are not restricted to a specific spectral efficiency. Codes with bandwidth efficiencies of 1, 2 and 3 bits/s/Hz were presented and their performance was evaluated. Coding gains of 2-5 dB at BER $= 10^{-5}$,

with respect to conventional coding schemes, are obtained.

3. The performance of some of these codes over the correlated fading model is shown. Finite interleaving results in relatively little degradation.

4. In designing the interleaver, it is found that a span size of four times the number of memory elements results in acceptable performance.

5. Although the I-Q codes are designed for the Rayleigh channel, some improvements are still obtained for the $K = 7$ and $K = 9$ Rician channel.

6. The performance of trellis coded systems with different diversity combining is tightly upper bounded using the transfer function approach.

7. When the diversity signals have equal energies, maximal ratio combining is the best combining scheme.

8. For low to moderate diversity orders, equal gain combining performance is within 1 dB from that of maximal ratio combining.

9. The highest relative coding gains are achieved in going from single to double diversity.

10. With increasing diversity, incremental coding gains diminish and the difference between the three combining schemes increases.

11. Correlated diversity branches does not degrade the performance substantially for $\rho \leq 0.5$.

12. Sequence MAP decoding gains when the channel is Rayleigh distributed are higher than the gains when the channel is AWGN.

13. The major portion of the MAP decoding gains is achieved with a low complexity encoder. The gains diminish with increasing the encoder constraint length.

14. Encoders with higher rates have significantly more MAP decoding gains, relative to ML decoding.

15. More MAP decoding gains (relative to ML decoding) are obtained for encoders with larger signal constellations.

6.3 Future Work

Future work can be summarized as follows

1. As mentioned by Zehavi in [55], bit interleaving can add more inherent minimum time diversity to the system at the expense of reduced minimum Euclidean distance. This approach could in principle be used along with the I-Q approach to further improve the error performance over Rayleigh fading channels.

2. I-Q QPSK schemes with differentially coherent demodulation should be investigated.

3. A general formula for the cutoff rate with diversity reception that includes the effect of space and time correlations should be derived.

4. The performance of I-Q schemes should be investigated over the frequency selective fading channel model.

5. The performance of the proposed codes with the existence of co-channel interference can be evaluated.

6. The tight upper bounds for the Rayleigh channel should be extended to include the Rician channel.

7. A more challenging problem is to find tight upper bounds for trellis coded systems with ML and MAP decoding at high BER values .

8. The performance of RAKE receivers can be similarly analyzed as is done for the performance of diversity systems.

Bibliography

[1] Proakis, J., Digital Communications,Mc-Graw Hill Book Company,1989.

[2] Bateman, A., "Feedforward Transparent Tone-in-band L its Implemenations and Applications," *IEEE Transactions on Vehicular Technology*, August 1990, 39(3), pp.235-243.

[3] M. L. Moher and J. H. Lodge "TCMP-A Modulation and Coding Strategy for Rician Fading Channels", *IEEE Journal on Selected Areas in Communications*, December 1989, 7(9), pp. 1347-1355.

[4] Cavers, J., "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels," *IEEE Transactions on Vehicular Technology*, November 1991, 40(40, pp. 686-693.

[5] Sampei, S. and T. Sunaga, "Rayleigh Fading Compensation for QAM in Land Mobile Radio Communications," *IEEE Transactions on Vehicular Technology*, May 1993, 42(2), pp.137-147

[6] , Y. Liu and S. Blostein, "Identification of Frequency Nonselective Fading Channels using Decision Feedback and Adaptive Linear Prediction," *IEEE Transactions on Communications*, April 1995

[7] Stein,S. "Fading Channel Issues in System Engineering," *IEEE Journal on Selected Areas in Communications*, February 1987, 5(2), pp. 68-89.

[8] Massey, J. L., "Coding and Modulation in Digital Communications," Proc. 1974 International Zurich Seminar on Digital Communications, Zurich, Switzerland, March 1974, pp. E2(1)-E2(4).

