Coding and Channel Estimation for Block Fading Channels

by

Salam A. Zummo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering: Systems)
in The University of Michigan
2003

Doctoral Committee:
Professor Wayne Stark, Chair
Assistant Professor Achilleas Anastasopoulos
Assistant Professor Brian Noble
Associate Professor Kim Winick

In memory of my parents

and

To my siblings

ACKNOWLEDGMENTS

First thankfulness and praises are for Allah, the most merciful, and most compassionate. He blesses me with his ever-enduring mercies.

My great and deep appreciation go to my advisor Prof. Wayne Stark for his time devoted to my thesis completion and research development. His way of thinking and his extensive knowledge in the field of communications have been valuable resources. I would like to thank my thesis committee members Prof. Achilleas Anastasopoulos, Prof. Brian Noble and Prof. Kim Winick for their valuable time devoted in reviewing and correcting my thesis.

I would like to thank King Fahd University of Petroleum and Minerals for supporting me in completing my PhD program. I thank the Ministry of Higher education and the Saudi Cultural Attache at the USA for their support. Special thanks go to his Royal Highness the ambassador of the Kingdom of Saudi Arabia to the USA and to Dr. Jamil Makhadmi for their support.

This work was supported in part by the National Science Foundation under grant ECS-9979347 and the Office of Navel Research under grant N00014-03-1-0232.

My heartful thanks go to the memory of my parents for their prayers and encouragement. I wish to express my deep appreciation to my siblings for their support throughout my academic life. The unlimited support of my family made difficult tasks achievable.

I am grateful to my colleagues at the wireless communications lab and my friends in Ann Arbor for the unforgettable times we spent together. I benefited from discussing diverse subjects related to work and life in general.

PREFACE

Wireless communication channels are commonly modeled by time-varying random processes that exhibit memory. A simple model is the block fading channel model. In this thesis, we derive a union bound on the performance of binary coded systems over block fading channels. Noncoherent and coherent receivers are considered with different assumptions on the channel side information (SI) at the receiver. We derive the union bound for Rayleigh, Rician and Nakagami distributed block fading channels. Systems employing single and multiple transmit antennas are considered. From the results, the tradeoff between channel diversity and channel estimation is investigated. Moreover, we study the effect of the parameters of the channel and the space diversity on the optimal channel memory.

As an effort to solve the channel estimation problem in multi-antenna transmission over block fading channels, we derive a pilot-aided iterative receiver for joint decoding and channel estimation. In the receiver, initial channel estimation is obtained using orthogonal pilot sequence insertion, and then soft information from the decoder is used to update the estimation. Results show that using 3 iterations in the iterative receiver results in a performance close to that of the best achievable performance.

Trellis space-time (ST) codes using the I-Q encoding scheme provide a large time diversity. For performance evaluation and decoding, the "super-trellis" of the composite code is necessary which is too complex in general. In this thesis, the performance of I-Q ST codes is analyzed using the transfer functions of the component codes. Moreover, two low-complexity iterative receivers for I-Q ST codes are proposed and compared to the optimal decoding in complexity and performance. Results show that using the iterative receivers with 3 iterations provide most of the coding gain of the optimal decoding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

Introduction

1.1 Motivation

Emerging multi-media and internet applications require transmitting at high data rates with good quality. The role of wireless communication in information exchange is growing rapidly because of increased demand for mobility. Since limited physical resources are shared by many users in wireless networks, new technologies that are extremely efficient with respect to both power consumption and bandwidth should be deployed. A serious challenge to having good communication quality in wireless systems is the time-varying fading environments that wireless systems experience. When the signal is transmitted, it is reflected and scattered over surrounding objects, which causes the signal to be received over many different paths. These paths can add constructively or destructively. When they add destructively the received signal-to-noise ratio (SNR) can drop severely. Error correcting codes and diversity are standard approaches to mitigate multipath fading.

The fundamental theory of error correcting codes is often traced back to Shannon, who proved in [1] that data communications with rates below *channel capacity* and arbitrarily low error rates can be achieved over noisy channels by encoding the information properly. Encoding refers to imposing a structured redundancy on the information prior to transmission. At the receiver, this redundancy is used to correct errors imposed by the channel. While Shannon's results demonstrated the existence of good error correction codes, it has not given guidelines on how to construct such codes. Since Shannon's paper, most of the research on communication theory focused on inventing channel codes with practical encoders and decoders. Recently, codes approaching the channel capacity for additive white Gaussian noise (AWGN) channels have been discovered such as turbo codes [2] and low-density parity-check (LDPC) codes [3, 4]. In an AWGN channel, which is often used to model wired communication systems, a noise sample from a white Gaussian random process is added to the received signal at the receiver.

In wireless environments, multipath reception causes the energy of the received signal to be varying randomly. A standard model for multipath fading is the *Rayleigh* distribution. Turbo and LDPC codes have achieved performance close to the channel capacity in memoryless Rayleigh fading environments [5, 6]. However, wireless communication channels are commonly modeled by slowly time-varying random processes. In this thesis, we are interested in narrowband wireless channels, where the transmission bandwidth is much smaller than the carrier frequency. Wireless channels are characterized mainly by two parameters; namely, the coherence time and the coherence bandwidth [7]. The *coherence time* of a fading channel is the time duration in which the fading remains almost constant. Similarly, the *coherence bandwidth* is the frequency band in which the fading is almost constant. The channel is said to be *time-selective* if the symbol duration is long compared to the coherence time of the channel. Also, *frequency-selective* channels arise when the transmission rate is large compared to the coherence bandwidth of the channel. Channels with memory result when the channel varies slowly compared to the symbol duration. Thus the channel may remain constant during the transmission of a block of symbols. Furthermore, if the fading changes independently from one block to another, the channel is referred to as a *block fading* channel [8, 9].

In modern wireless communications, digital information signals are divided into small size frames and then transmitted. Each frame is usually encoded, modulated and then carried by a high-frequency carrier over a radio link. At the receiver, the reverse processes are performed by a demodulator and a channel decoder. The *effective channel diversity* can be thought of as the number of independent fading realizations available at the receiver to decode a frame. If insufficient number of realizations is available, the decoder will not be able to average over the channel statistics. In such environments, the performance of coded systems is degraded severely. However, independent fading realizations can be provided using diversity techniques [10]. In the simplest form, diversity can be obtained by transmitting the signal more than once and using the multiple copies of the received signal to improve the error rate. Repeating the transmission can be achieved over time, frequency or space. From a coding perspective this can be viewed as a repetition code, and hence more bandwidth efficient codes can be exploited to improve the performance. In general, encoding the information using an error correcting code is a way to provide time diversity at the receiver.

A common approach to break the channel memory and to spread burst errors in the decoder is to interleave the coded sequence prior to transmission. Conventionally, infinite interleaving is assumed in the literature [5, 11, 12] in order to simplify code design and performance analysis. However, infinite interleaving is impossible practically for delay-sensitive applications. Besides, using the infinite interleaving assumption in the performance analysis may not reflect the asymptotic behavior of the coded system at high SNR. If a coherent receiver is used, the phase of the fading process is needed for decoding. In general, *Channel side information* (SI) is defined as the phase and amplitude of the fading process. Thus if the receiver knows the channel SI perfectly, large channel diversity improves the system performance resulting in an optimal channel memory of unity. On the other hand, if the receiver estimates the channel, long channel memory provides more observations for each fading realization which permits a better channel estimation. Therefore, longer channel memory improves the performance if the frame size is infinite. However, if the frame size is finite, there exists a fundamental tradeoff between the channel diversity and channel estimation [13]. As the channel memory length increases, the channel diversity is reduced but the channel estimation becomes easier since more observations of each fading realization is available. On the converse, short channel memory increases the number of independent fading realizations available to the decoder, and hence it is able to average out the channel behavior at the cost of less accurate channel estimation.

In [13], Worthen *et al.* used the error exponent to find the optimal memory length of a communication system over some simple block memory channels. However, a method to analyze the performance of specific codes over block fading channels with arbitrarily chosen frame size and channel memory length is needed. Also, such a method is crucial in optimizing the channel memory of a coded system employing iterative decoding and channel estimation for example. In this thesis, we propose a union bound on the performance of binary convolutional and turbo codes over block fading channels. In deriving the bound, we assume uniform interleaving of the coded sequence prior to transmission over the channel and compute the distribution of error bits over the fading blocks. In order to evaluate the bound, the pairwise error probability corresponding to specific distribution patterns of the fading blocks is derived under different assumptions on the channel SI at the receiver. The proposed bound is used with imperfect SI at the receiver to investigate the tradeoff between the channel diversity and channel estimation, and hence optimize the memory length at which the system should operate.

If a line-of-site exists between the transmitter and receiver in addition to the multipath reception, the fading process is modeled by a Rician distribution [14]. In this model, the received signal is composed of two signal-dependent components; namely the specular and diffuse components. The specular component is due to the line-of-site reception and the diffuse component results from multipath reception. As the ratio of specular-to-diffuse component energy increases, the channel approaches the Gaussian channel, i.e., no fading. For Rician channels the importance of channel diversity becomes less significant as the specular-to-diffuse ratio increases because the fading channel becomes less random. As in the Rayleigh fading case, a method is needed to analyze the performance of coded systems over block fading channels with Rician distribution. In this thesis, the performance of coded systems over Rician block fading channels is studied using the union bound for block fading channels. Furthermore, we investigate the effect of channel memory on the system performance and its relation to the parameters of the channel such as the specular-to-diffuse ratio of the channel. The pairwise error probabilities for coherent and noncoherent receivers are derived. Furthermore, the effect of the specular-to-diffuse ratio of a Rician channel on the optimal channel memory is investigated.

Another popular model for the fading process is the Nakagami distribution [15], which provides a family of distributions that are well matched to measurements under different propagation environments [16, 17]. Nakagami distribution is characterized by the Nakagami parameter which indicates the fading severity. As the fading parameter of a Nakagami distributed channel is increased, the significance of diversity decreases because the channel approaches the no fading channel. Block fading channels with Nakagami distribution are encountered in many communication systems. Thus it is essential to analyze the performance of such systems. In this thesis, the performance of coded systems over Nakagami block fading channels is analyzed using the union bound for block fading channels. The effect of channel memory on the system performance and its relation to the channel fading characteristics is investigated. The pairwise error probability is derived for both coherent and noncoherent receivers.

An alternative approach to using error correcting codes to provide the receiver with diversity is to use multiple antennas at the transmitter or receiver. In transmit diversity [18], the information can be sent over different transmit antennas. When the information is encoded and different signals are transmitted over the transmit antennas, the resultant system is referred to as a space-time (ST) code [19]. A simple and elegant space-time block code (STBC) was proposed by Alamouti [20] to provide diversity at the transmitter. This idea was soon generalized by Tarokh *et al.* [21] to general number of transmit antennas. A differential scheme for STBCs that relaxes the need to estimate the channel was proposed in [22] at the cost of a 3-dB loss in the received SNR. Since the use of ST codes was originally initiated to mitigate fading channels with block fading behaviour, it is of interest to analyze the performance of such systems. In this thesis, the union bound for block fading channels is extended to coded STBC with perfect and imperfect SI at the receiver. From this, the effect of increasing the number of transmit antennas on the SNR degradation due to channel memory is investigated. It is well known [23] that channel estimation becomes more crucial to the performance of space-time trellis codes as the number of transmit antennas increases. In this thesis, a similar result for STBCs is derived and used with the union bound to investigate the tradeoff between channel diversity and channel estimation for binary coded STBCs over block fading channels. It is shown that increasing the number of transmit antennas provides more space diversity at the cost of more difficult channel estimation. Again, an interesting tradeoff between channel diversity and channel estimation exists. This tradeoff is investigated as well as the effect of the number of transmit antennas on the optimal channel memory.

Conventionally, channel estimation is performed independently from decoding. After the astonishing performance of the low complexity iterative decoders of turbo and LDPC codes, iterative receivers that jointly decode and estimate the channel were considered by several research groups. Examples of iterative receivers for joint decoding and channel estimation for single-antenna systems appeared in [24–27]. Similar receivers for multiple antenna systems are found in [28–32]. These iterative receivers use hard decisions from the decoder to update the channel estimation. In this thesis, an iterative receiver for joint decoding and channel estimation of coded multi-antenna systems is proposed.

The receiver uses soft information from the decoder to update channel estimation.

For applications that require low-complexity receivers and short delays, trellis codes are good candidates. If a trellis code is used over a block fading channel and the coded sequence is interleaved, then for low-to-medium SNR values, the channel can be approximated by a memoryless channel provided that the number of independent fading blocks is several times larger than the constraint length of the code. For this observation and due to the difficulty of optimizing trellis codes for block fading channels, we consider in this thesis a class of ST trellis codes that are appropriate for independent fading channels. These codes are referred to as I-Q ST codes [33], in which the I and Q channels are encoded independently using two independent encoders. The super-trellis of the composite code is necessary for performance evaluation and decoding, which is too complex in general. In this thesis, the pairwise error probability as well as the coding and diversity gains of I-Q ST codes are expressed as functions of the transfer functions of the component codes. Furthermore, the geometrical uniformity of I-Q ST codes is derived from the geometrical uniformity of the component codes. We propose two iterative receivers for I-Q ST codes. The two receivers are compared to the optimal decoding in complexity and performance. Results show that using 3 iterations in both receivers provide performance close to optimal decoding.

1.2 Block Fading Channel Model

In this section, we describe the channel model used throughout the thesis, i.e., the block fading channel model. Wireless communication channels are usually modeled by discrete-time systems. In this model, the signal is filtered using a matched filter and the output is sampled every T seconds. An equivalent model for the matched filter sampled output is given by

$$y_l = h_l s_l + z_l, \qquad l = 1, \ldots, N \tag{1.1}$$

where N is the frame size, s_l is the transmitted signal during the time interval l and z_l is a complex Gaussian random variable with zero mean and variance N_0, i.e., $\mathcal{CN}(0, N_0)$. The coefficient $h_l = a_l e^{j\theta_l}$ is a sample of the fading process whose phase θ_l is uniformly distributed over $[0, 2\pi)$. Here, $j = \sqrt{-1}$. In this thesis, the amplitude a_l is assumed to have either Rayleigh, Rician or Nakagami distribution.

A channel is said to be independent fading channel if $\{h_l\}$ are independent random variable. On the other hand, if the fading process is constant for a block of symbols, and changing independently from a block to another, the channel is said to be a block fading channel [8, 9]. The block of symbols for which the fading process is constant is called a *fading block*. Examples of systems modeled by block fading channels include orthogonal-frequency division multiplexing (OFDM), frequency hopping (FH) and time-division multiplexing (TDM) systems. In these systems, a transmission frame of length N symbols is affected by F independent fading realizations, i.e., $\{h_f\}_{f=1}^{F}$, resulting in a block of length $m = \lceil \frac{N}{F} \rceil$ symbols being affected by the same fading realization. The fading process remains constant for m symbols if the coherence time of the channel is much longer than the transmission duration of m symbols. This is a valid assumption for a practical range of mobility speeds and carrier frequencies. The independent assumption between different fading blocks is justified in the cases when FH or OFDM systems are used with the spacing between the carriers is larger than the coherence bandwidth of the channel. In these cases, each carrier will experience independent fading. In this model, F represents the number of hops in a FH system or the number of carriers in an OFDM transmission. Equivalently, m is the hop length in a FH system and the burst length in an OFDM system.

1.3 Channel Diversity vs. Channel Estimation

In this section, the tradeoff between the channel diversity and channel estimation is explained. In the block fading channel model, the effective channel diversity available at the decoder is defined as the number of independent channel realizations observed by a frame, i.e., F. If channel SI is available at the receiver, maximizing the effective channel diversity maximizes the performance. In general, a larger channel diversity makes the system performance close to the performance of the system over an AWGN channel with an SNR value equal to the expected value of the fading process. This is because the probability of bad channel conditions becomes smaller as the number of independent fading realizations increases. Moreover, fewer number of independent fading realizations increases the chance for experiencing poor channel realizations and hence the performance of a coded system degrades. Therefore, maximum channel diversity is obtained when the channel memory is smallest, i.e., $m = 1$. This is equivalent to the infinite interleaving assumption.

On the other hand, if channel estimation is performed at the receiver, longer memory permits better observation of the channel and hence improves the estimation quality. This suggests that there is a fundamental tradeoff between effective diversity and channel estimation and an optimal memory length exists for each coded system. Information theoretic bounding techniques were used in [27] to demonstrate this tradeoff for single-antenna systems. In this thesis, the tradeoff between channel diversity and channel estimation is investigated for different coded systems employing single and multiple transmit antennas, and the corresponding optimal channel memory is approximated.

1.4 Thesis Outline

In Chapter 2, we briefly review binary and ST coded systems. The maximum likelihood receivers of these codes are considered with different assumptions on the availability of channel SI at the receiver. Two binary codes are considered; namely, convolutional and turbo codes. Binary coded systems employing single and multiple transmit antennas are described. In addition, multi-antenna systems employing trellis and turbo ST codes are presented.

In Chapter 3, a union bound on the performance of binary coded systems employing single and multiple transmit antennas over block fading channels is derived. The pairwise error probability required for computing the bound is derived for the cases of coherent detection with perfect SI, imperfect SI and no amplitude SI at the receiver. Moreover, we consider the case of noncoherent detection using square-law combining. The union bound is evaluated for turbo and convolutional codes for different number of transmit antennas. From the obtained results, the effect of increasing the space diversity on the system performance is investigated. Moreover, the tradeoff between channel effective diversity and channel estimation is discussed as well as the relation of this tradeoff to the system space diversity.

In Chapter 4, the union bound for block fading channel is extended to more general fading models; namely, the Rician and Nakagami fading models. Expressions for the corresponding pairwise error probabilities for coherent detection with different assumptions on the channel SI at the receiver are derived. Also, noncoherent receivers employing square-law combining are considered. The effect of channel memory on the system performance and its relation to the parameters of the channel such as the specular-to-diffuse ratio in Rician channels and the fading characteristics in Nakagami channels are investigated. Furthermore, the tradeoff between the channel diversity and channel estimation is investigated as well as the effect of the specular-to-diffuse ratio of the channel on the optimal channel memory.

In Chapter 5, we propose an iterative receiver for joint decoding and channel estimation in ST coded systems. The performance of the receiver is simulated for turbo and trellis ST codes and the effect of the different parameters of the receiver on its performance and convergence is discussed. In addition, the cases of large frame sizes and large of number transmit antennas are considered. The optimal memory length for trellis and turbo ST code is investigated.

Chapter 6 considers the performance analysis and iterative receivers for I-Q trellis ST codes. The pairwise error probability of I-Q ST codes is derived as a function of the transfer functions of the I and Q codes, instead of that of the product of their trellises, i.e., the “super-trellis”. Moreover, two low-complexity iterative receivers are proposed and their performance and complexity are compared to those of the optimal decoder.

Finally in Chapter 7, we briefly summarize the main conclusions from the thesis and point out possible future research directions.

CHAPTER 2

Channel Codes for Fading Channels

Error correcting codes have been widely used in communication systems to enhance performance. Depending on the application and the channel under consideration, the system designer uses different channel codes. In this chapter, we review two classes of channel codes that are used throughout the thesis; namely, binary and space-time (ST) codes. In Section 2.1, binary coded systems utilizing single and multiple transmit antennas are described. Convolutional and turbo encoders are described and different receivers are derived for different assumptions on the channel side information (SI). Particularly, coherent receivers with perfect SI and no amplitude SI are considered as well as noncoherent square-law combining receivers. In Section 2.2, ST coded systems are considered. Trellis and turbo ST encoders and their corresponding receivers are described.

2.1 Binary Coded Systems

In this section, binary coded systems employing single and multiple transmit antennas are described. The general block diagram of a binary coded system over a fading channel is shown in Figure 2.1. The transmitter consists of a binary encoder (e.g., convolutional or turbo), random interleaver, a modulator and a multi-antenna transmission matrix. Time is divided into frames of

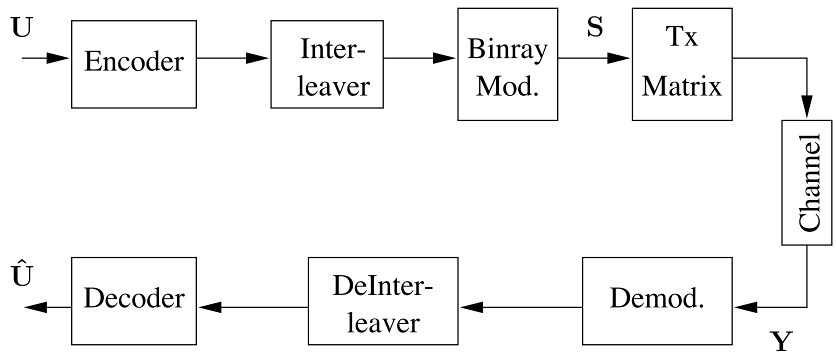

Figure 2.1: The structure of the binary coded system with a possible use of multi-antenna transmission.

duration NT, where T is the *transmission interval* of a bit. Throughout the thesis, the words "sequence" and "codeword" will be used interchangeably to mean a transmission frame. In each time interval of duration kT, a rate-R_c encoder maps k information bits into n coded bits, where $R_c = \frac{k}{n}$ is the code rate. Each coded bit is modulated to generate a signal using an equal-energy binary modulation. The modulation schemes considered in this thesis are binary phase-shift keying (BPSK) for coherent detection and binary frequency-shift keying (BFSK) for noncoherent detection. The channel we adopt is a block fading channel. Recall that in a block fading channel, each frame is subject to F independent fading realizations, where a block of length $m = \lceil \frac{N}{F} \rceil$ signals undergoes the same fading realization. The coded bits are interleaved prior to transmission over the channel in order to spread burst errors in the decoder, which result from low instantaneous SNR at the output of the demodulator due to fading. In the following sections, transmitters and receivers for binary coded systems employing single and multiple transmit antennas are described.

2.1.1 Single-Antenna Systems

In single-antenna systems, the modulated sequence is transmitted using one antenna. Coherent or noncoherent detection can be used at the receiver to detect the received sequence. In coherent receivers, the matched filter sampled output at time l in the f^{th} fading block is given by

$$y_{f,l} = \sqrt{E_s} h_f s_{f,l} + z_{f,l}, \tag{2.1}$$

where E_s is the average received energy, $s_{f,l} = (-1)^{c_{f,l}}$, where $c_{f,l}$ is the corresponding coded bit, and $z_{f,l}$ is an additive white noise modeled as independent zero-mean complex Gaussian random variables with variance N_0, i.e., $z_{f,l} \sim \mathcal{CN}(0, N_0)$. The coefficient h_f is the channel gain in fading block f which is modeled as complex Gaussian $\mathcal{CN}(0,1)$ and is written as $h_f = a_f \exp(j\theta_f)$, where a_f and θ_f are Rayleigh and uniform distributed, respectively.

The receiver employs maximum likelihood (ML) sequence decoding which is optimal for frame error probability. In this rule, the decoder chooses the codeword $\mathbf{S} = \{s_{f,l}, f = 1, \ldots, F, l = 1, \ldots, m\}$ that maximizes the likelihood function $p(\mathbf{Y}|\mathbf{S})$, where $\mathbf{Y} = \{y_{f,l}, f = 1, \ldots, F, l = 1, \ldots, m\}$. In coherent reception, some information about the channel phase and amplitude are available. The case where perfect SI is available at the receiver is a hypothetical scenario that predicts the best performance of the code. A practical assumption considers estimating the channel amplitude and phase resulting in imperfect SI at the receiver. This case is discussed in Chapter 3. If perfect SI is available at the receiver, the decoder chooses the codeword S that maximizes the following metric

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = \sum_{f=1}^{F} \sum_{l=1}^{m} a_f s_{f,l} \mathrm{Re}\{y_{f,l}\}, \tag{2.2}$$

where $\mathrm{Re}\{.\}$ represents the real part of a complex number. In the literature, coherent detection with no amplitude SI was considered extensively as an intermediate case between coherent detection with perfect SI and noncoherent detection, where both amplitude and phase are not available at the receiver. A suboptimal decoding metric was used in [12, 34] due to its analytical tractability, where the decoder chooses the codeword S that maximizes

$$\mathbf{m}(\mathbf{Y},\mathbf{S}) = \sum_{f=1}^{F}\sum_{l=1}^{m} s_{f,l}\mathrm{Re}\{y_{f,l}\}. \tag{2.3}$$

In noncoherent systems, BFSK signaling is used where the carrier frequency of the modulated signal is set to be one of two frequencies according to whether the coded bit $c = 0, 1$. The carrier frequencies are chosen such that the resultant signals are orthogonal. At the receiver, square-law combining [35] is employed for each received signal resulting in a suboptimal receiver with respect to minimizing the frame error probability. The outputs of the demodulator are represented by

$$\begin{aligned} r_{f,l}^{(I,0)} &= \sqrt{E_s}a_f\delta(c_{f,l},0)\cos(\theta_f) + \eta_{f,l}^{(I,0)} \\ r_{f,l}^{(Q,0)} &= \sqrt{E_s}a_f\delta(c_{f,l},0)\sin(\theta_f) + \eta_{f,l}^{(Q,0)} \end{aligned} \tag{2.4}$$

$$\begin{aligned} r_{f,l}^{(I,1)} &= \sqrt{E_s}a_f\delta(c_{f,l},1)\cos(\theta_f) + \eta_{f,l}^{(I,1)} \\ r_{f,l}^{(Q,1)} &= \sqrt{E_s}a_f\delta(c_{f,l},1)\sin(\theta_f) + \eta_{f,l}^{(Q,1)} \end{aligned} \tag{2.5}$$

where $r_{f,l}^{(I,c)}$ and $r_{f,l}^{(Q,c)}$ for $c = 0, 1$ correspond to the correlation of the received signal with the inphase and quadrature dimensions of the signal corresponding to a coded bit c. In (2.5), θ_f is the unknown phase of the received signals in block f, $\delta(x, y) = 1$ if $x = y$ and $\delta(x, y) = 0$ otherwise; and $\eta_{f,l}^{(I,0)}$, $\eta_{f,l}^{(Q,0)}$ $\eta_{f,l}^{(I,1)}$ and $\eta_{f,l}^{(Q,1)}$ are independent variables with $\mathcal{N}(0, \frac{N_0}{2}$ distribution. The decoder chooses the codeword S that maximizes

$$\mathbf{m}(\mathbf{R},\mathbf{S}) = \sum_{f=1}^{F}\sum_{l=1}^{m}(r_{f,l}^{(I,c)})^2 + (r_{f,l}^{(Q,c)})^2, \tag{2.6}$$

where $\mathbf{R} = \{r_{f,l}^{(I,c)}, r_{f,l}^{(Q,c)}, f = 1, \ldots, F, l = 1, \ldots, m\}$. Note that this decoder makes no use of channel SI in decoding.

2.1.2 Multi-antenna Systems

In the discussion to follow we describe multi-antenna transmitters that concatenate error correcting codes with space-time block codes (STBC)s [20]. Throughout the thesis, small letters in bold are used to denote column vectors. Moreover, we consider systems employing n_t transmit antennas and single receive antenna, but all derivations and results to be presented apply to multiple receive antennas. After encoding and interleaving, each n_t signals are mapped into a $n_t \times n_t$ transmission matrix $\mathcal{G}_{n_t}$ as shown below for the cases of $n_t = 2$ and 4

$$\mathcal{G}_2 = \begin{pmatrix} s_1 & s_2 \\ -s_2 & s_1 \end{pmatrix}, \qquad \text{and} \qquad \mathcal{G}_4 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ -s_2 & s_1 & -s_4 & s_3 \\ -s_3 & s_4 & s_1 & -s_2 \\ -s_4 & -s_3 & s_2 & s_1 \end{pmatrix}. \tag{2.7}$$

More examples of real and complex orthogonal matrices were presented in [21] for different values of n_t. Note that the rows of $\mathcal{G}_{n_t}$ are orthogonal to each other to enable linear complexity detection [20]. The transmission of $\mathcal{G}_{n_t}$ takes place in a *time slot* of duration n_tT, where the i^{th} row of $\mathcal{G}_{n_t}$ is transmitted in the i^{th} time interval of the time slot using the n_t transmit antennas. Equivalently, the i^{th} column of $\mathcal{G}_{n_t}$ is transmitted over the i^{th} transmit antenna during the time slot of duration n_tT. Thus the resulting STBC has full rate, i.e., one coded bit is transmitted every T seconds.

To be able to detect STBCs, the fading process from each antenna should remain constant for at least one time slot, i.e., n_t time intervals. This constrains the channel memory length to be a multiple of n_t, where each fading block contains $\frac{m}{n_t}$ time slots each of length n_t. In the rest of the thesis, the subscript of $\mathcal{G}_{n_t}$ is omitted to simplify notation. The received vector in time slot l of the

f^{th} fading block is given by

$$\mathbf{y}_{f,l} = \sqrt{E_s}\mathcal{G}_{f,l}\mathbf{h}_f + \mathbf{z}_{f,l}, \tag{2.8}$$

where $\mathbf{z}_{f,l}$ is a length-n_t column random vector with a distribution $\mathcal{CN}(\mathbf{0}, N_0\mathbf{I})$ and $\mathbf{I}$ denotes the $n_t \times n_t$ identity matrix. The vector $\mathbf{h}_f$ contains the channel gains from the transmit antennas in fading block f and is modeled as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$. This indicates that the gains from different transmit antennas are uncorrelated, which is reasonable if the distance separating different antennas exceeds half the wavelength of the carrier [10]. The receiver employs sequence decoding based on the detection scheme of STBC in [20]. Equivalently, the decoder chooses the codeword S that maximizes the metric

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = \sum_{f=1}^{F} \sum_{l=1}^{m/n_t} \mathrm{Re}\{\mathbf{y}^*_{f,l}\mathcal{G}_{f,l}\mathbf{h}_f\}, \tag{2.9}$$

where $(.)^*$ denotes the complex conjugate of a complex vector. The channel codes that are concatenated with the single and multiple antennas transmitters are described as follows.

2.1.3 Convolutional Codes

A popular channel code with an easy encoding scheme is the convolutional code. A convolutional encoder is a finite state machine consisting of v shift registers. During a time interval of length T, the encoder receives an input vector $\mathbf{u} = \{u_l\}_{l=1}^{k}$ of length k and produces a code vector $\mathbf{c} = \{c_l\}_{l=1}^{n}$ of length n. The l^{th} coded bit among the n coded bits at the output of the encoder is given by

$$c_l = u_1 + \sum_{j=1}^{v} g_{l,j} r_j, \qquad \text{mod } 2, \tag{2.10}$$

where r_j is the content of the j^{th} shift register, and $\mathbf{g}_l = \{g_{l,j}\}_{j=0}^{v}$ is the generator polynomial of the l^{th} coded bit where $g_{l,j} \in \{0, 1\}$. The generator polynomials are

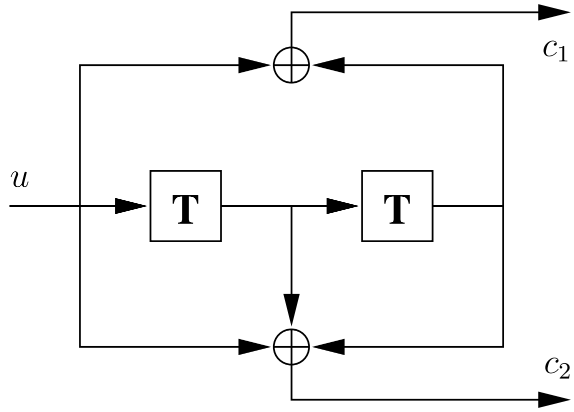

Figure 2.2: The encoder of a rate-$\frac{1}{2}$ (5,7) NSC.

usually expressed in octal numbers, e.g., a generator polynomial $\mathrm{g}_i = 1101$ is represented as $\mathrm{g}_i = 15$. In general, a convolutional code is represented using its generator polynomials as $(\mathrm{g}_1, \mathrm{g}_2, \ldots, \mathrm{g}_n)$. The encoder for a (5,7) convolutional code is shown in Figure 2.2. This code is non-systematic code (NSC) since the coded bits are not partitioned into information (systematic) bits and parity bits. Convolutional codes which partition the information and parity bits are referred to as systematic codes.

A recursive systematic code (RSC) is a systematic code with feedback in the encoder. The encoder of a RSC with $k = 1$ is shown in Figure 2.3. An RSC code is represented using its generator polynomials as $(\mathrm{g}_1, \ldots, \mathrm{g}_n/\mathrm{g}_b)$, where $\mathrm{g}_b = \{g_{b,j}\}_{j=1}^{v}$ is the feedback generator polynomial. In Figure 2.3, the modified input $\tilde{u}$ is given by

$$\tilde{u} = u + \sum_{j=1}^{v} g_{b,j} r_j, \qquad \text{mod 2}. \tag{2.11}$$

The coded bits are given by

$$c_l = \tilde{u} + \sum_{j=1}^{v} g_{l,j} r_j, \qquad \text{mod 2}. \tag{2.12}$$

The frame error probability is minimized by employing a ML sequence decoder using a Viterbi algorithm [11]. The Viterbi algorithm uses the metrics described in Section 2.1.1 for different assumptions on channel SI at the receiver.

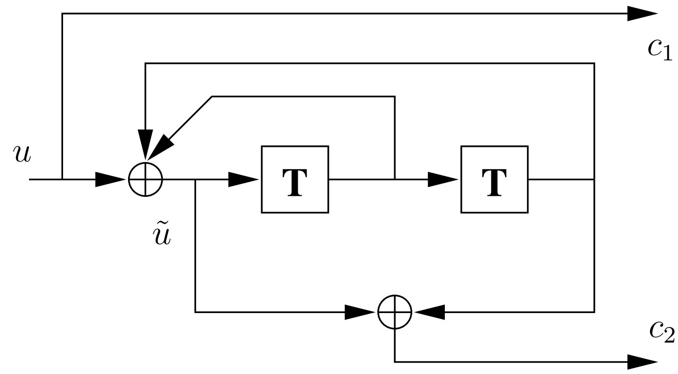

Figure 2.3: The encoder of a rate-$\frac{1}{2}$ (5/7) RSC.

2.1.4 Turbo Codes

Originally, turbo codes referred to parallel concatenated codes [2]. The general turbo encoder is shown in Figure 2.4, which consists of two constituent codes separated by an interleaver. The constituent code can be a convolutional or a block code. If the constituent codes are block codes, the resultant code is referred to as a product code [36], where parallel concatenation of convolutional codes results in turbo codes [2]. A turbo code employing a RSC, with a respective feedforward and feedback generator polynomials of g and g_b, is denoted by $(1, (\mathrm{g}/\mathrm{g}_b), (\mathrm{g}/\mathrm{g}_b))$. The information sequence is encoded by the first constituent code, and by second constituent code after being randomly interleaved. The systematic (information) sequence $\mathbf{C}$ as well as the parity sequences at the output of the two constituent codes, $\mathbf{C}_1$ and $\mathbf{C}_2$ are multiplexed to form a length-N codeword $\mathbf{C}$. The overall code rate is $R_c = \frac{k}{2n+k}$, where n is the number of parity bits at the output of each RSC and k is the number of input bits to the turbo encoder in each time interval. Note that the interleaver size is $\tilde{N} = R_c N$.

The block diagram of the turbo decoder is shown in Figure 2.5. It consists of two soft-input soft-output (SISO) decoders, one for each constituent code, an interleaver similar to that used in the encoder and a deinterleaver that reverses the effect of the interleaver. The decoder works iteratively and the SISO blocks exchange soft information about information bits in each iteration. Each SISO block employs a maximum aposteriori probability (MAP) rule which

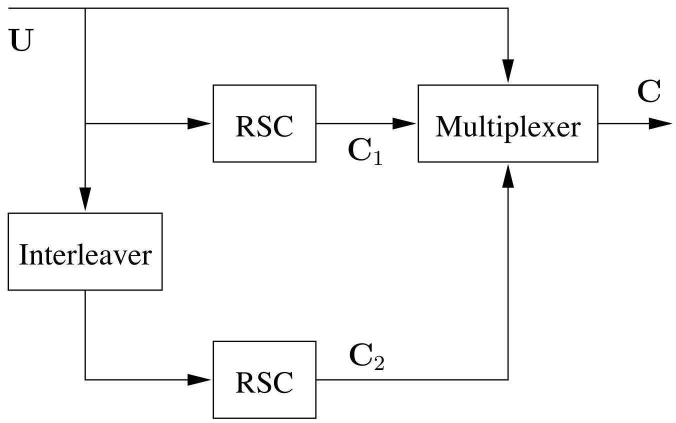

Figure 2.4: The block diagram of a turbo encoder.

computes the bit aposteriori probability $p(u_l|\mathbf{Y})$ of each bit. Using the BCJR algorithm [2], the likelihood function of the bit u_l is given by

$$\Lambda_l = \log \frac{p(u_l = 1|\mathbf{Y})}{p(u_l = 0|\mathbf{Y})}, \qquad l = 1, \ldots, \tilde{N}, \tag{2.13}$$

$$= \log \frac{\sum\limits_{(m,m^{'}):u_l=1} \gamma_l(m, m^{'})\alpha_{l-1}(m^{'})\beta_l(m)}{\sum\limits_{(m,m^{'}):u_l=0} \gamma_l(m, m^{'})\alpha_{l-1}(m^{'})\beta_l(m)} + \Lambda_l^e, \qquad l = 1, \ldots, \tilde{N}, \tag{2.14}$$

where $\tilde{N}$ is the length of the information sequence and $\gamma_l(m, m^{'}) = p(y_{s,l}, y_{i,l}|u_l)$ is the channel transition probability for a trellis transition from state m at time l to state $m^{'}$ at time $l+1$ in the i^{th} constituent code. The first part of Λ_l is due to the contribution of the systematic bit and the constraint in the i^{th} constituent code, where the second part, i.e., Λ_l^e is the extrinsic information given by

$$\Lambda_l^e = \log \frac{p(u_l = 1)}{p(u_l = 0)}, \qquad l = 1, \ldots, \tilde{N}, \tag{2.15}$$

where $p(u_l = 0)$ and $p(u_l = 1)$ are the apriori probabilities for the bit u_l to be 0 or 1, respectively, which are assumed to be equal in the first iteration. The variables $\alpha_l(m^{'})$ and $\beta_l(m)$ are the standard forward and backward recursions in the BCJR algorithm [37]. For SISO decoder i, they are given by $\alpha_l(m) = p(S_l = m, \{y_s\}_1^l, \{y_i\}_1^l)$ and $\beta_l(m) = p(S_{l+1} = m, \{y_s\}_{l+1}^{\tilde{N}}, \{y_i\}_{l+1}^{\tilde{N}})$, where $p(S_l =$

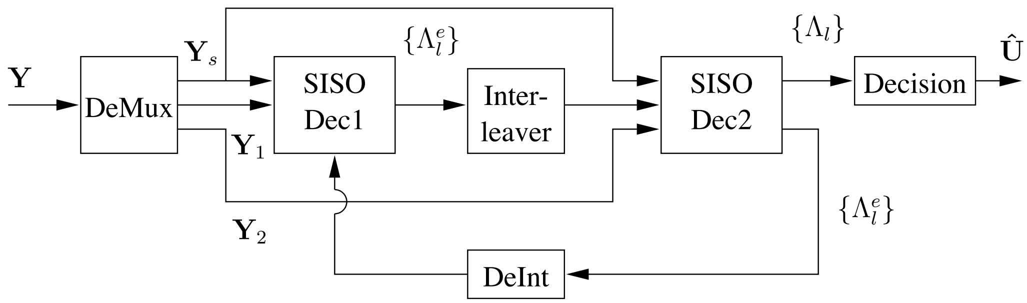

Figure 2.5: The turbo decoder.