[9] Imai, H. and S. Hirakawa "A New Multilevel Coding Method Using Error-Correcting Codes," *IEEE Transactions on Information Theory*, May 1977, 23(3), pp. 371-377.

[10] Ungerboeck, G. "Channel Coding with Multilevel/Phase Signals," *IEEE Transactions on Information Theory*, January 1982, 28(1), pp.55-67.

[11] Hagenauer, J. and E. Lutz "Forward Error Correction Coding for fading Compensation in Mobile Satellite Channels," *IEEE Journal on Selected Areas in Communications*, February 1987, 5(2), pp. 215-225.

[12] Calderbank, R. and J. E. Mazo, "A New Description of Trellis Codes," *IEEE Transactions on Information Theory*, November 1984, 30(6), pp. 784-791.

[13] Zehavi, E. and J. K. Wolf, "On the Performance Evaluation of Trellis Codes," *IEEE Transactions on Information Theory*, March 1987, 33(3), pp. 196-201.

[14] Biglieri, E. and P. J. McLane, "Uniform Distance and Error Properties of TCM Schemes," *IEEE Transactions on Communications*, January 1991, 39(1), pp. 41-53.

[15] Wilson, S.G., Sleeper, H.A. and N.K. Srinath, "Four-Dimensional modulation and Coding," Proc. ICC'84, Amestrdam, The Netherlands, May 1984, pp.919-923.

[16] Calderbank, R, and N. Sloane, "Four-Dimensional Modulation with an Eight State Trellis Code," Bell System technical Journal, vol. 64, 1985, pp. 1005-1018.

[17] Calderbank, R. and N. Sloane, "An Eight-Dimensional Trellis Code," proceedings of IEEE, vol.74, 1986, pp. 757-759.

[18] Wei, L.-F. , "Trellis-Coded Modulation with Multidimensional Constellations," IEEE Transactions on Information Theory, July 1987, 33(4), pp. 483-501.

[19] Calderbank, A. R. and N. J. Sloane, "New Trellis Codes Based on Lattices and Cosets," IEEE Transactions on Information Theory, March 1987, 33(2), pp.177-195.

[20] Forney, G. D., "Coset Codes- Part I : Introduction and Geometrical Classification," IEEE Transactions on Information Theory, September 1988, 34(5), pp. 1123-1151.

[21] Benedetto, S. M. Mondin and G. Montrosi "Performance Evaluation of Trellis-Coded Modulation Schemes," *Proceedings of IEEE*, June 1994, 82(6), pp. 833-855.

[22] Tanner, R.M., "Algebraic construction of Large Euclidean Distance Combined Modulation Schemes," IEEE International Symposuim on Information Theory, 1986.

[23] Yamaguchi, K. and H. Imai "Highly Reliable Multilevel Channel Coding System Using Binary Convolutional Codes," *Electronics Letters*, 27th August 1987, 23(18), pp. 939-941.

[24] Pottie, G.J. and D.P. Taylor, "Multilevel Codes Based on Partitioning," *IEEE Transactions on Information Theory*, January 1989, 35(1),pp. 87-98.

[25] Calderbank, A. R., "Multilevel Codes and Multistage Decoding," *IEEE Transactions on Communications*, March 1989, 37(3), pp. 222-229.

[26] Su, S., J. Wu and J. Lee "A New Combination of Block Coded Modulation and Trellis Coded Modulation," International Conference on Communications, 1991, pp. 1080-1084.

[27] Kasami, T., T. Takata, T. Fujiwara and S. Lin "On Multilevel Block Nodulation Codes," *IEEE Tranmsactions on Information Theory*, July 1991, 37(4), pp. 965-975.

[28] Sundberg, C.-E. W. and N. Seshardi, "Coded Modulations for Fading Channels: An Overview," *European Transactions on Telecommunications*, May-June 1993, 4(3), pp. 309-324.