$m, \{y_s\}_1^l, \{y_i\}_1^l)$ is the joint probability density function of the encoder being at state m at time l and the sequences $\{y_s\}_1^l$ and $\{y_i\}_1^l$. A similar definition holds for the function $p(S_{l+1} = m, \{y_s\}_{l+1}^{\tilde{N}}, \{y_i\}_{l+1}^{\tilde{N}})$. The variables $\alpha_l(m)$ and $\beta_l(m)$ are computed in each SISO decoder, respectively as

$$\alpha_l(m) = \sum_{m'} \gamma_l(m, m')\alpha_{l-1}(m'), \qquad l = 1, \ldots, \tilde{N}, \tag{2.16}$$

$$\beta_l(m) = \sum_{m'} \gamma_{l+1}(m, m')\beta_{l+1}(m'), \qquad l = 1, \ldots, \tilde{N}. \tag{2.17}$$

In the final iteration, a decision is made on information bits according to

$$\hat{u}_l = \begin{cases} 0, & \text{if } \Lambda_l \leq 0, \\ 1, & \text{if } \Lambda_l > 0. \end{cases} \tag{2.18}$$

2.2 Space-Time Coded Systems

In this section we describe the ST coded systems. The section starts with a general system description. Then, trellis and turbo ST codes are discussed. The block diagram of the general ST coded system is shown in Figure 2.6. The transmitter consists of a ST encoder, an interleaver, a modulator and n_t transmit antennas. During a frame transmission period of length NT, the input to the transmitter is a length-N sequence $\mathbf{U} = \{\mathbf{u}_l\}_{l=1}^N$ of input vectors each of length k. Each component of the input vector $\mathbf{u}_l$ is assumed to be from a binary

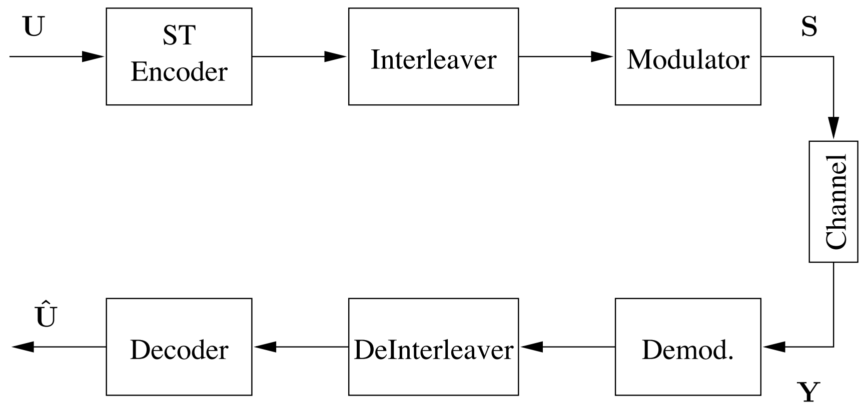

Figure 2.6: A general ST coded system.

alphabet. The ST encoder produces a length-N sequence $\mathrm{S} = \{\mathrm{s}_l\}_{l=1}^{N}$ of signal vectors each of length n_t. The i^{th} element of each signal vector s_l^i is an element of an M-ary signal constellation, such as MPSK or M-QAM, which is modulated and transmitted using the i^{th} transmit antenna in the time interval l. Therefore, the overall system throughput is k/T bits/s. Before being modulated, the signal vectors are interleaved to avoid burst errors in the decoder. Note that we used s instead of c to denote the output of the encoder because each element of s is a signal from a constellation as opposed to being a coded bit in the case of binary codes. This notation is used for ST coded systems throughout the thesis.

The received signal at time interval l in fading block f is

$$y_{f,l} = \sqrt{E_s} \sum_{i=1}^{n_t} h_f^i s_{f,l}^i + z_{f,l}, \qquad (2.19)$$

where E_s is the average transmitted energy at each transmit antenna and $z_{f,l}$ is an additive white noise modeled as $\mathcal{CN}(0, N_0)$. The coefficient h_f^i is the channel gain from the i^{th} transmit antenna in fading block f, which is modeled as $\mathcal{CN}(0,1)$. As in the case of STBCs, it is assumed that the channel gains from different transmit antennas are uncorrelated, i.e. $\mathrm{E}[h_f^p h_f^{q*}] = 0$ for $p \neq q$, where $(.)^*$ denotes the complex conjugate. The decoder employs a ML decoding rule to minimize the frame error probability. Assuming perfect channel SI at the

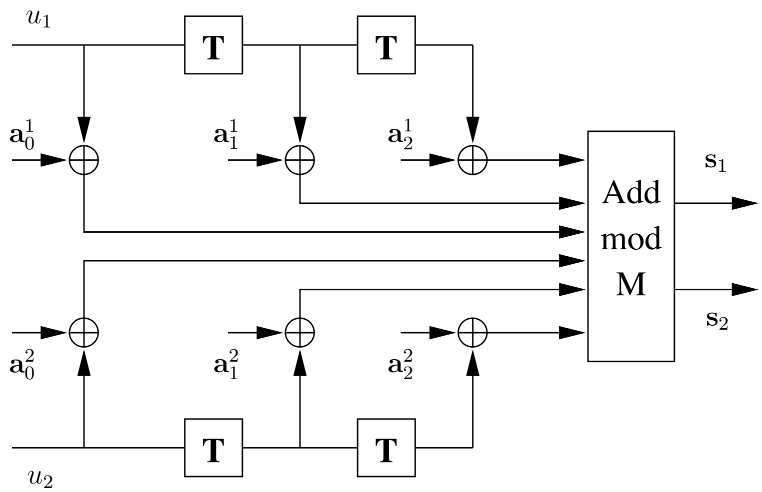

Figure 2.7: The encoder of a trellis ST code with $k = 2$ and $v = 4$.

receiver, the ML decoding rule chooses a codeword S that maximizes

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = -\sum_{f=1}^{F}\sum_{l=1}^{m}\left|y_{f,l} - \sqrt{E_s}\sum_{i=1}^{n_t} h_f^i s_{f,l}^i\right|^2. \tag{2.20}$$

2.2.1 Trellis ST Codes

The trellis an ST encoder considered in this thesis consists of k parallel trellis subcodes, each encodes one of the k input bits. The resulting code has a throughput of k/T bits/s. This is a generalization of the encoder described in [38]. In general, each subcode has $\tilde{v}$ shift registers. The contents of the shift registers of the q^{th} subcode are denoted by $\{r_l^q\}_{l=1}^{\tilde{v}}$, for $q = 1, \ldots, k$. Note that the total code memory is $v = k\tilde{v}$. Figure 2.7 shows a trellis encoder with $k = 2$ and $v = 4$. The signal vector at the output of the encoder at time l is given by

$$\mathbf{s}_l = \sum_{q=1}^{k}\sum_{j=1}^{\tilde{v}} u_{l-j}\mathbf{a}_{l-j}^q \qquad \bmod M, \tag{2.21}$$

where $\mathbf{a}^q = \{a^{q,i}\}_{i=1}^{n_t}$ is a length-n_t vector with elements drawn from a set with cardinality M, i.e., $a^{q,i} \in \{0, 1, \ldots, M-1\}$.

For the purpose of iterative joint decoding and channel estimation to be discussed in Chapter 5, a SISO decoder is necessary. Using the BCJR algorithm

[37], soft information about signal vectors in the frame is computed using

$$p\left(\mathbf{s}_l \middle| \mathbf{Y}, \mathbf{H}\right) = C \sum_{(m,m^{'}):\mathbf{s}_l} \gamma_l(m, m^{'}) \alpha_{l-1}(m^{'}) \beta_l(m), \qquad l = 1, \ldots, N, \tag{2.22}$$

where C is a normalization constant, $\gamma_l(m, m^{'}) = p(y_l|\mathbf{s}_l, \mathbf{h}_l)$ is the branch metric for a trellis transition from state m at time l to state $m^{'}$ at time $l+1$, where $\mathbf{h}_l$ contains the channel gains of the fading block in which time interval l lies. The variables α_l and β_l are the standard forward and backward recursions in the BCJR algorithm computed as in Section 2.1.4.

2.2.2 Turbo ST Codes

In [39], Yuan *et al.* presented ST turbo scheme consisting of two constituent codes and an interleaver as shown in Figure 2.8. During a transmission interval of length NT, the input to the encoder is a length-$N/2$ sequence of information vectors $\mathbf{U} = \{\mathbf{u}_l\}_{l=1}^{N/2}$, each of length k bits. The information sequence is encoded using the first constituent code into $\mathbf{S}_1 = \{\mathbf{s}_l\}_{l=1}^{N/2}$. The information sequence is randomly interleaved and encoded by the second constituent code into $\mathbf{S}_2$. The two signal sequences out of the constituent codes are multiplexed to form the signal frame $\mathbf{S} = \{\mathbf{s}_l\}_{l=1}^{N}$. Hence, the throughput of the resultant turbo code is $\frac{k}{2}$ bits/s.

The constituent codes are recursive trellis ST codes that output signal vectors of length n_t, with elements from an M-ary signal constellation [39]. Figure 2.9 shows a recursive ST trellis with $k = 2$ and $v = 4$. In the l^{th} time interval, the q^{th} recursive subcode in the ST encoder generates a modified input bit as

$$r_0^q = u_l + \sum_{j=1}^{\tilde{v}} r_{l-j}^q \qquad \text{mod } 2. \tag{2.23}$$

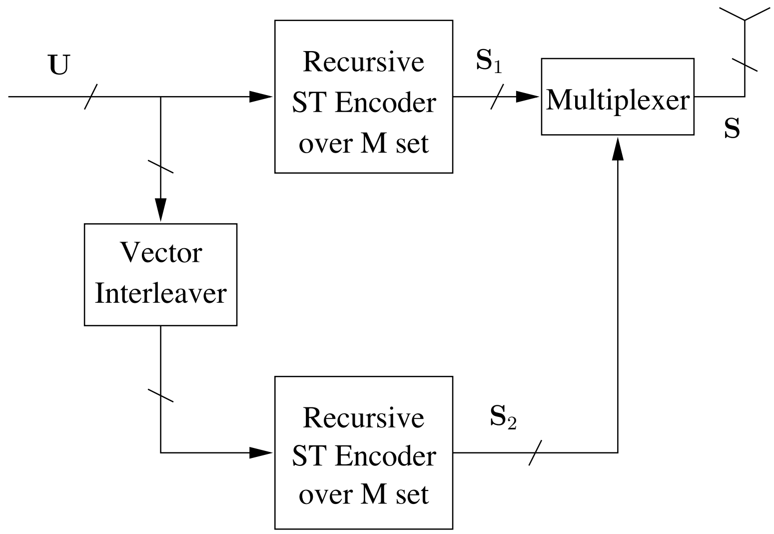

Figure 2.8: The structure of a turbo ST encoder.

The output signal vector is given by

$$\mathbf{s}_l = \sum_{q=1}^{k}\sum_{j=1}^{\tilde{v}} r_{l-j}^{q}\mathbf{a}_{l-j}^{q} \qquad \bmod M, \tag{2.24}$$

where $\mathbf{a}^h$ is the same as that for non-recursive trellis ST codes.

The decoder of ST turbo codes is based on the turbo decoding principle [2], in which two SISO modules are used, one for each constituent code. Aposteriori probabilities of input vectors in the frame are computed as

$$p\left(\mathbf{u}_l \middle| \mathbf{Y}, \mathbf{H}\right) = K \sum_{(m,m^{'}):\mathbf{u}_l} \gamma_l(m, m^{'})\alpha_{l-1}(m^{'})\beta_l(m), \qquad l = 1, \ldots, N, \tag{2.25}$$

where $\gamma_l(m, m^{'}) = p(y_l|\mathbf{u}_l, \mathbf{h}_l)p(\mathbf{u}_l)$ and $p(\mathbf{u}_l)$ is the apriori probability of the input vector $\mathbf{u}_l$ computed in the other SISO block, which is set equal for all input vectors in the first iteration. Information passed between the SISO modules is extrinsic information which is the probability in (2.25) after removing the contribution of $p(\mathbf{u}_l)$. The algorithm runs for a number of iterations, and in the last iteration the decoder chooses the information vector with the largest aposteriori probability. For the purpose of iterative decoding and channel estimation,

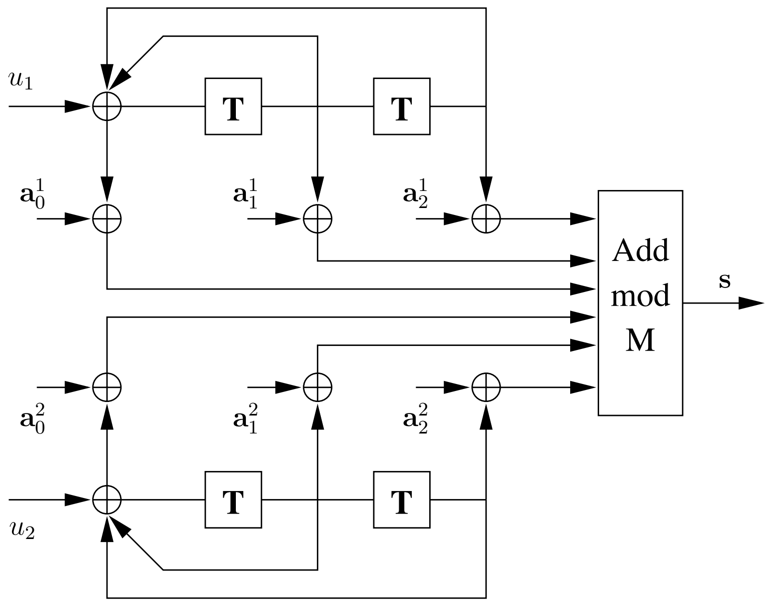

Figure 2.9: The encoder of a recursive trellis ST code with $k = 2$ and $v = 4$.

SISO modules compute soft information about signal vectors using (2.22).

CHAPTER 3

Performance of Binary Coded Systems over Rayleigh Block Fading Channels

In this chapter, a union bound for binary coded systems over block fading channel is derived. Systems employing single and multiple transmit antennas are considered. Conventionally, the performance of coded systems over ideally interleaved channels is analyzed using the union bound, which was computed by Viterbi [11] for specific convolutional codes using the weight enumerator of the code. In [12], Divsalar *et al.* derived similar bounds for trellis codes with perfect and no amplitude SI assumptions, where union bound for turbo codes were presented by Hall *et al.* [5]. Multi-frequency trellis codes [40] are special codes for block fading channels. At each state transition in multi-frequency trellis codes, the encoder produces signals, each is transmitted over one fading block. Hence, the number of output signals at each transition equals the number of fading blocks in the frame. This makes these codes too complicated for arbitrary number of fading blocks. In [41], the union bound for multi-frequency convolutional codes was evaluated. Several block and trellis codes designed for the block fading channels were presented in [42]. Also, in [42], Knopp *et al.* derived the outage probability of binary coded systems over block fading channels. Malkamaki *et al.* [43] derived random coding upper bounds on the average error probability of coded diversity over block fading channels. Also, an expression for the maximum achievable diversity as a function of the number of fading blocks and the code rate was presented.

Despite the efforts to analyze the performance of coded systems over block fading channels, a method for evaluating the error probability of binary codes is needed. In this chapter, we derive a union bound on the error probability of binary coded systems over block fading channels. Using the union bound the channel memory can be optimized for a coded system employing iterative decoding and channel estimation.

This chapter is organized as follows. In Section 3.1 the union bound for communication systems employing arbitrary convolutional and turbo codes over block fading channels is derived. Communication systems employing single and multiple transmit antennas are considered in Sections 3.2 and 3.3, respectively. Expressions for pairwise error probability of the coded systems are derived for noncoherent and coherent receivers with different assumptions about the channel SI at the receiver. Results are presented and discussed for each case following the derivation of the corresponding result.

3.1 Union Bound

In this section, a union bound on the bit and frame error probability of convolutional and turbo codes over block fading channels is derived. Throughout the thesis, the subscripts c, u and b are used to denote conditional, unconditional and bit error probabilities, respectively. For linear convolutional codes with k input bits, the bit error probability is upper bounded [11] as

$$P_b \leq \frac{1}{k} \sum_{d=d_{\min}}^{N} w_d P_u(d), \tag{3.1}$$

where $d_{\min}$ is the minimum distance of the code, $P_u(d)$ is the unconditional pairwise error probability defined as the probability of decoding a received sequence as a weight-d codeword given that the all-zero codeword is transmitted.

In (3.1), $w_d = \sum_{i=1}^{N} iA_{i,d}$ is the number of codewords with output weight d, where $A_{i,d}$ is the number of codewords with output weight d and input weight i. The weight distribution $\{w_d\}_{d=d_{\min}}^{N}$ is obtained directly from the weight enumerator of the code [11]. For turbo codes with code interleaver size $\tilde{N} = R_c N$, the union bound for a particular interleaver is difficult to evaluate. However, if we consider the ensemble of codes generated by all possible interleavers, then we can obtain a bound on the bit error probability of such a code by averaging over all possible interleavers [44]

$$P_b \leq \sum_{i=1}^{\tilde{N}} \frac{i}{\tilde{N}} \binom{\tilde{N}}{i} \sum_{d=d_{\min}}^{N} p(i,d) P_u(d), \tag{3.2}$$

where $p(i,d)$ is the probability of having an input sequence with weight i and an output codeword with weight d. For a turbo code with two component codes, $p(i,d)$ is given by

$$p(i,d) = \sum_{\{d_0,d_1,d_2\}:d_0+d_1+d_2=d} p_0(i,d_0) p_1(i,d_1) p_2(i,d_2), \tag{3.3}$$

where $p_0(i,d_0) = \delta(i,d_0)$ represents the systematic bit; and $p_j(i,d_j) = A_{i,d_j} / \binom{\tilde{N}}{i}$ for $j = 1,2$, accounts for the code interleaver. The frame error probability of turbo codes is given by

$$P_f \leq \sum_{i=1}^{\tilde{N}} \binom{\tilde{N}}{i} \sum_{d=d_{\min}}^{N} p(i,d) P_u(d). \tag{3.4}$$

3.1.1 Union Bound for Block Fading Channels

Recall that in the block fading channel model, a frame of size N is affected by F fading realizations. Thus the fading realization stays constant for a duration of a fading block composed of $m = \lceil \frac{N}{F} \rceil$ signals. In this case, the pairwise error probability $P_u(d)$ is a function of the distribution of the d nonzero bits over the F fading blocks. In the following, this distribution is quantified assuming uniform channel interleaving of the coded bits over the fading blocks. Figure 3.1 shows the distribution of the d nonzero bits in a weight-d erroneous codeword over the F fading blocks. Denote the number of fading blocks with weight v by f_v and define $w = \min(m, d)$, then the fading blocks are distributed according to the pattern $\mathbf{f} = \{f_v\}_{v=0}^{w}$ if the following conditions are satisfied

$$F = \sum_{v=0}^{w} f_v, \qquad d = \sum_{v=1}^{w} v f_v. \tag{3.5}$$

Denote by $L = F - f_0$ the number of fading blocks with nonzero weights. Then $P_u(d)$ averaged over all possible fading block patterns is given by

$$P_u(d) = \sum_{L=\lceil d/m \rceil}^{d} \sum_{f_1=1}^{L_1} \sum_{f_2=1}^{L_2} \cdots \sum_{f_w=1}^{L_w} P_u(d|\mathbf{f}) p(\mathbf{f}), \tag{3.6}$$

where

$$L_v = \text{min} \left\{ L - \sum_{r=1}^{v-1} f_r, \frac{d - \sum_{r=1}^{v-1} r f_r}{v} \right\}, \qquad 1 \leq v \leq w. \tag{3.7}$$

The probability of a fading block pattern $p(\mathbf{f})$ is computed using combinatorics as

$$p(\mathbf{f}) = \frac{\binom{m}{1}^{f_1} \binom{m}{2}^{f_2} \cdots \binom{m}{w}^{f_w}}{\binom{mF}{d}} . \frac{F!}{f_0! f_1! \dots f_w!}. \tag{3.8}$$

The left factor of $p(\mathbf{f})$ in (3.8) is the probability of distributing d nonzero bits over F blocks with f_v blocks having v bits, for possible values of v. The right term of $p(\mathbf{f})$ is the probability of having such combinations $\mathbf{f} = \{f_v\}_{v=0}^{w}$ among the F fading blocks. Using (3.6)-(3.8), the union bounds on the bit error probabilities of convolutional and turbo codes over a block fading channels are found by substituting (3.6) in (3.1) and (3.2), respectively. Also, substituting (3.6) in (3.4) results in the union bound on the frame error probability of turbo codes over block fading channels.

The number of summations involved in computing $P_u(d)$ in (3.6) increases as the channel memory length increases. Computing the bound by summing all values of $d \leq N$ for a large channel memory length becomes a time consum-

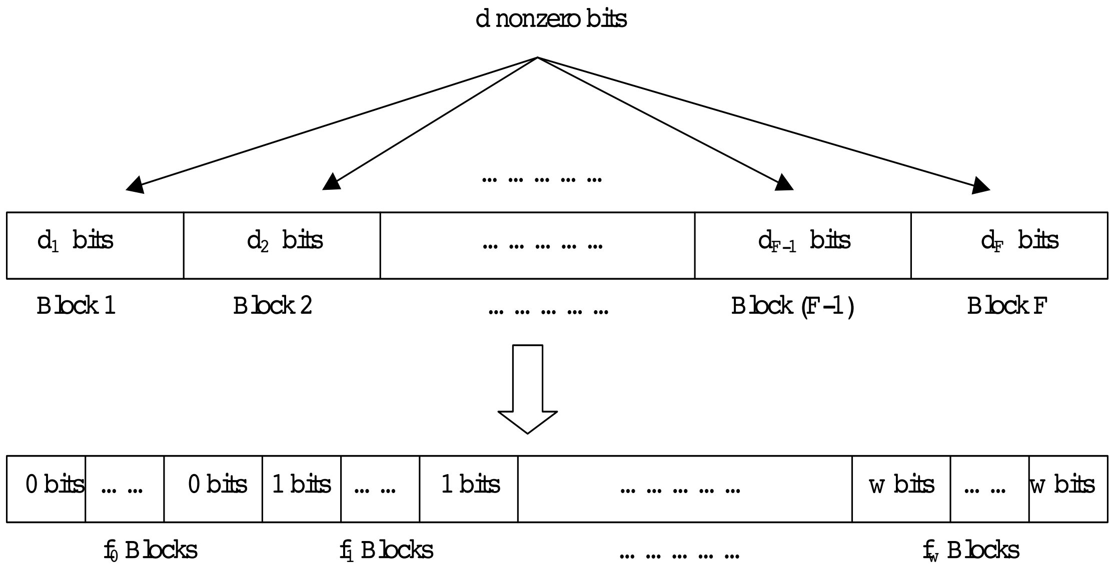

Figure 3.1: The distribution of the d nonzero bits in a d-weight error codeword over the F fading blocks.

ing task. However, a good approximation to the union bound is obtained by truncating it for a small value of $d < N$. This results in an approximation to the error probability rather than an upper bound.

3.1.2 Pairwise Error Probability

The conditional pairwise error probability $P_c(d|\mathbf{f})$ is defined as the probability of decoding a received sequence $\mathbf{Y}$ as a weight-d codeword $\hat{\mathbf{S}}$ given that the all-zero codeword $\mathbf{S}$ was transmitted and conditioned on the channel fading gains. It is given by

$$P_c(d|\mathbf{f}) = \Pr\Big(\mathbf{m}(\mathbf{Y}, \mathbf{S}) - \mathbf{m}(\mathbf{Y}, \hat{\mathbf{S}}) < 0 \Big| \mathbf{H}, \mathbf{S}\Big), \tag{3.9}$$

where $\mathbf{H} = \{h_f\}_{f=1}^{F}$. Note that the d nonzero errors are distributed over the F fading blocks according to a pattern $\mathbf{f}$. For a specific receiver, the unconditional pairwise error probability $P_u(d|\mathbf{f})$ is found by substituting the corresponding decoding metric (3.9) and then averaging over the fading gains. The rest of this chapter is devoted to deriving expressions of the unconditional pairwise error probability for coded single and multiple transmit antennas systems employing different receivers with different SI assumptions at the receiver.

3.2 Single-Antenna Systems

In this section, the pairwise error probability $P_u(d|\mathbf{f})$ is derived for coded single-antenna systems employing coherent and noncoherent receivers. For coherent detection, we consider the cases of perfect and imperfect SI at the receiver as well as the case of no amplitude SI. For noncoherent detection, receivers employing a square-law combining are considered.

3.2.1 Coherent Detection - Perfect SI

Recall that the received signal over a block fading channel is given by (2.1) and the corresponding ML decoding rule is given by (2.2). Substituting the metric (2.2) in (3.9), the conditional pairwise error probability for coherent detection with perfect SI is given by

$$P_c(d|\mathbf{f}) = \Pr\left(\sum_{f=1}^{L} a_f \sum_{l=1}^{m} \mathrm{Re}\{y_{f,l}\} < 0 \middle| \mathbf{H}, \mathbf{S}\right). \tag{3.10}$$

The distribution of $\mathrm{Re}\{y_{f,l}\}$ conditioned on a_f is Gaussian with mean $\sqrt{E_s} a_f s_{f,l}$ and variance N_0. The conditional pairwise error probability simplifies to

$$P_c(d|\mathbf{f}) = \mathrm{Q}\left(\sqrt{2R_c\gamma_b \sum_{v=1}^{w} v \sum_{i=1}^{f_v} a_i^2}\right), \tag{3.11}$$

where $\gamma_b = \frac{E_b}{N_0}$ is the SNR per information bit. Note that the average energy per bit is given by $E_b = R_c E_s$, where R_c is the encoder rate. To find $P_u(d|\mathbf{f})$, (3.11) is averaged over the fading amplitudes as

$$P_u(d|\mathbf{f}) = \mathrm{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\sqrt{2R_c\gamma_b\sum_{v=1}^{w} v\sum_{i=1}^{f_v} a_i^2}\right)\right], \tag{3.12}$$

where $\mathbf{A} = \{a_f\}_{f=1}^{F}$. Using the Chernoff bound, $\mathrm{Q}(x) \leq \frac{1}{2}e^{-x^2/2}$, the unconditional pairwise error probability is upper bounded as

$$P_u(d|\mathbf{f}) \leq \frac{1}{2}\prod_{v=1}^{w}\left(\frac{1}{1+vR_c\gamma_b}\right)^{f_v}, \tag{3.13}$$

where the product results from the independence of fading in different fading blocks. An exact expression of the pairwise error probability can be found by using the integral expression of the Q-function, $\mathrm{Q}(x) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} e^{(-x^2/2\sin^2\theta)}d\theta$ [45]

$$\begin{aligned} P_u(d|\mathbf{f}) &= \frac{1}{\pi}\mathrm{E}_{\mathbf{A}}\left[\int_0^{\frac{\pi}{2}} \exp\left(\frac{R_c\gamma_b}{\sin^2\theta}\sum_{v=1}^{w} v\sum_{i=1}^{f_v} a_i^2\right) d\theta\right] \\ &= \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{v=1}^{w}\left(\frac{1}{1+vR_c\gamma_b/\sin^2\theta}\right)^{f_v} d\theta. \end{aligned} \tag{3.14}$$

The union bound was evaluated for a rate-$\frac{1}{2}$ (23,35) convolutional code with a frame size of $N = 2 \times 512$ coded bits, and a rate-$\frac{1}{3}$ (1,5/7,5/7) turbo code with a frame size $N = 3 \times 1024$ coded bits. Note that the convolutional code has 4 memory elements, whereas the component codes of the turbo have 2 memory elements. As discussed in Section 3.1.1, the union bound is truncated to reduce computation complexity. For convolutional codes, the union bound was truncated after codewords with distances $d > 12$. However, the union bound for turbo codes was truncated after $d > 52$. The bound is compared to simulation results. In the simulations, the channel interleaver is chosen randomly and is changed every 10 frames to account for the uniform interleaving argument. In turbo codes, the code interleaver is an s-random interleaver. In the literature, s-random interleavers are used due to their good performance. An s-random interleaver permutes the data sequence in such a way that bits within distance of s from each other are separated by more than s positions in the interleaved sequence. In coherent systems, BPSK signaling is employed and (3.14) is used to compute the union bound.

Figure 3.2 shows the results for the convolutional code with perfect SI for channel memory ranging from $m = 1$ to $m = 64$. In the figure we plot the curves corresponding to simulation and union bound results. Since simulating very low error rates is too difficult, we plot simulation curves up to error rates of around $P_b = 10^{-6}$. However, the union bound curves are plotted for all SNR values. From the figure, the bound is very close to the simulation curves for a wide range of channel memory lengths. Also, the union bound starts to be loose as the SNR decreases, where the union bound is known to diverge at SNR values lower than the cutoff rate of the channel [46]. However, since simulation is easy for low error rates, the importance of the performance analysis is emphasized at high SNR values. In the following results, the union bound is plotted for high SNR values to make the presentation more clear.

The frame error probability of the turbo code is shown in Figure 3.3. We observe that the bound is able to predict the performance loss due to channel memory in the error floor region, i.e., at the high SNR. Also, the bound is not as close to simulation results as in the case of convolutional codes. This is mainly because the turbo code interleaver is an s-random interleaver, where the union bound for turbo codes in (3.4) assumes uniform code interleaver which averages the performance of bad and good code interleavers. However, using the bound provide some insight about the performance of turbo codes with channel memory. Note that the performance of turbo codes in the water-fall region is not predictable using the distance spectrum of the code. In this region iterative decoding is the major influence on the performance and results for ML decoding does not hold. Generally, the performance with perfect SI degrades as the channel memory increases.

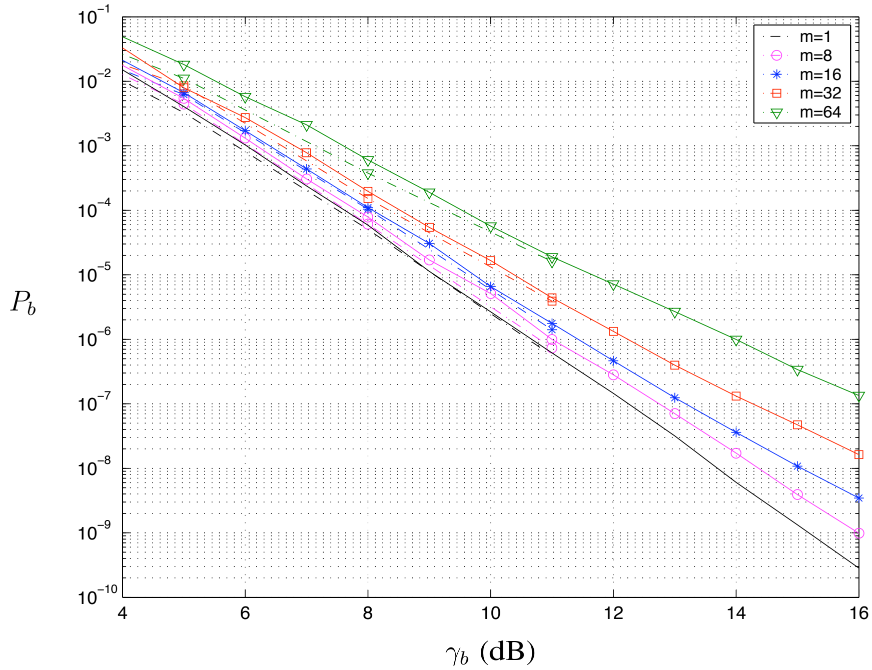

Figure 3.2: Bit error probability of a rate-$\frac{1}{2}$ (23,35) convolutional code with perfect SI and a frame size $N = 1024$ for channel memory lengths $m = 1, 8, 16, 32, 64$ (solid: approximation using the union bound, dash: simulation).

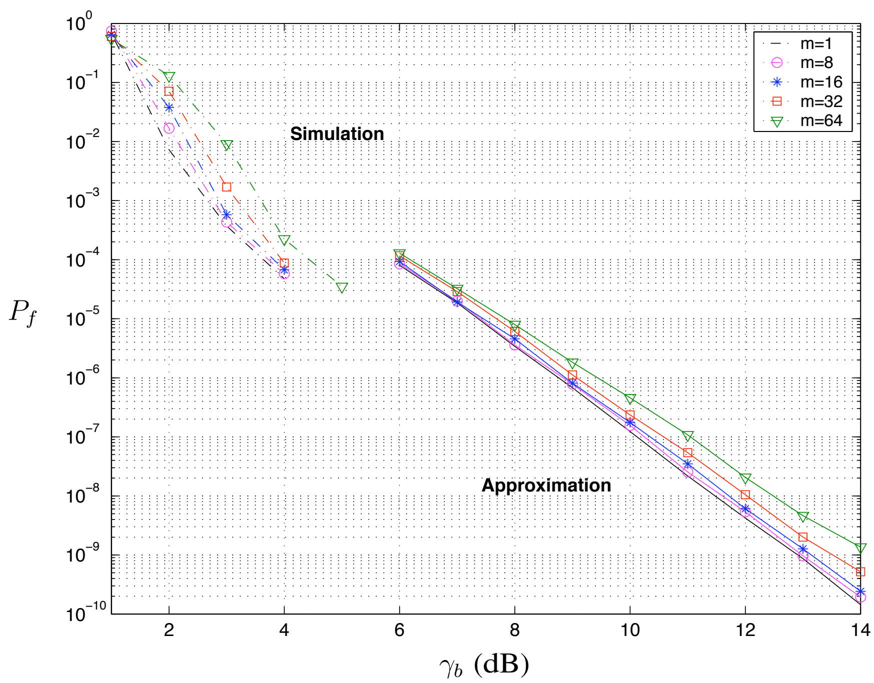

Figure 3.3: Frame error probability of a rate-$\frac{1}{3}$ (1,5/7,5/7) turbo code with perfect SI and an code interleaver size $\tilde{N} = 1024$ for channel memory lengths $m = 1, 8, 16, 32, 64$.

3.2.2 Coherent Detection - Imperfect SI

For coherent detection with imperfect SI it is necessary to estimate the channel SI. This is achieved by transmitting a pilot signal with energy E_p in each fading block. The corresponding received signal is given by

$$y_{f,p} = \sqrt{E_p} h_f + z_{f,p}. \tag{3.15}$$

The ML estimator for h_f is given by $\hat{h}_f = \frac{y_{f,p}}{\sqrt{E_p}} = h_f + e_f$, where $e_f = \frac{z_{f,p}}{\sqrt{E_p}}$ is the estimation error. The distribution of e_f is $\mathcal{CN}(0, \sigma_e^2)$ where $\sigma_e^2 = \frac{N_0}{E_p}$. The correlation coefficient between the actual channel gain and its estimate is given by

$$\mu = \frac{E[h_f \hat{h}_f^*]}{\sqrt{\text{Var}(h_f)\text{Var}(\hat{h}_f)}} = \frac{1}{\sqrt{1+\sigma_e^2}}. \tag{3.16}$$

In order implement a ML decoding rule, the likelihood function of the channel observations (received and pilot signals) conditioned on the transmitted codeword should be maximized. Define $\mathbf{y}_p = \{y_{f,p}\}_{f=1}^F$ to be the vector containing pilot signals in a frame, then the likelihood function is written as

$$p(\mathbf{Y}, \mathbf{y}_p|\mathbf{S}) = \text{E}_{\mathbf{H}}\left[p(\mathbf{Y}, \mathbf{y}_p|\mathbf{H}, \mathbf{S})\right] = C_1 \prod_{f=1}^{F} \int_0^\infty \exp\left(-\frac{1}{N_0}|y_{f,p} - \sqrt{E_p} h_f|^2\right) \tag{3.17}$$

$$\times \exp\left(-\frac{1}{N_0}\sum_{l=1}^{m} |y_{f,l} - \sqrt{E_s} h_f s_{f,l}|^2\right) e^{-|h_f|^2} dh_f, \tag{3.18}$$

where C_1 is a constant. Simplifying

$$p(\mathbf{Y}, \mathbf{y}_p|\mathbf{S}) = C_2 \prod_{f=1}^{F} \exp\Big(-\frac{1}{N_0}(|y_{f,p}|^2 + \sum_{l=1}^{m} |y_{f,l}|^2)\Big)$$

$$\times \int_0^\infty \exp\Big[-\frac{1}{N_0}\Big(\alpha|h_f|^2 - 2\text{Re}\{h_f^* \beta_f\}\Big)\Big] dh_f, \tag{3.19}$$

where $\alpha = E_p + mE_s + N_0$, $\beta_f = \sqrt{E_p} y_{f,p} + \sqrt{E_s}\sum_{l=1}^m y_{f,l} s_{f,l}^*$ and C_2 is another constant. Simplifying and completing the squares by adding and subtracting

$\frac{|\beta_f|^2}{\alpha N_0}$ **results in**

$$p(\mathbf{Y}, \mathbf{y}_p|\mathbf{S}) = C_3 \prod_{f=1}^{F} \exp\Big(-\frac{1}{N_0}(|y_{f,p}|^2 + \sum_{l=1}^{m} |y_{f,l}|^2)\Big) \exp\Big(\frac{1}{\alpha N_0}|\beta_f|^2\Big)$$
$$\times \int_0^{\infty} \exp\Big(-\frac{\alpha}{N_0}\Big|h_f - \frac{\beta_f}{\alpha}\Big|^2\Big) dh_f, \tag{3.20}$$

where C_3 is a constant. Now, the integral becomes a constant since it is an integral of a Gaussian density function. Note that the only part that depends on the transmitted codeword is the second exponential term. Thus based on observing the received and the pilot signals, the ML decoder chooses the codeword S that maximizes the metric

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = \sum_{f=1}^{F} |\beta_f|^2 = \sum_{f=1}^{F} E_s|\sum_{l=1}^{m} y_{f,l}s_{f,l}^*|^2 + 2\sqrt{E_s}\sum_{l=1}^{m} \mathrm{Re}\{y_{f,l}^* s_{f,l}\hat{h}_f\} + C_4, \tag{3.21}$$

where $\hat{h}_f$ have been substituted for $\sqrt{E_p}y_{f,p}$ and C_4 is another constant term that is independent of the transmitted codeword. This receiver is difficult to implement in a Viterbi decoder of a convolutional code. Also, it is difficult to analyze. Therefore, a suboptimal decoding metric that maximizes the likelihood function $p(\mathbf{Y}|\hat{\mathbf{H}}, \mathbf{S})$ is used. It is given by

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = \sum_{f=1}^{F}\sum_{l=1}^{m} \mathrm{Re}\{y_{f,l}^* \hat{h}_f s_{f,l}\}. \tag{3.22}$$

In order to find the pairwise error probability, the distribution of the received signal $y_{f,l}$ conditioned on the estimated channel gain $\hat{h}_f$ is required. Applying basic probability results, this distribution is found to be a complex Gaussian with mean $\frac{\mu}{\sigma}\sqrt{E_s}\hat{h}_f s_{f,l}$ and variance $N_0 + (1-\mu^2)E_s$, where $\sigma^2 = \mathrm{Var}(\hat{h}_f) =$

Table 3.1: Rates, minimum distances and puncturing patterns of the punctured rate-$\frac{1}{2}$ codes.

m	**Code Rate** $\tilde{R}_c$	**Puncturing Location**	$d_{\min}$
4	0.667	3	4
8	0.571	7	5
16	0.533	15	6
32	0.516	31	6
64	0.508	63	6

$1 + \sigma_e^2$. Thus the conditional pairwise error probability becomes

$$P_c(d|\mathbf{f}) = \mathrm{Q}\left(\sqrt{\frac{2\mu^2 R_c \gamma_b \sum_{v=1}^{w} v \sum_{i=1}^{f_v} |\zeta_i|^2}{1 + R_c\gamma_b(1-\mu^2)}}\right), \tag{3.23}$$

where $\zeta = \frac{\hat{h}}{\sigma}$. Define $\hat{\gamma}_b = \frac{\mu^2\gamma_b}{1+R_c\gamma_b(1-\mu^2)}$ to be the effective SNR after taking into account the additional noise in the channel estimation, the unconditional pairwise error probability simplifies to (3.14) with γ_b replaced by $\hat{\gamma}_b$.