[29] M. Nakagami, "The m-distribution – A General Formula of Intensity Distribution of Rapid Fading," Statistical Methods in Radio Wave Propagation, W. G. Hoffman, Ed. Oxford, England, Pergamon, 1960, pp. 3-36.

[30] P. Crepeau, "Coding Performance on Generalized Fading Channels," MILCOM'88 conference proceedings, Sand Diego, CA, Oct. 23-26, 1988, pp. 15.1.1-15.1.7.

[31] W. Jakes, Microwave Mobile Communications, John Wiley and Sons., 1974

[32] L. J. Mason, "Error Probability Evaluation for Systems Employing Differential Detection in a Rician Fast Fading Environment and Gaussian Noise," *IEEE Transactions on Communications*, January 1987, vol. 35, pp. 39-46.

[33] C. E. Shannon, "A Mathematical Theory of Communication" *Bell System Technical Journal*, vol. 27, pp. 379-423, July 1948.

[34] T. Ericsson, "A Gaussian Channel with Slow Fading," *IEEE Transactions of Information Theory*, vol. 16, pp. 353-355, May 1970

[35] W. C. Y. Lee, "Estimate of Channel Capacity in Rayleigh Fading Environment," *IEEE Transactions on Vehicular Technology*, vol. 39, pp. 187-189, August 1990

[36] R. Buz, "Information Theoretic Limits on communication over Multipath Channels," Proceedings of the 1995 IEEE International Symposium on Information Theory, Whistler, Canada, pp. 151, September 1995

[37] J. M. Wozencraft and R. S. Kennedy, "Modulation and demodulation for probabilistic coding," *IEEE Transactions on Information Theory*, vol. IT-12, pp.291-297, 1966.

[38] J. Belfiore and K. Leeuwin-Boullé, "The Cutoff Rate of Time Correlated Fading Channels," *IEEE Transactions on Information Theory*, vol. IT-39, March 1993, pp. 612-617.

[39] D. Divsalar and M. Simon "Trellis Coded Modulation for 4800-9600 bits/s Transmission over a Fading Mobile Satellite Channel," *IEEE Journal on Selected Areas in Communications*, February 1987, 5(2), pp.162-175.

[40] D. Divsalar and M. Simon "The Design of Trellis Coded MPSK for Fading Channels : Performance Criteria," *IEEE Transactions on Communications*, September 1988, 36(9), pp. 1004-1012.

[41] J. Cavers and P. Ho "Analysis of the Error Performance of Trellis-Coded Modulations in Rayleigh-Fading Channels," *IEEE Transactions on Communications*, January 1992, 40(1), pp. 74-83.

[42] K. Chan and A. Bateman "The Performance of Reference-Based M-ary PSK with Trellis Coded Modulation in Rayleigh Fading," *IEEE Transactions on Vehicular Technology*, May 1992, 41(2), pp. 190-198.

[43] Y. Chen and C. Wei "Performance Evaluation of Convolutional Codes with MPSK on Rician fading Channels," *IEE Proceedings-F*, April 1987, 134(2), pp. 166-172.

[44] S. Wilson and Y. Leung "Trellis-Coded Phase Modulation on Rayleigh Channels," Proceedings of the International Conference on Communications, 1987, pp.739-743.

[45] D. Divsalar and M. Simon "The Design of trellis Coded MPSK for Fading Channels: Set Partitioning for Optimum Code design," *IEEE Transactions on Communications*, September 1988, 36(9), pp. 1013-1021.

[46] C. Schlegel and D. J. Costello, Jr. "Bandwidth Efficient Coding for Fading Channels : Code Construction and Performance Analysis," *IEEE Journal on Selected Areas in Communications*, December 1989, 7(9), pp. 1356-1368.

[47] S. Jamali and T. Le-Ngoc "A New 4-State 8PSK Scheme for Fast Fading, Shadowed Mobile Radio Channels," *IEEE Transaction on Vehicular Technology*, January 1991, 40(1), pp. 216-222.