Two scenarios can be considered for channel estimation using pilots with $E_p = E_s$. The first one results from only pilot estimation (OPE) with an estimation error variance of $\sigma_e^2 = \frac{N_0}{E_s}$. The second case considers a lower bound on the performance of receivers employing iterative joint decoding and channel estimation. In such receivers the decoding results are used to improve the channel estimates, which are used to improve the decoding results. This process is repeated iteratively. In general, the more reliable the decoding results, the more accurate is channel estimation. A lower bound on the performance of iterative receivers is obtained if the signals in each fading block are known with probability one. In this case they can be considered as pilots resulting in an estimation error variance of $\sigma_e^2 = \frac{N_0}{mE_s}$. This case is referred to as correct data estimation (CDE). Similar channel estimation scenarios were used in [47–49] for channel estimation for LDPC codes.

In simulating systems with channel estimation via pilot insertion, one coded bit is punctured every m coded bits to account for the rate reduction resulting from inserting a pilot signal every $m-1$ signals. This affects the whole distance distribution of the resulting code and may reduce the minimum distance of the code. In general, the resultant code rate after puncturing is given by

$$\tilde{R}_c = \frac{mR_c}{m - n_p}, \tag{3.24}$$

where n_p is the number of pilot signals inserted in each fading block which is set to be $n_p = 1$ in single-antenna systems. In Table 3.1 we show the code rates and the minimum distances of the punctured codes for different channel memory lengths. Also, the location of the punctured coded bit in a m-length fading block. According to the table, the code rate increases with reducing the channel memory length, which decreases the error correcting capabilities of the code. Thus systems with short channel memory are expected to have more channel diversity at the cost of lower minimum distance and worse channel estimation quality. On the other hand, longer channel memory results in more a powerful code as well as better channel estimation at the cost of less channel diversity.

Figures 3.4 and 3.5 show the results for convolutional code with imperfect SI under the OPE and CDE assumptions, respectively. Note that the energy of the pilot is taken into account in the SNR axis. We observe that the SNR degradation due to channel memory is less compared to the case of perfect SI. This is expected since the number of transmitted pilot signals is reduced as the channel memory gets longer, and hence the system becomes more energy efficient. Also, the estimation quality under the CDE assumption improves with increasing channel memory length as appears in the expression of σ_e^2. From the figures, the cases of $m = 16$ and $m = 32$ are the best systems, where the former becomes better than the later for an SNR values exceeding 14 dB. This suggests that the optimal channel memory length is between $m = 16$ and

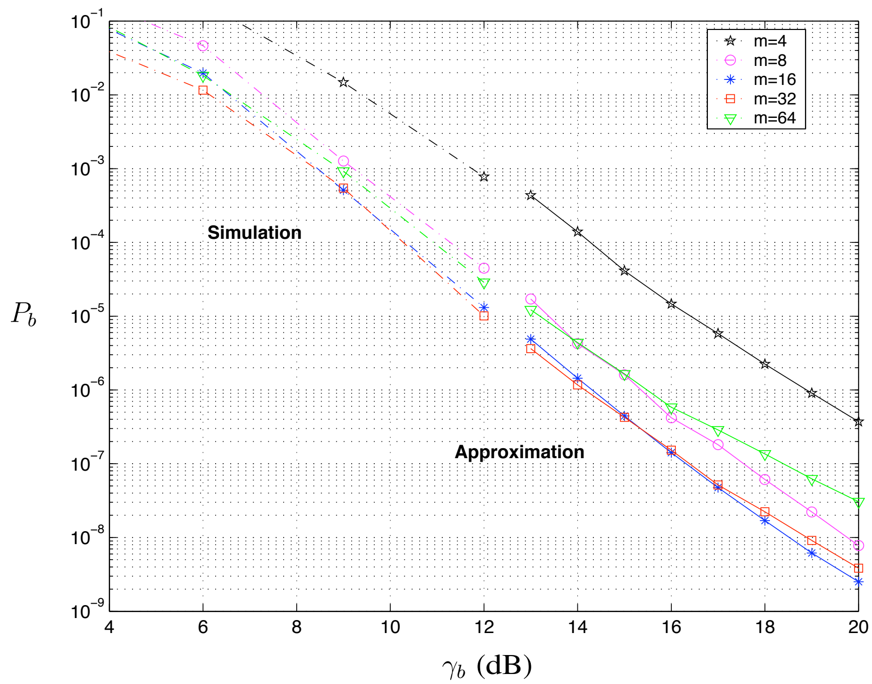

Figure 3.4: Bit error probability of a (23,35) convolutional code with imperfect SI (OPE receiver with $E_p = E_s$) and a frame size $N = 1024$ for channel memory lengths $m = 4, 8, 16, 32, 64$.

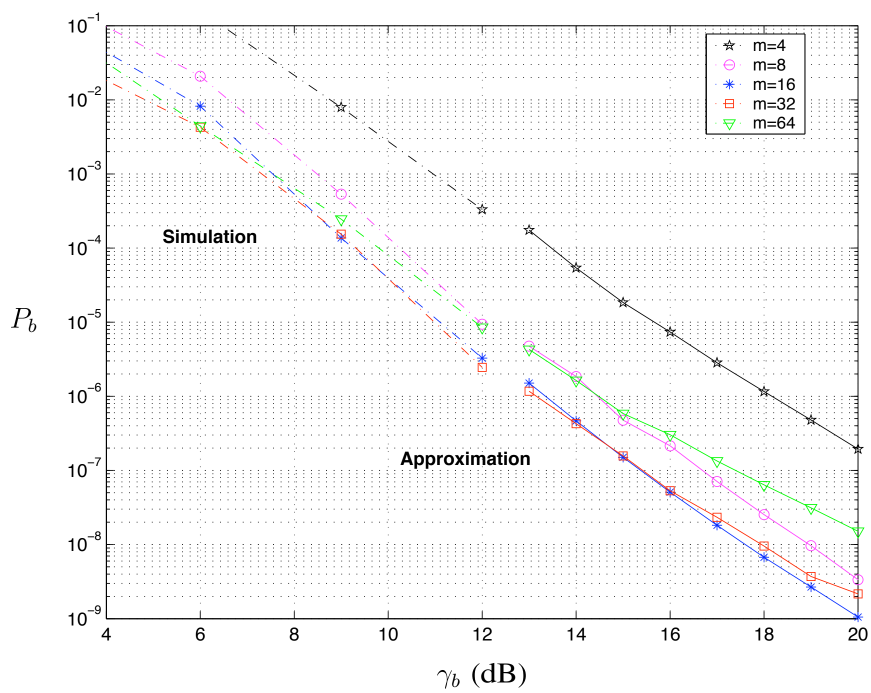

Figure 3.5: Bit error probability of a (23,35) convolutional code with imperfect SI (CDE assumption with $E_p = E_s$) and a frame size $N = 1024$ for channel memory lengths $m = 4, 8, 16, 32, 64$.

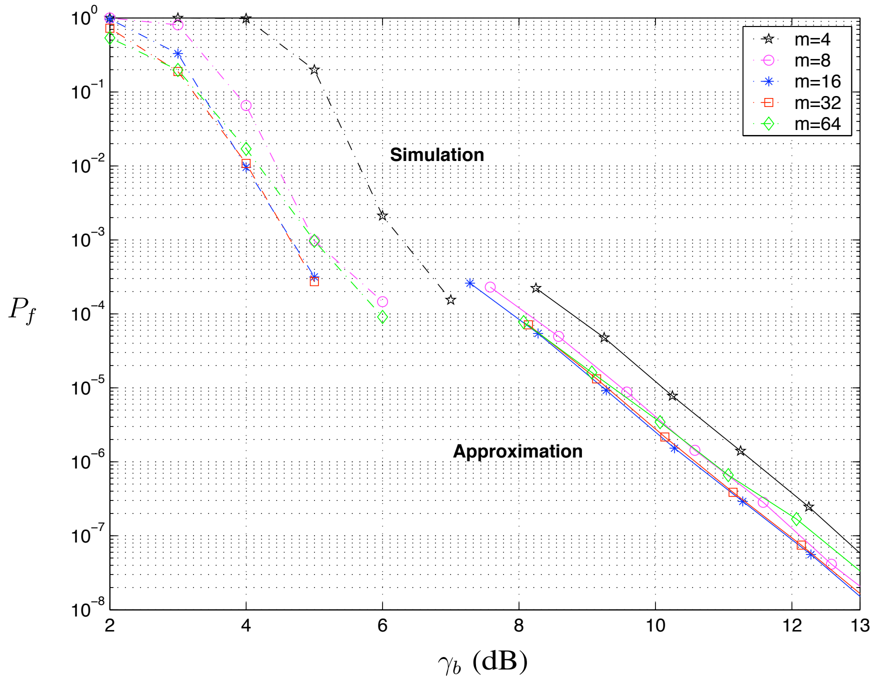

Figure 3.6: Frame error probability of a (1,5/7,5/7) turbo code with imperfect SI (OPE receiver with $E_p = E_s$) and an code interleaver size $\tilde{N} = 1024$ for channel memory lengths $m = 4, 8, 32, 64$.

$m = 32$. Also, the cases of $m = 8$ is the worse than the case of $m = 64$ at low SNR and starts to improve as the SNR increases. This is because the resulting code for $m = 8$ is less powerful than the code for $m = 64$ but has larger amount of channel diversity. Although the case of $m = 64$ has the best code, it lacks enough channel diversity to perform better than the other cases. Under the CDE assumption, the case of $m = 64$ performs the best at low SNR because longer memory permits better estimation. In general, the optimal memory tends to increase under the CDE assumption compared to the OPE receiver due to the improved channel estimation. In all cases the channel memory $m = 4$ performs the worst because the resulting code is weak due to puncturing one coded bit every 4 coded bits.

Results for turbo code with imperfect SI using an OPE receiver and the CDE assumption are shown in Figures 3.6 and 3.7, respectively. In OPE receivers, we observe that the curves of the cases of $m = 8, 16, 32$ cross at different SNR points. As the SNR increases, short channel memory provides better perfor-

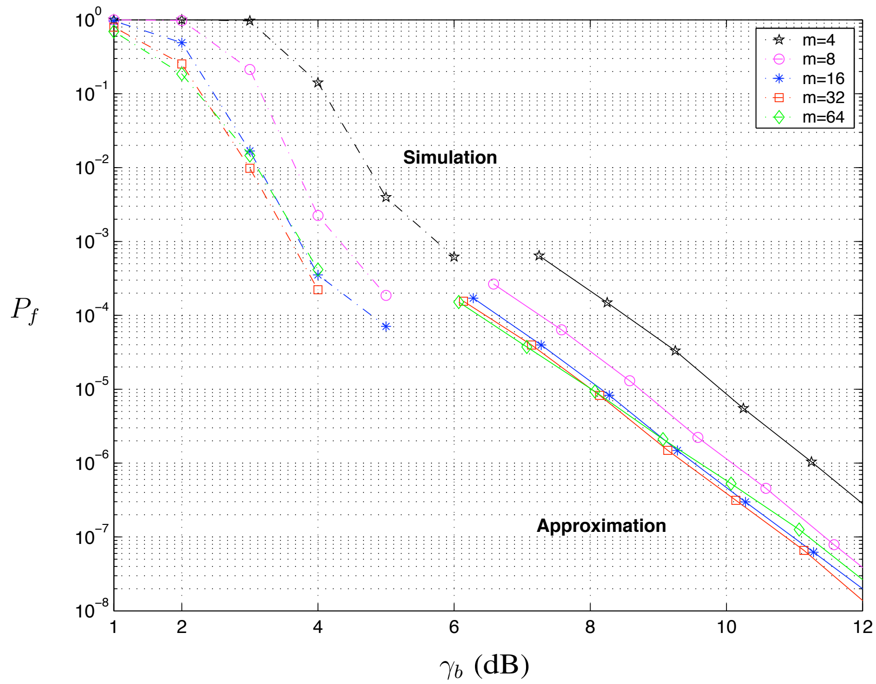

Figure 3.7: Frame error probability of for a (1,5/7,5/7) turbo code with imperfect SI (CDE assumption with $E_p = E_s$) and an code interleaver size $\tilde{N} = 1024$ for channel memory lengths $m = 4, 8, 32, 64$.

mance than long channel memory, which is apparent in the analytical curves. From this we conclude that the optimal memory length changes with the operating SNR value. In the CDE assumption, the channel memory lengths $m = 32$ and $m = 64$ are very close where the former becomes the best at high SNR. Figure 3.8 shows a comparison between the performance of the convolutional code with channel memory lengths $m = 8$ and $m = 32$ for the cases of perfect SI, OPE receiver and the CDE assumption. It is clear that as the channel memory gets longer, the SNR degradation due to imperfection in the channel SI reduces. This is because long channel memory causes less penalty in the rate and energy than short channel memory does, as well as an improved channel SI under the CDE assumption.

When the energy allocated for the pilot signal is varied, the performance of an OPE receiver is expected to change as a function of the channel memory. The energy per information bit is written as a function of the energy allocated

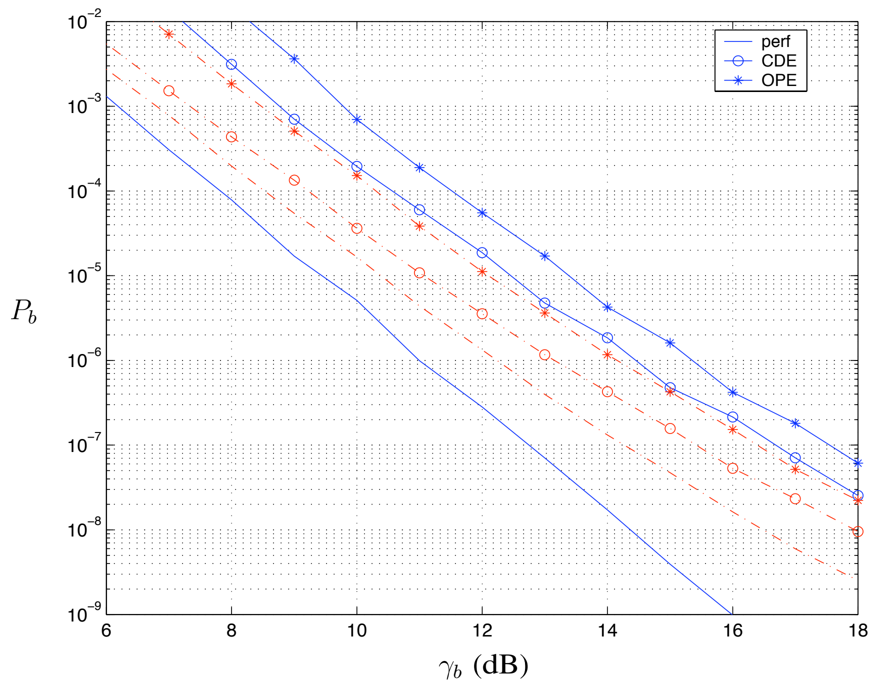

Figure 3.8: The performance of convolutional coded systems with a frame size $N = 1024$ and channel memory lengths $m = 8, 32$ using perfect and imperfect SI with $E_p = E_s$ (solid: $m = 8$, dash: $m = 32$).

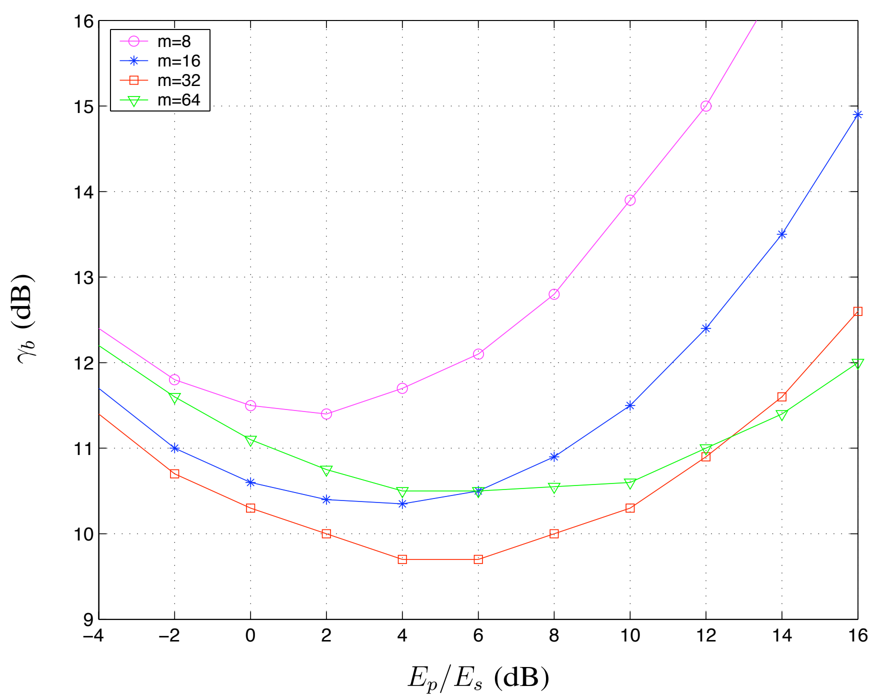

Figure 3.9: SNR required for the (23,35) convolutional code to achieve $P_b = 10^{-4}$ versus E_p/E_s for the OPE receiver and channel memory lengths $m = 8, 16, 32, 64$.

for transmitting signals and pilots as

$$E_b = \frac{(m - n_p)E_s + n_p E_p}{mR_c}. \tag{3.25}$$

Thus for a fixed channel memory, there exists an optimal value for pilot energy. This is illustrated in Figure 3.9, where the SNR required for the system to achieve bit error probability of $P_b = 10^{-4}$ is plotted versus the pilot-to-signal energy ratio E_p/E_s in dB. We observe that as the channel memory length increases the optimum value for E_p/E_s increases. This is expected since longer channel memory permits more possible energy to be allocated for the pilot signal. On the other hand, when the channel memory is short, a more wise usage of the offered energy is to transmit the information signals rather than to estimate the channel. Note that optimizing the pilot energy results in an SNR gain as large as 1 dB over the case where $E_p = E_s$. This SNR gain increases as the channel memory increases since longer memory increases the amount of energy that can be devoted for channel estimation, which improves the overall performance.