[48] J. Du and B. Vucetic "New MPSK Trellis Codes for fading Channels," *Electronics Letters*, 2nd August 1990, 26(16), pp. 1267-1269.

[49] J. Du, and B. Vucetic, "New 16-QAM Trellis Codes for Fading Channels," *Electronics Letters,* 6th June 1991, pp. 1009-1010.

[50] N. Seshardi and C. W. Sundeberg "Multi-level Trellis Coded Modulations with Large Time Diversity for the Rayleigh Fading Channel," Proceedings of the Conference on Information Sciences and Systems (CISS), Princeton, March 1990, pp. 853-857.

[51] T. Woerz and J.Hagenauer "Multistage Coding and Decoding for a M-PSK System," GLOBECOM'90, pp. 698-703.

[52] N. Seshardi and C. W. Sundberg "Multi-level Block Coded Modulations for the Rayleigh Fading Channel", GLOBECOM'91, pp. 047-051.

[53] N. Seshardi and C. W. Sundberg "Multilevel Trellis Coded Modulations for The Rayleigh Fading Channel," *IEEE Transactions on Communications*, Spetember 1993, 41(9), pp. 1300-1310.

[54] J. Wu and S. Lin "Multilevel Trellis MPSK Modulation Codes for the Rayleigh Fading Channel," *IEEE Transactions on Communications*, September 1993, 41(9), pp. 1311-1318.

[55] E. Zehavi "8-PSK Trellis Codes for a Rayleigh Channel," *IEEE Transactions on Communications*, May 1992, 40(5), pp. 873-884.

[56] L.-F. Wei "Coded M-DPSK with Built-in Time Diversity for Fading Channels," *IEEE Transactions on Information Theory*, Nov. 1993, 39(6), pp.1820-1839.

[57] E. Leonardo L. Zheng and B. Vucetic, "Multidimensional MPSK Trellis Codes for Fading Channels," submitted to *IEEE Transactions on Information Theory*, 1994.

[58] A. Viterbi, E. Zehavi, J. K. Wolf and R. Padovani, "A Pragmatic Approach to Trellis Coded Modulation", *IEEE Communications Magazine*, 27(7) July, 1989, pp. 11-19.

[59] C. Heegard, S. Lery and W. Paik, "Practical coding for QAM transmission of HDTV", *IEEE Journal on Selected Areas in Communications*, 11(1), January 1993, pp. 111-118.

[60] P. Ho, J. Cavers and J. Varaldi, "The Effects of Constellation Density on Trellis-Coded Modulation on Fading Channels," *IEEE Transactions on Vehicular Technology*, August 1993, 42(3), pp. 318-325

[61] E. Biglieri, D. Divsalar, P. Mclane and M. Simon, Introduction to Trellis-Coded Modulation with Applications. Macmillan Publishing company, 1991.

[62] S. Jamali, and T. Le-Ngoc, Coded-Modulation Techniques for Fading Channels, Kluwer Academic Publishers, Boston, 1994.

[63] S. Lin and D. J. Costello, Jr., Error Control Coding: Fundamentals and Applications, Prentice-Hall, 1983

[64] P. Ho and D. Fung, "Error Performance of Interleaved Trellis-Coded PSK Modulations in Correlated Rayleigh Fading Channels," *IEEE Transactions on Communications*, December 1992, 40(12), pp. 1800-1809.

[65] M. Schwartz W. Bennett and S. Stein, Communication Systems and Techniques, McGraw-Hil, New York, 1966.

[66] M. D. Yacoubs, Foundations of Mobile Radio Engineering, CRC press, 1993.

[67] W. Lindsey, "Error probabilities for Rician fading multichannel reception of binary and N-ary signals," *IEEE Transactions on Information Theory*, vot. IT-10, Oct. 1964, pp. 339-350.

[68] A. A. Abu-Dayya and N. C. Beaulieu, "Microdiversity on Rician fading channels," *IEEE Transactions on Communications*, 42(6), June 1994, pp.2258-2267.