3.2.3 Coherent Detection - No Amplitude SI

For completeness we consider the case where the channel phase is known but the amplitude is unknown. A suboptimal decoding metric (2.3) is used due to its mathematical tractability. Substituting this metric (3.9) and using the Chernoff bound to upper bound the pairwise error probability [12] as

$$P_c(d|\mathbf{f}) \leq \prod_{f=1}^{L} \mathrm{E}_y \Big[\exp\Big(-2\lambda \sum_{l=1}^{m} \mathrm{Re}\{y_{f,l}\} |s_{f,l} - \hat{s}_{f,l}| \Big) \Big], \tag{3.26}$$

where $\lambda > 0$ is the Chernoff parameter. Following the derivation in Appendix A.1, the pairwise error probability with coherent detection and no amplitude

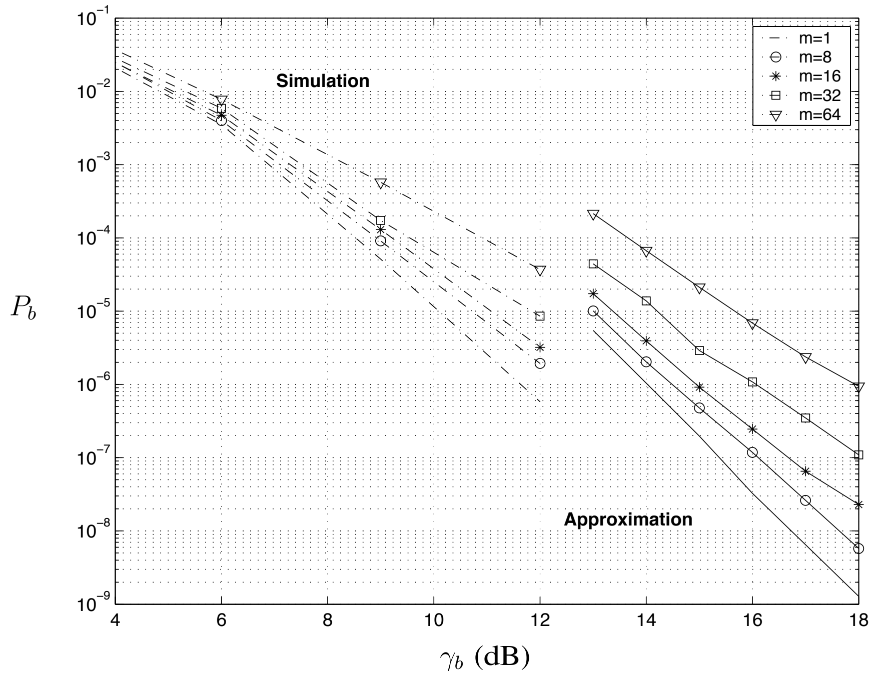

Figure 3.10: Bit error probability of a rate-$\frac{1}{2}$ (23,35) convolutional code with no amplitude SI and a frame size $N = 1024$ for channel memory lengths $m = 1, 8, 16, 32, 64$.

SI is given by

$$P_u(d|\mathbf{f}) \lesssim \left(\frac{ed}{2R_c\gamma_b L}\right)^L \left(\prod_{v=1}^{w} v^{2f_v}\right)^{-1}, \tag{3.27}$$

where $e = \exp(1)$. The union bound is evaluated for the case of coherent detection with no amplitude SI using (3.27) and results are shown in Figure 3.10. From the figure, we observe that the bound is less tight than in the perfect SI case. This is due to the use of the Chernoff bound which is known to be a little loose in low SNR region. However, it dictates the trend in the performance over block fading channels, where the SNR degradation due to memory is clear and close to that shown in simulations.

3.2.4 Noncoherent Detection

In noncoherent systems the channel phase is unknown at the receiver. A receiver that does not need any channel estimation (phase and amplitude) is the square-law combining receiver. Recall that the outputs of the square-law combiner are given by (2.5) and the corresponding decoding metric appears in (2.6). The conditional pairwise error probability is found by substituting the metric (2.6) in (3.9)

$$P_c(d|\mathbf{f}) = \Pr\left(\sum_{f=1}^{F} d_f \left(|r_f^{(I,1)}|^2 + |r_f^{(Q,1)}|^2 - |r_f^{(I,0)}|^2 - |r_f^{(Q,0)}|^2 \right) > 0 \Big| \mathbf{H}, \mathbf{S} \right), \tag{3.28}$$

where d_f is the number of error bits in the fading block f. The variables $\{r_f^{(I,0)}, r_f^{(Q,0)}\}$ and $\{r_f^{(I,1)}, r_f^{(Q,1)}\}$ are zero mean Gaussian random variables with variances equal to $\frac{1}{2}(E_s + N_0)$ and $\frac{1}{2}N_0$, respectively. Let $|r_f^{(s)}|^2 = |r_f^{(I,s)}|^2 + |r_f^{(Q,s)}|^2$ for $s = 0, 1$ and define $\kappa = \sum_{f=1}^{F} d_f(|r_f^{(1)}|^2 - |r_f^{(0)}|^2)$. Then, the unconditional pairwise error probability is upper bounded using the Chernoff bound and the density function $p(\kappa)$ as

$$P_u(d|\mathbf{f}) = \int_0^{\infty} p(\kappa) d\kappa \leq \mathrm{E}_\kappa \left[e^{\lambda\kappa} \right], \tag{3.29}$$

where $\lambda > 0$ is the Chernoff parameter that should be optimized to result in the tightest bound. Substituting for κ and collecting terms having the same distance

$$P_u(d|\mathbf{f}) \leq \prod_{v=1}^{w} \mathrm{E}\left[e^{\lambda v |r_v^{(1)}|^2} \right]^{f_v} \mathrm{E}\left[e^{-\lambda v |r_v^{(0)}|^2} \right]^{f_v}. \tag{3.30}$$

The Chernoff parameter λ is optimized as in [35] and the resulting Chernoff bound for the pairwise error probability simplifies to

$$P_u(d|\mathbf{f}) \leq \prod_{v=1}^{w} \left[4D_v(1 - D_v) \right]^{f_v}, \tag{3.31}$$

where $D_v = \frac{1}{2 + vR_c\gamma_b}$. The union bound for convolutionally encoded BFSK signals with square-law combining is evaluated using (3.31) and shown in Figure 3.11. The bound is less tight than the perfect SI case due to the use of the Chernoff bounding technique. However, the performance trend and the SNR

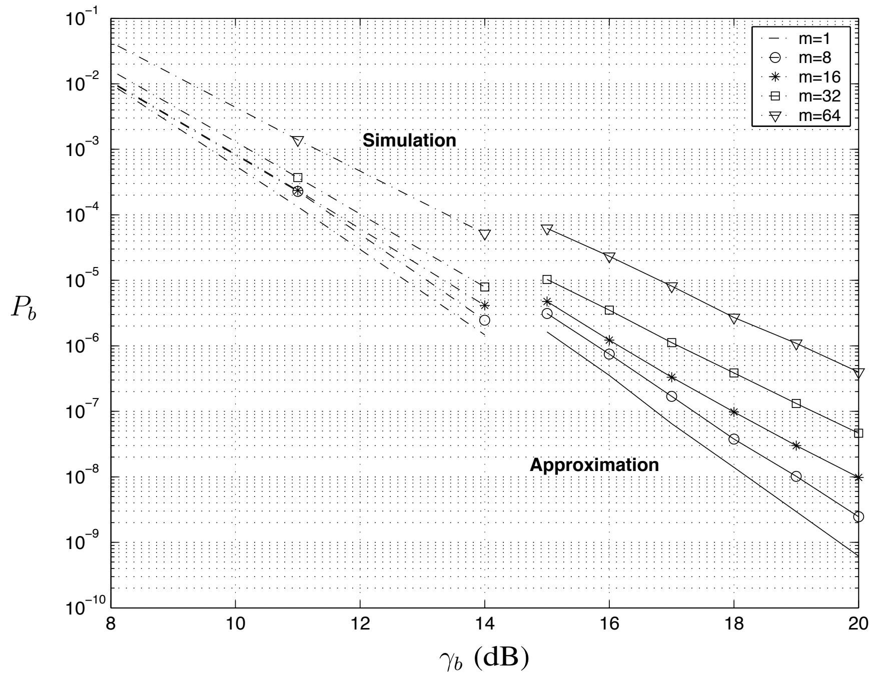

Figure 3.11: Bit error probability of a rate-$\frac{1}{2}$ (23,35) convolutional code with noncoherent detection and a frame size $N = 1024$ for channel memory lengths $m = 1, 8, 16, 32, 64$.

degradation due to memory are predicted well from the bound.

3.3 Multi-Antenna Systems

In this section, the pairwise error probabilities for coded STBC systems are derived. Coherent detection with perfect and imperfect SI is considered. Note that the results for the case of perfect SI apply directly to coded differential STBCs taking into account a penalty of 3 dB in the SNR [22].

3.3.1 Perfect SI

Recall that the received vector due to transmitting a STBC transmission matrix $\mathcal{G}_{f,l}$ is given by (2.8). Also, the simple decoding metric that is equivalent to ML metric is given by (2.9). The conditional pairwise error probability for coded STBCs with perfect SI is found by substituting the metric (2.9) in the expression of the pairwise error probability in (3.9) resulting in

$$P_c(d|\mathbf{f}) = \Pr\left(\sum_{f=1}^{L}\sum_{l=1}^{w}\kappa_{f,l} < 0\middle|\mathbf{H},\mathbf{S}\right), \tag{3.32}$$

where $\kappa_{f,l} = \mathrm{Re}\{\mathbf{y}^*_{f,l}\mathcal{E}_{f,l}\mathbf{h}_f\}$ and $\mathcal{E}_{f,l} = \mathcal{G}_{f,l} - \hat{\mathcal{G}}_{f,l}$. Here, $\mathcal{G}_{f,l}$ and $\hat{\mathcal{G}}_{f,l}$ are the transmission matrices in time slot l of the f^{th} fading block corresponding to the all-zero codeword and a weight-d error codeword, respectively. We use $(.)^*$ and $(.)^T$ to denote the complex conjugate of a complex vector and the transpose of a real matrix, respectively. In (3.32), $\kappa_{f,l}$ is a Gaussian random variable with conditional mean and variance given respectively by

$$\mathrm{E}\left[\kappa_{f,l}\middle|\mathcal{G}_{f,l},\mathbf{h}_f\right] = \sqrt{E_s}\mathrm{Re}\left\{\mathbf{h}^*_f\mathcal{E}^T_{f,l}\mathcal{E}_{f,l}\mathbf{h}_f\right\} = \sqrt{E_s}d_{f,l}\sum_{i=1}^{n_t}|h^i_f|^2, \tag{3.33}$$

$$\mathrm{Var}\left[\kappa_{f,l}\middle|\mathcal{G}_{f,l},\mathbf{h}_f\right] = \mathrm{E}\left[\mathrm{Re}\left\{\mathbf{z}^*_{f,l}\mathcal{E}_{f,l}\mathbf{h}_f\mathbf{h}^*_f\mathcal{E}^T_{f,l}\mathbf{z}_{f,l}\right\}\middle|\mathcal{G}_{f,l},\mathbf{h}_f\right] = d_{f,l}N_0\sum_{i=1}^{n_t}|h^i_f|^2, \tag{3.34}$$

where $d_{f,l}$ is the number of error bits in the time slot l in the f^{th} fading block. In (3.33) and (3.34), the cross terms are zero due to the orthogonality of the rows in $\mathcal{E}_{f,l}$. The error probability in (3.32) simplifies to

$$P_c(d|\mathbf{f}) = \mathrm{Q}\left(\sqrt{2R_c\gamma_b\sum_{v=1}^{w}v\sum_{l=1}^{f_v}\sum_{i=1}^{n_t}|h^i_l|^2}\right). \tag{3.35}$$

As in Section 3.2.1, the unconditional pairwise error probability $P_u(d|\mathbf{f})$ is found by averaging over the fading gains and using the exact expression of the Q function. The resulting expression of $P_u(d|\mathbf{f})$ is given by

$$P_u(d|\mathbf{f}) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{v=1}^{w}\left(\frac{1}{1+vR_c\gamma_b/\sin^2\theta}\right)^{n_t f_v}d\theta. \tag{3.36}$$

As study cases, a rate-$\frac{1}{2}$ (23,35) convolutional code with 4 memory elements is concatenated with a STBC employing two and four transmit antennas with the

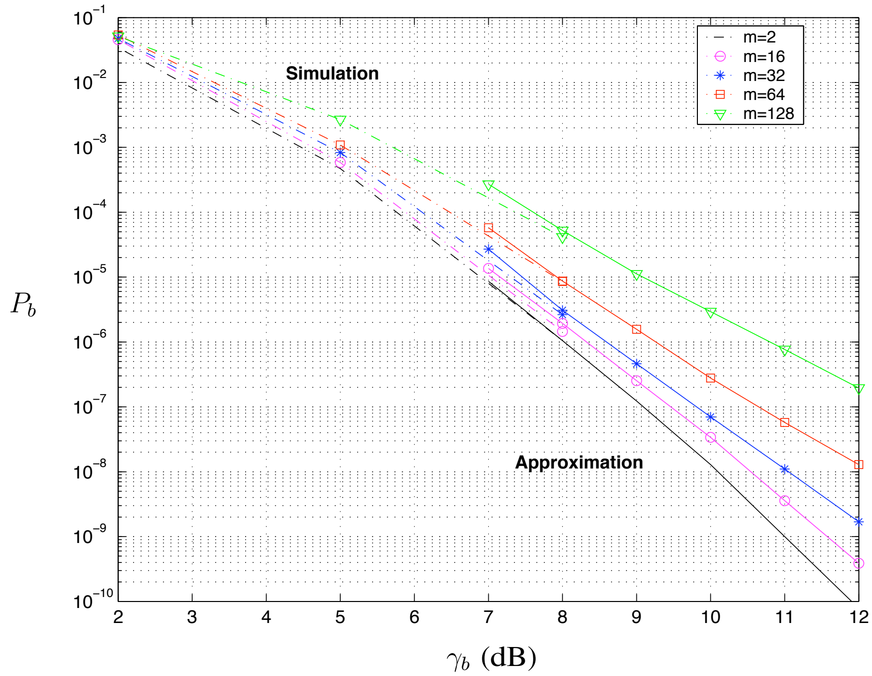

Figure 3.12: Bit error probability of a convolutionally coded STBC with perfect SI, $n_t = 2$ and a frame size $N = 1024$ for channel memory lengths $m = 2, 16, 32, 64, 128$.

same simulation settings as in the single-antenna case. The union bound was evaluated by substituting (3.36) in (3.1) and summing over codewords with distances $d \leq 12$ for fast computation of the bound. The results for the cases of two and four transmit antennas are shown in Figures 3.12 and 3.13, respectively for different channel memory lengths. From the figures, we observe that the performance degradation due to increasing the memory length reduces as the number of transmit antennas is increased. This is expected since there is more space diversity as the number of transmit antennas increases, which reduces the sensitivity of the performance to the number of independent fading blocks. This is clear in Figure 3.14, where the SNR required to achieve $P_b = 10^{-4}$ is plotted versus the number of transmit antennas for different channel memory lengths. We see that as the number of transmit antennas increases, the SNR loss due to long channel memory reduces.

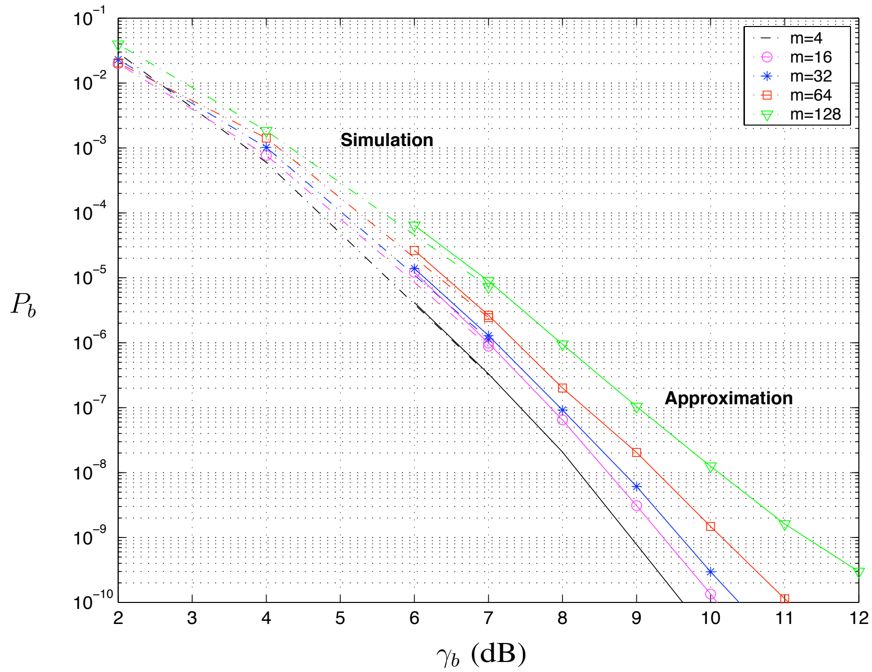

Figure 3.13: Bit error probability of a convolutionally coded STBC with perfect SI, $n_t = 4$ and a frame size $N = 1024$ for channel memory lengths $m = 4, 16, 32, 64, 128$.

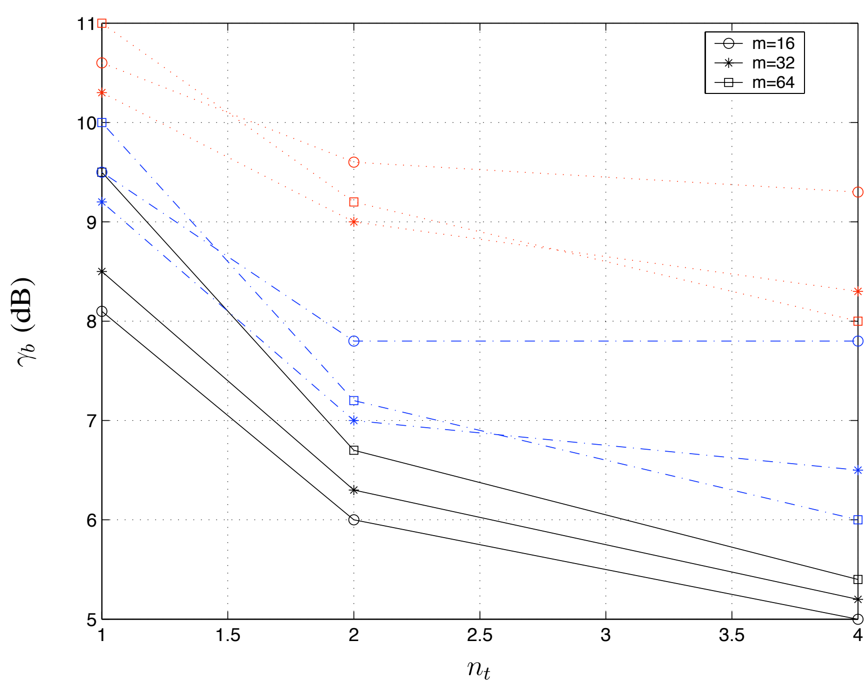

Figure 3.14: SNR required for a convolutionally coded STBC to achieve $P_b = 10^{-4}$ versus the number of transmit antennas n_t for channel memory lengths $m = 16, 32, 64$ (solid: perfect SI, dash: CDE, dots: OPE).

3.3.2 Imperfect SI

Now we consider the case of coherent receivers with channel SI generated from pilot sequences. This is achieved by transmitting n_t orthogonal pilot sequences [23], each of length $n_p \geq n_t$, over the n_t antennas in each fading block. Thus a known pilot sequence $\mathbf{p}^i = \{p_l^i\}_{l=1}^{n_p}$ is transmitted from the i^{th} transmit antenna once in each fading block. Denote by $\mathbf{y}_f^p = \{y_{f,l}^p\}_{l=1}^{n_p}$ the received column vector corresponding to the pilot sequences at each receive antenna in fading block f. It is given by

$$\mathbf{y}_f^p = \sqrt{E_p}\sum_{i=1}^{n_t} h_f^i \mathbf{p}^i + \mathbf{z}_f, \qquad 1 \leq f \leq F, \tag{3.37}$$

where E_p is the pilot energy. If pilot sequences from different transmit antennas are orthogonal, i.e., $\mathbf{p}^{i*}.\mathbf{p}^j = 0$ when $i \neq j$, then the ML estimator of the channel gain h_f^i is obtained by projecting $\mathbf{y}_f^p$ on $\mathbf{p}^i$ as

$$\mathbf{y}_f^p.\mathbf{p}^{i*} = h_f^i(\mathbf{p}^i.\mathbf{p}^{i*}) + \mathbf{z}_f.\mathbf{p}^{i*}. \tag{3.38}$$

Thus the ML estimator of the channel gain h_f^i is given by

$$\hat{h}_f^i = \frac{\mathbf{y}_f^p.\mathbf{p}^{i*}}{\parallel \mathbf{p}^i \parallel^2} - \frac{\mathbf{z}_f.\mathbf{p}^{i*}}{\parallel \mathbf{p}^i \parallel^2} \;=\; h_f^i + e_f^i, \tag{3.39}$$

where $e_f^i = (\mathbf{z}_f.\mathbf{p}^{i*}/ \parallel \mathbf{p}^i \parallel^2)$ is the estimation error associated with the channel from the i^{th} transmit branch in fading block f. From (3.39), the distribution of e_f^i follows $\mathcal{CN}(0, \sigma_e^2)$ where $\sigma_e^2 = \frac{N_0}{n_p E_p}$. The correlation coefficient between the true and estimated channel gains is defined as

$$\mu = \frac{\mathrm{E}\left[h_f^i \hat{h}_f^{i*}\right]}{\sqrt{\mathrm{Var}(h_f^i)\mathrm{Var}(\hat{h}_f^i)}} = \frac{1}{\sqrt{1+\sigma_e^2}}, \tag{3.40}$$

where $\mathrm{Var}(\hat{h}_f^i) = \sigma^2 = 1 + \sigma_e^2$.

From Section 3.2.2 we conclude that the ML decoding metric is difficult to implement in a Viterbi-like decoder as well as being difficult to analyze. Hence a suboptimal decoding rule that maximizes the likelihood function $p(\mathbf{Y}|\hat{\mathbf{H}}, \mathbf{S})$ is employed. It chooses the codeword S that maximizes the metric

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = \sum_{f=1}^{F} \sum_{l=1}^{m/n_t} \mathrm{Re}\{\mathbf{y}_{f,l} G_{f,l} \hat{\mathbf{h}}_f\}. \tag{3.41}$$

The received signal vector $\mathbf{y}_{f,l}$ conditioned on the estimated channel gains is a complex Gaussian random vector with mean $\frac{\mu}{\sigma}\sqrt{E_s}\mathcal{G}_{f,l}\hat{\mathbf{h}}_f$ and covariance matrix $(N_0 + n_t E_s(1-\mu^2))\mathbf{I}$. Thus the pairwise error probability conditioned on the estimated fading gains is given by (3.32) with replacing $\mathbf{H}$ and $\kappa_{f,l}$ by $\hat{\mathbf{H}}$ and $\hat{\kappa}_{f,l} = \mathrm{Re}\{\mathbf{y}^*_{f,l}\mathcal{E}_{f,l}\hat{\mathbf{h}}_f\}$, respectively. Similar to the case of perfect SI, $\hat{\kappa}_{f,l}$ is a Gaussian random variable with mean and variance given respectively by

$$\mathrm{E}\left[\hat{\kappa}_{f,l}\middle|\mathcal{G}_{f,l}, \hat{\mathbf{h}}_f\right] = \frac{\mu}{\sigma}\sqrt{E_s}\mathrm{Re}\left\{\hat{\mathbf{h}}^*_{f,l}\mathcal{E}^T_{f,l}\mathcal{E}_{f,l}\hat{\mathbf{h}}_f\right\} = \frac{\mu}{\sigma}\sqrt{E_s} d_{f,l}\sum_{i=1}^{n_t}|\hat{h}^i_f|^2, \tag{3.42}$$

$$\mathrm{Var}\left[\hat{\kappa}_{f,l}\middle|\mathcal{G}_{f,l}, \hat{\mathbf{h}}_f\right] = \left(N_0 + n_t E_s(1-\mu^2)\right) d_{f,l}\sum_{i=1}^{n_t}|\hat{h}^i_f|^2. \tag{3.43}$$

Using the mean and variance of $\hat{\kappa}_{f,l}$, the pairwise error probability conditioned on the estimated fading gains is given by (3.35), with $\hat{\gamma}_b = \frac{\mu^2\gamma_b}{1+n_t R_c \gamma_b(1-\mu^2)}$ replacing γ_b. As in the single-antenna case, $\hat{\gamma}_b$ represents the effective SNR taking into account the additional noise in the channel estimation. Thus the unconditional error probability is found by averaging over the estimated fading gains, resulting in (3.36) with the SNR value being $\hat{\gamma}_b$. For the special case of $E_p = E_s$, the estimation error variance in an OPE receiver is $\sigma_e^2 = \frac{N_0}{n_t E_s}$, whereas it is $\sigma_e^2 = \frac{N_0}{m E_s}$ under the CDE assumption.

In multi-antenna systems with pilot-aided channel estimation, n_t coded bits are punctured every m coded bits and replaced by a pilot sequence of length n_t. This reduces the error correcting capability of the code as the channel memory length becomes shorter which degrades the performance. The code rate of the punctured codes is given by (3.24). Tables 3.2 and 3.3 show the code rates and

Table 3.2: Rates, minimum distances and puncturing patterns of the punctured rate-$\frac{1}{2}$ codes for multi-antenna systems with $n_t = 2$.

m	**Code Rate** $\tilde{R}_c$	**Puncturing Locations**	$d_{\min}$
8	0.667	3,7	5
16	0.571	7,15	5
32	0.533	15,31	6
64	0.516	31,63	6
128	0.508	63,127	6

Table 3.3: Rates, minimum distances and puncturing patterns of the punctured rate-$\frac{1}{2}$ codes for multi-antenna systems with $n_t = 4$.

m	**Code Rate** $\tilde{R}_c$	**Puncturing Locations**	$d_{\min}$
16	0.667	3,7,11,15	5
32	0.571	7,15,23,31	6
64	0.533	15,31,47,63	6
128	0.516	31,63,95,127	6

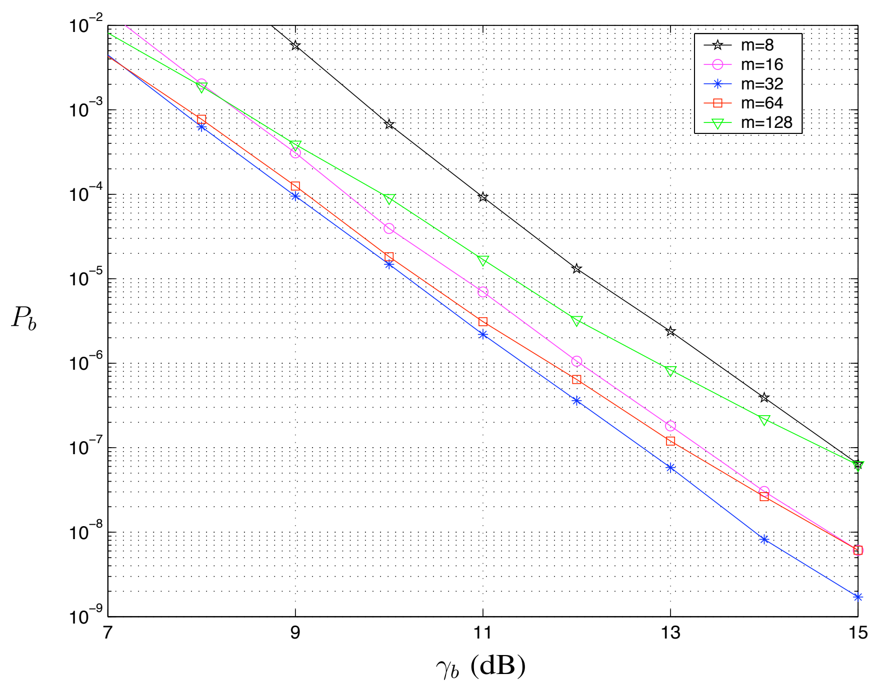

Figure 3.15: Approximation of bit error probability of a convolutionally coded STBC with imperfect SI (OPE receiver with $E_p = E_s$), $n_t = 2$ and a frame size $N = 1024$ for channel memory lengths $m = 8, 16, 32, 64, 128$.

minimum distance of the punctured codes used in systems with two and four transmit antennas, respectively with different channel memory lengths. Also the tables show the locations of the punctured coded bits in a length-m fading block. In Figure 3.14, we see that for single-antenna systems at $P_b = 10^{-4}$, systems with $m = 32$ performs the best, whereas in multi-antenna systems, the memory lengths $m = 32$ and $m = 64$ provide the best performance for the cases of $n_t = 2$ and $n_t = 4$, respectively. Moreover, we observe that as the channel memory length increases the gain of the CDE assumption over the OPE receiver increases. This is basically due to the enhanced channel estimation as the memory length increases. The results for two transmit antennas with imperfect SI using an OPE receiver and the CDE assumption are shown in Figures 3.15 and 3.16, respectively. Again, the energy of the pilot sequences is taken into account in the SNR axis. In all cases the memory length $m = 8$ performs the worst among the shown curves because the resulting code is weak due to puncturing two coded bits every 8 coded bits. Using an OPE receiver, the case of $m = 64$ outperforms all other cases in the low SNR region, whereas the case of $m = 32$ starts to improve and becomes the best after an SNR value of 7 dB. Also, observe that the case of $m = 16$ outperforms the $m = 128$ case, which is reversed at low SNR. The same phenomena are observed in the CDE assumption, where the case of $m = 64$ performs the best at low SNR and then degrades as the SNR increases.

Figures 3.17 and 3.18 show the results for the case of four transmit antennas with imperfect SI using an OPE receiver and the CDE assumption, respectively. The optimal channel memory seems to be between $m = 64$ and $m = 128$, with a cross over at around 7 dB. In general we conclude that the optimal memory tends to increase as the number of transmit antennas increases for the following reasons. First, as the number of transmit antennas increases more channels are needed to be estimated, which requires longer observation period for each channel. Second, more space diversity reduces the effect of diversity provided by the independent fading blocks in the channel making channel esti-

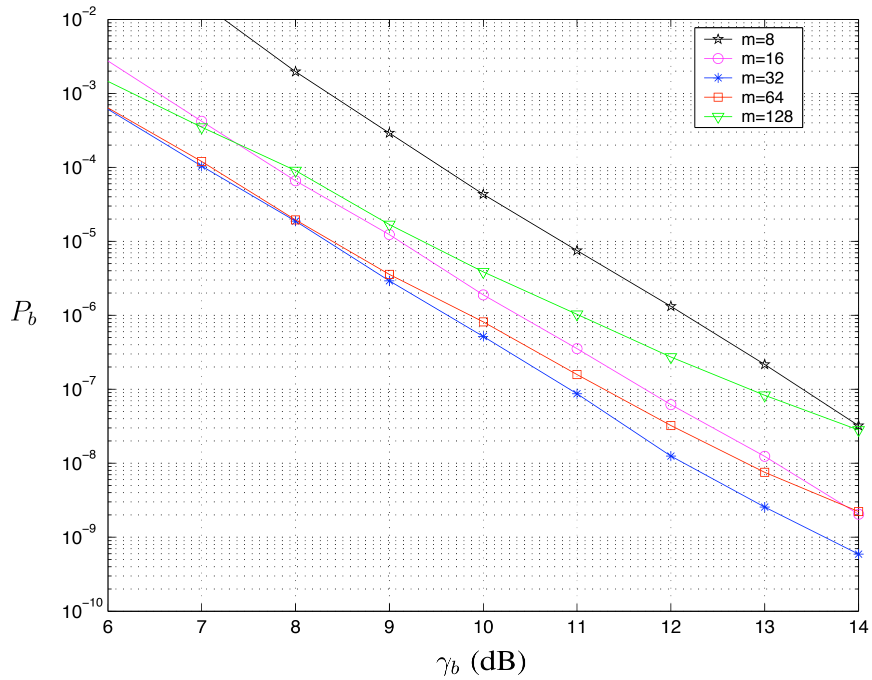

Figure 3.16: Approximation of bit error probability of a convolutionally coded STBC with imperfect SI (CDE assumption with $E_p = E_s$), $n_t = 2$ a frame size $N = 1024$ for channel memory lengths $m = 8, 16, 32, 64, 128$.

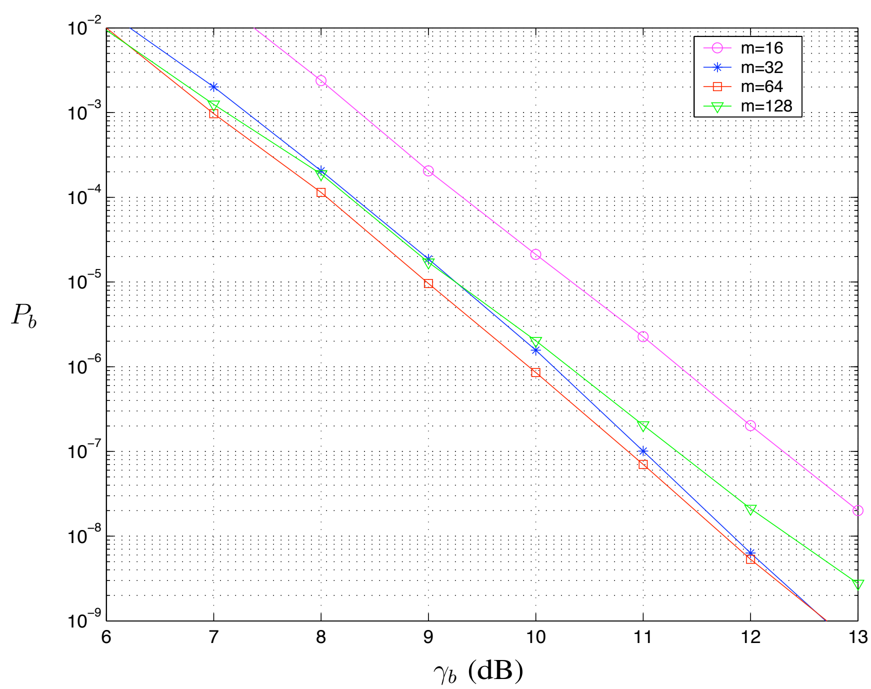

Figure 3.17: Approximation of bit error probability of a convolutionally coded STBC with imperfect SI (OPE receiver with $E_p = E_s$), $n_t = 4$ and a frame size $N = 1024$ for channel memory lengths $m = 16, 32, 64, 128$.

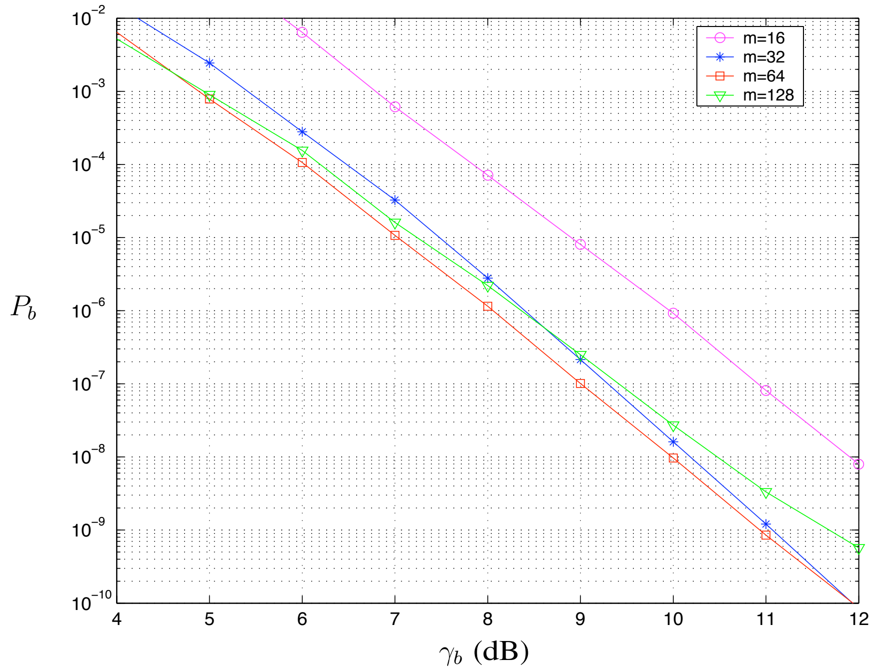

Figure 3.18: Approximation of bit error probability of a convolutionally coded STBC with imperfect SI (CDE assumption with $E_p = E_s$), $n_t = 4$ a frame size $N = 1024$ for channel memory lengths $m = 16, 32, 64, 128$.

mation more crucial. Finally, the length of the pilot sequences increases as the number of transmit antennas increases, which reduces the energy efficiency of the system. In order to accommodate long pilot sequences the channel memory has to increase, which results in longer value for the optimal channel memory as the number of transmit antennas increases.

In Figures 3.19 and 3.20, we show a comparison of systems with channel memory lengths $m = 16$ and $m = 64$ for the cases of $n_t = 2$ and $n_t = 4$, respectively. As in the single-antenna case, the SNR degradation due to channel estimation reduces as the channel memory increases. Also, this SNR degradation increases with the number of transmit antennas making it more crucial to estimate the channel for larger number of transmit antennas. Moreover, using an OPE receiver results in an SNR loss of about 2 and 3 dB for the cases of $n_t = 2$ and $n_t = 4$, respectively. This performance loss is comparable to the loss encountered in differential STBCs. However, using efficient iterative joint decoding and channel estimation receivers may reduce this loss resulting in a

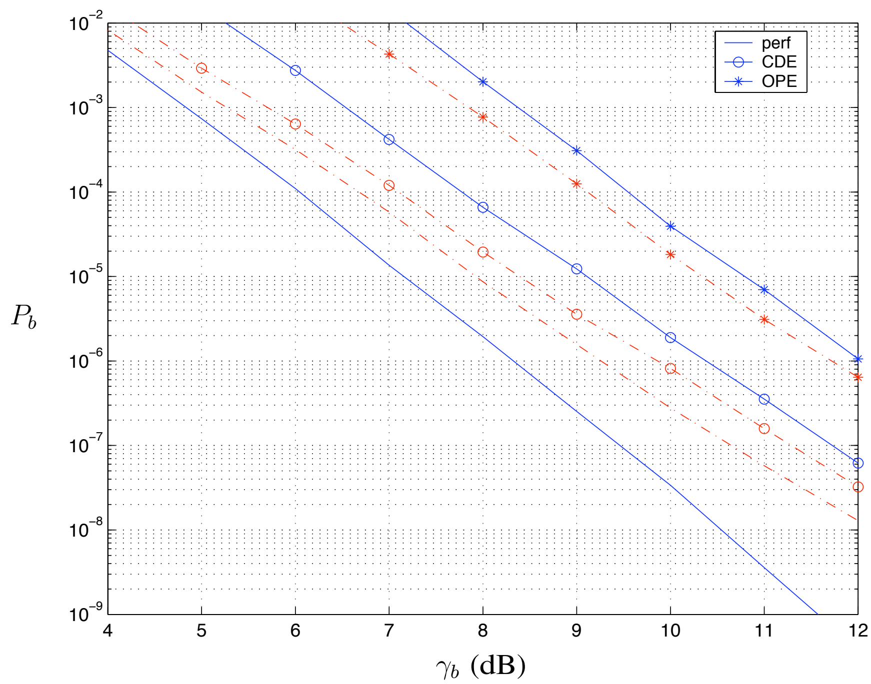

Figure 3.19: Approximation of bit error probability of a convolutional coded STBC systems with a frame size $N = 1024$ and channel memory lengths $m = 16, 64$ using perfect and imperfect SI with $E_p = E_s$ (solid: $m = 16$, dash: $m = 64$).

performance close to that of the system under the CDE assumption. An iterative receiver that performs close to the performance of the CDE assumption is described in Chapter 5.

The energy allocated for the pilot signal is optimized for the multi-antenna systems as shown in Figures 3.21 and 3.22 for the cases of two and four transmit antennas, respectively. From the figures, we observe that the optimal pilot energy allocation is almost independent of the number of transmit antennas. This is mainly because the optimal pilot energy allocation is governed by the ratio of energy spent on estimating the channel, and this ratio is a function of the channel memory length only. As in single-antenna systems, optimizing the pilot energy results in an SNR gain as large as 1 dB over the case where $E_p = E_s$.