[69] L. Rasmussen and S. Wicker, "A Comparison of Two Combining Techniques for Equal Gain, Trellis Coded Diversity Receivers," *IEEE Transactions on Vehicular Technology,* May 1995, 44(2), pp.291-295.

[70] G. Femenias and R. Agusti, "Analysis of Predetection Diversity TCM-MPSK and Postdetection Diversity TCM-MDPSK Systems on

a Rayleigh Fading Channel," *IEEE Transaction on Communications,* Feb./Mar./Apr. 1995, pp. 374-385.

[71] J. Ventura-Traveset, G. Caire, E. Biglieri and G. Taricco, "Impact of Diversity Reception on Fading Channels with Coded Modulation. Part I: Coherent Detection," preprint.

[72] N. C. Beaulieu, "An infinite series for the computation of the complementary probability distribution function of a sum of independent random variables and its application to the sum of Rayleigh random variables," *IEEE Transactions on Communications*, 38(9), Sept. 1990, pp.1463-1474.

[73] C. S. Chang and P. J. Mclane, "Convolutional Codes with Correlated Diversity and Non-Coherent Orthogonal Digital Modulation," presented at the 1995 international conference on communications ICC'95, Seattle May 1995.

[74] N. Farvardin and V. Vaishampayan, "Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source-Channel Coding," *IEEE Transactions on Information Theory*, vol. 33, pp.827-838, November, 1987.

[75] E. Ayanoglu and R. Gray, "The Design of Joint Source and Channel Trellis Waveform Coders," *IEEE Transactions on Information Theory*, vol. 33, pp. 885-865, November 1987.

[76] N. Farvardin, "A Study of Vector Quantization for Noisy Channels," *IEEE Transactions on Information Theory*, vol. 36, pp. 799-809, July 1990.

[77] J. Hagenauer, "Source Controlled Channel Decoding, " accepted for publication, *IEEE Transactions on Communications*, March 1994.

[78] F. Alajaji, N. Phamdo and T. Fuja, "Channel Codes That Exploit the Residual Redundancy in CELP-Encoded Speech," submitted to *IEEE Transactions on Speech and Audio Proceesing*, May 1995.

[79] C. Hartmann and L. Rudolph, "An Optimum Symbol-by-Symbol Decoding Rule for Linear Codes," *IEEE Transactions on Information Theory*, vol. 22 pp. 514-517, September 1976.

[80] L. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284-287, March 1974.

[81] P. McAdam, L. Welch and C. Weber, "M.A.P. Bit Decoding of Convolutional Codes" Proceedings of the IEEE International Symposium on Information Theory, Asilomar, CA, pp. 91, January 1972.

[82] National Institute of Standards and Technology (NIST), "The DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus CD-ROM," NIST, October 1990.

[83] F. Alajaji, N. Phamdo, N. Farvardin and T. Fuja, "Detection of Binary Markov Sources Over Channels with Additive Markov Noise," *IEEE Transactions on Information Theory*, to appear, January 1996.

[84] A. Viterbi and J. Omura, Principles of Digital Communications and Coding, McGraw-Hill, 1979.

[85] M. Burnashev, "Improved Union Bound for Viterbi Decoder of Convolutional Codes," Proceedings of the 1995, International Symposium on Information Theory, Whistler, Canada pp.167

[86] S. Pietrobon, R. Deng, A. Lafanchere, G. Ungerboeck, and D. Costello, Jr., " Trellis-Coded Multidimensional Phase Modulation," *IEEE Transactions on Information Theory*, vol. 36, pp. 63-88, January 1990.

[87] National Communications System (NCS), "Details to Assist in Implementation of Federal Standard 1016 CELP," *Technical Information Bulletin 92-1*, Office of the Manager, NCS, Arlington, VA 22204-2198.

[88] R. Laroia, N. Phamdo and N. Farvardin, "Robust and Efficient Quantization of Speech LSP Parameters Using Structured Vector Quantization," *Proceedings ICASSP-91*, pp. 661-664, 1991