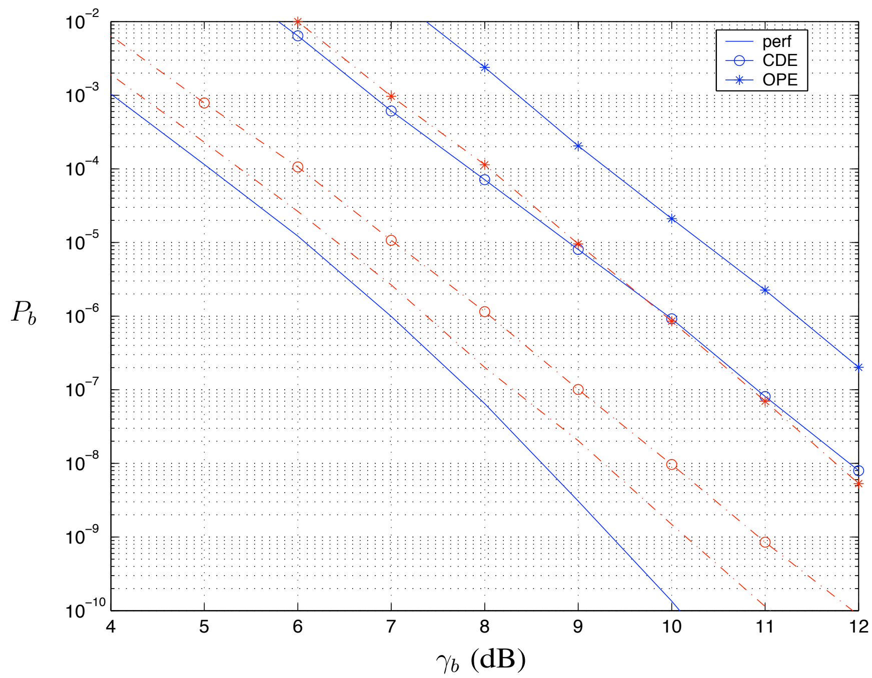

Figure 3.20: The performance of convolutional coded STBC systems with a frame size $N = 1024$ and channel memory lengths $m = 16, 64$ using perfect and imperfect SI with $E_p = E_s$ (solid: $m = 16$, dash: $m = 64$).

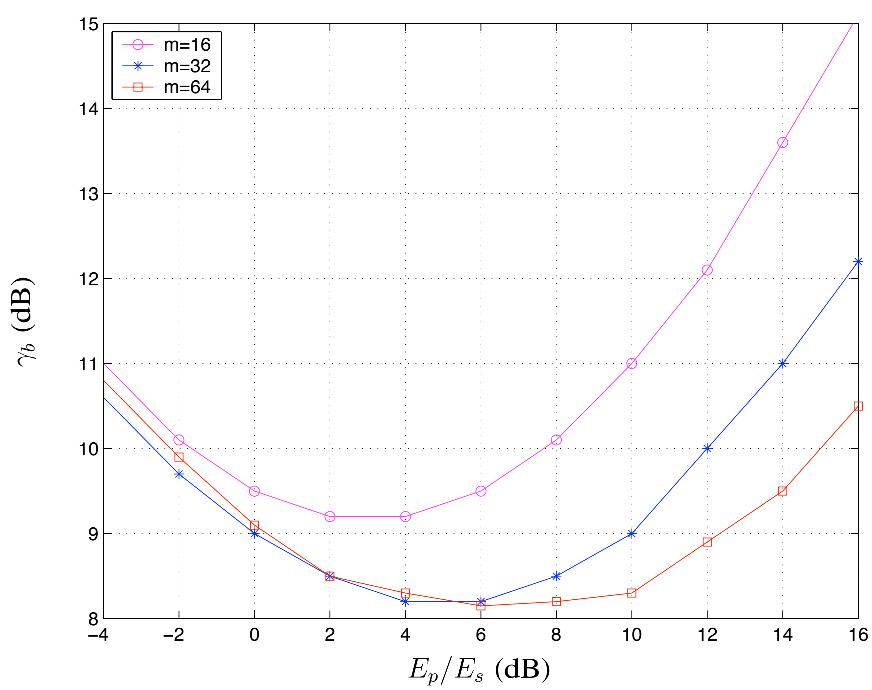

Figure 3.21: SNR required for the convolutionally coded STBC with $n_t = 2$ to achieve $P_b = 10^{-4}$ versus E_p/E_s for the OPE receiver and channel memory lengths $m = 16, 32, 64$.

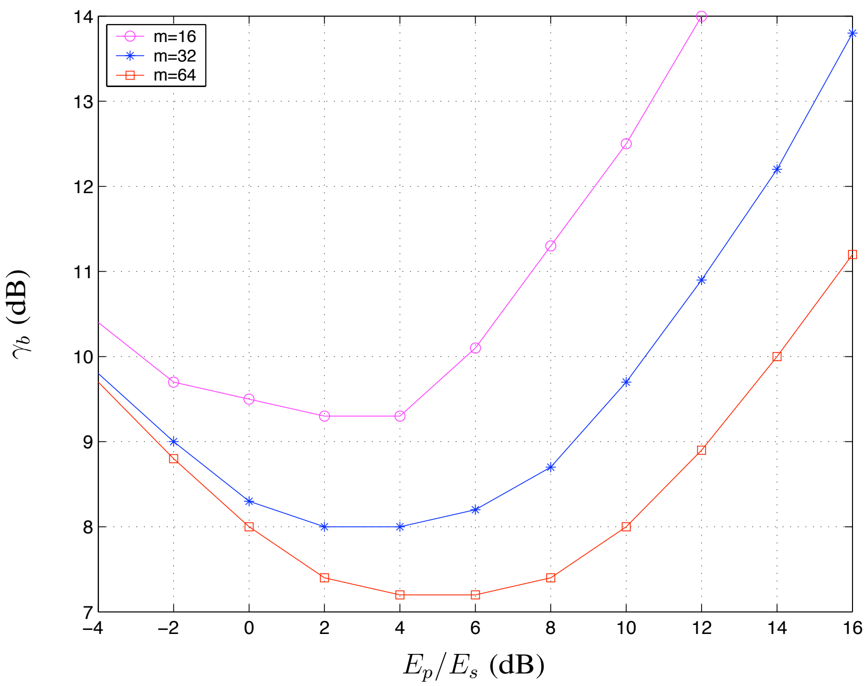

Figure 3.22: SNR required for the convolutionally coded STBC with $n_t = 4$ to achieve $P_b = 10^{-4}$ versus E_p/E_s for the OPE receiver and channel memory lengths $m = 16, 32, 64$.

3.3.3 Correlated Transmit Antennas

In the following discussion, the performance of multi-antenna systems with correlated transmit antennas is derived. The effect of correlation between the antennas in a receive diversity system was modeled in [10, 50] as a function of the distance separating antennas and angles of arrival of the signal beam. In the following discussion, we consider the correlation coefficient as the correlation measure in order to simplify the analysis. In this case, the received signal vector is given by (2.8), where $\mathbf{h}_f$ is a correlated complex Gaussian random vector with a covariance matrix $K_\mathbf{h}$ whose $(i, j)^{th}$ element is given by

$$K_\mathbf{h}(i, j) = \mathrm{E}[h^{i*}h^j] = \begin{cases} 1, & i = j, \\ \rho_{ij}, & i \neq j. \end{cases} \tag{3.44}$$

When perfect SI is available at the receiver, the conditional pairwise error probability is still given by (3.35). Clearly this probability is a function of the inner product $\sum_{i=1}^{n_t} |h_f^i|^2 = \mathbf{h}_f^* \mathbf{h}_f$. Thus the unconditional error probability is found by averaging over the joint probability density function of $\mathbf{h}_f$ given by

$$p(\mathbf{h}_f) = \frac{1}{\pi^{n_t} \det K_{\mathbf{h}}} \exp\left(\mathbf{h}_f^* K_{\mathbf{h}}^{-1} \mathbf{h}_f\right). \tag{3.45}$$

Performing the averaging is difficult due to the complicated form of $p(\mathbf{h}_f)$. To resolve this issue, an uncorrelated random vector $\mathbf{g}_f$ is generated by pre-whitening $\mathbf{h}_f$. Recall the eigenvalue decomposition of the covariance matrix $K_{\mathbf{h}} = U \Lambda U^T$, where Λ is a diagonal matrix containing the eigenvalues of $K_{\mathbf{h}}$, i.e., $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_{n_t}\}$, and U is a unitary matrix that contains the eigenvectors of $K_{\mathbf{h}}$ in its rows. Thus an uncorrelated Gaussian random vector $\mathbf{g}_f$ with a covariance matrix $K_{\mathbf{g}} = \Lambda$ is generated by applying the linear transformation $\mathbf{g}_f = U^T \mathbf{h}_f$ to (3.35) resulting in

$$P_c(d|\mathbf{f}) = \mathrm{Q}\left(\sqrt{2 R_c \gamma_b \sum_{v=1}^{w} v \sum_{l=1}^{f_v} \sum_{i=1}^{n_t} \lambda_i |g_f^i|^2}\right). \tag{3.46}$$

Now, the vector $\mathbf{g}_f$ is a complex Gaussian with distribution given by $\mathcal{CN}(\mathbf{0}, \Lambda)$. By averaging (3.46) over the distribution of $\{\mathbf{g}_f\}$, the unconditional pairwise error probability becomes

$$P_u(d|\mathbf{f}) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{v=1}^{w} \prod_{i=1}^{n_t} \left(\frac{1}{1 + v \lambda_i R_c \gamma_b / \sin^2\theta}\right)^{f_v} d\theta. \tag{3.47}$$

In the case of imperfect SI, there are two methods for estimating the channel as in [51]. In the first method the receiver knows the channel covariance matrix and finds the channel estimates that minimizes the estimation mean square error. This is the optimal linear receiver in the mean square error sense. The second method uses a suboptimal receiver, where the channel is estimated assuming uncorrelated transmit antennas as in Section 3.3.2. In [51], it was shown that the performance of the suboptimal receiver is very close to that of the optimal receiver. In the rest of this Section, we investigate the performance of the suboptimal receiver for a coded STBC systems over block fading channels as a function of the channel covariance matrix.

When the channel is estimated as in Section 3.3.2, the conditional pairwise error probability is given by (3.35) with $\mathbf{h}_f$ replaced by $\mathbf{q}_f = \frac{1}{\sigma}\hat{\mathbf{h}}_f$, where $\sigma^2 = 1 + \sigma_e^2$. Thus the covariance matrix of $\mathbf{q}_f$ is given by

$$K_{\mathbf{q}} = \frac{1}{1+\sigma_e^2} K_{\mathbf{h}} + \frac{\sigma_e^2}{1+\sigma_e^2}\mathbf{I}, \tag{3.48}$$

where $K_{\mathbf{q}}$ is the effective covariance matrix whose diagonals are unity and the off-diagonal elements are $\frac{\rho_{ij}}{1+\sigma_e^2}$. Let $\{\hat{\lambda}_i\}$ be the eigenvalues of $K_{\mathbf{q}}$, then $K_{\mathbf{q}}$ can be diagonalized as above and the pairwise error probability is given by (3.47) with replacing $\{\lambda_i\}$ by $\{\hat{\lambda}_i\}$.

Figure 3.23 shows the SNR required for uncoded STBCs to achieve error rate of $P_b = 10^{-3}$ versus the correlation coefficient between the transmit antennas for the cases of perfect and estimated SI. It is assumed that all the antennas are correlated by the same amount ρ. Thus the diagonal elements of $K_{\mathbf{h}}$ are unity and its off-diagonal elements are ρ. From the figure, we observe that the SNR loss is almost negligible for a correlation coefficient less than $\rho = 0.6$ and $\rho = 0.3$ for the cases of two and four transmit antennas, respectively. This is expected since space diversity increases with increasing the number of antennas, and the permitted amount of correlation to achieve this diversity becomes smaller. Also, under channel estimation environments, the SNR loss due to correlation reduces as the correlation coefficient increases. This suggests that higher antenna correlation results in better channel estimation.

In Figures 3.24 and 3.25, the performance of the coded STBC systems with antenna correlation of $\rho = 0.9$ is shown for the cases of two and four transmit antennas, respectively. By comparing with Figures 3.16 and 3.18 respectively, we observe that antenna correlation degrades the performance of systems with long channel memory more than it does for systems with short channel memory. This is because of two main reasons. First, long channel memory reduces

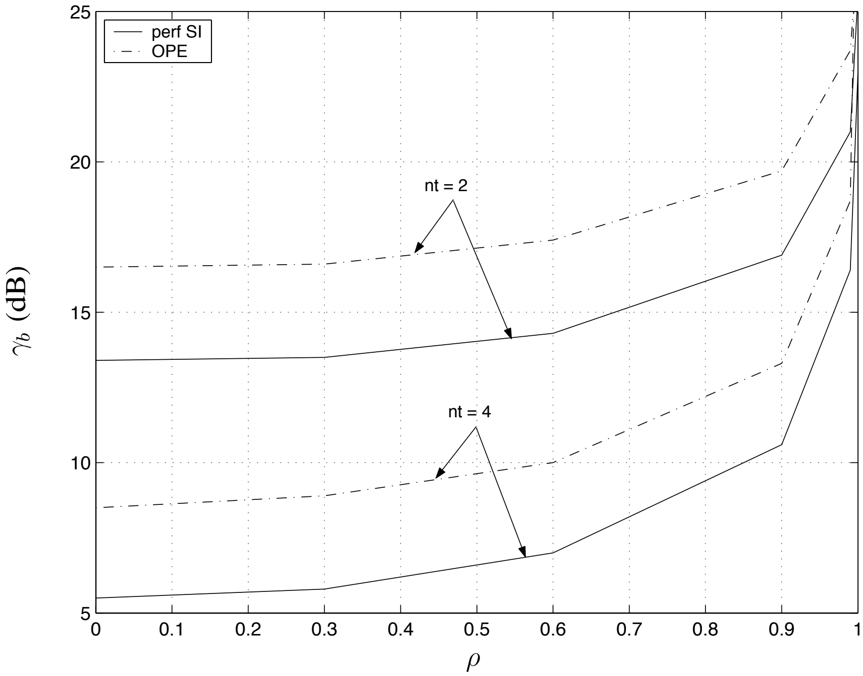

Figure 3.23: SNR required for uncoded STBC to achieve $P_b = 10^{-3}$ versus the correlation coefficient between the transmit antennas for perfect SI and an OPE receiver.

the channel diversity provided by the independent fading blocks causing space diversity to become more crucial to the performance of the system. Thus reducing space diversity by imposing antenna correlation reduces the overall channel diversity which degrades the performance. The second reason is that long channel memory produces better channel estimation, the task that becomes easier due to antenna correlation. Therefore, an advantage of systems with long memory have disappeared, resulting in a degraded performance more than in the systems with short channel memory.

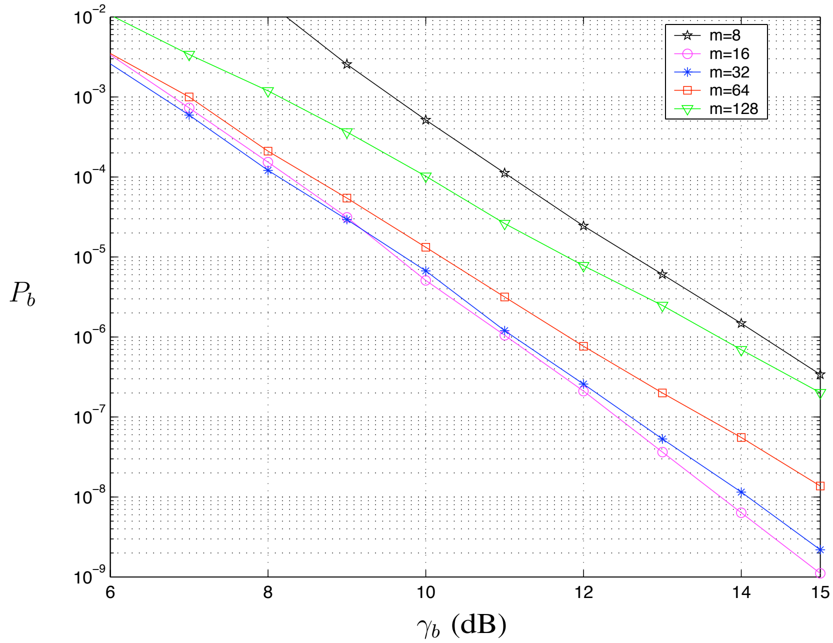

Figure 3.24: Approximation of bit error probability of a convolutionally coded STBC using $n_t = 2$ with antenna correlation coefficient of $\rho = 0.9$, imperfect SI (CDE assumption with $E_p = E_s$) and a frame size $N = 1024$ for channel memory lengths $m = 8, 16, 32, 64, 128$.

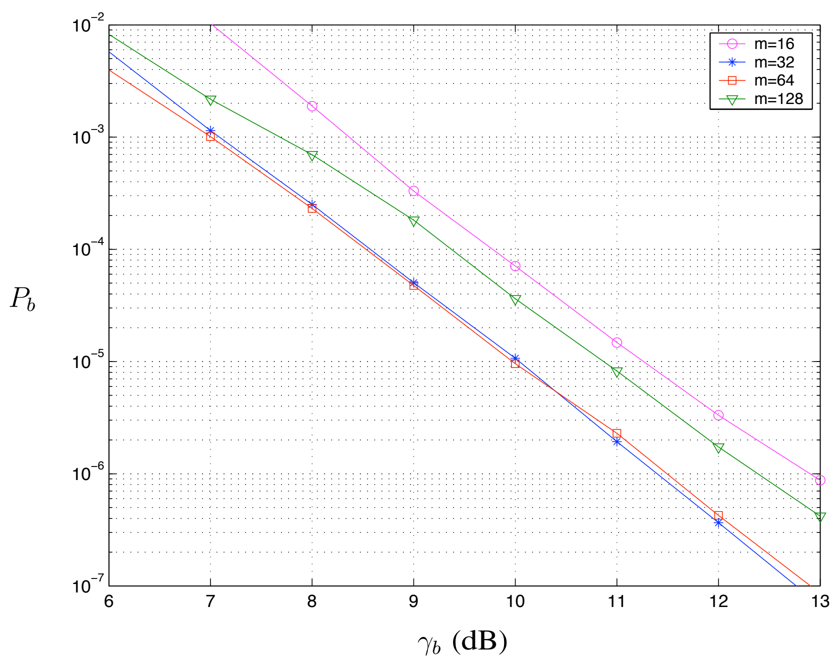

Figure 3.25: Approximation of bit error probability of a convolutionally coded STBC using $n_t = 4$ with antenna correlation coefficient of $\rho = 0.9$, imperfect SI (CDE assumption with $E_p = E_s$) and a frame size $N = 1024$ for channel memory lengths $m = 16, 32, 64, 128$.

CHAPTER 4

Performance of Binary Coded Systems over Rician and Nakagami Block Fading Channels

In this chapter, the union bound for binary coded systems over block fading channel introduced in Chapter 3 is extended to more general fading models. In particular, Rician and Nakagami distributions are used to model the fading process in each fading block. The performance of diversity reception over Rician channels with noncoherent detection was derived by Jacobs [52]. In [12], coherent detection of trellis coded systems was analyzed for the cases of perfect and no amplitude SI available at the receiver. In [53], Charash analyzed the performance of noncoherent communication over multipath Nakagami fading channels with random delays. The bit error probability of coherent diversity reception over Nakagami was derived by Al-Hussaini *et al.* [16], where the block error probability was derived by Noga in [54].

In Chapter 3, a union bound for binary coded single-antenna systems over Rayleigh block fading channels was derived with coherent and noncoherent detection. The union bound is based on uniform interleaving of the coded sequence prior to transmission over the channel, and the distribution of the erroneous bits over the fading blocks is computed. Results showed that the bound dictates the performance of coded systems for a wide range of SNR. For Rician channels, the importance of channel diversity becomes less significant as the specular-to-diffuse ratio increases because the fading random variable becomes less random. Similarly, as the fading severity of a Nakagami distributed channel is increased, the significance of diversity reduces. As in Rayleigh fading channels, a method to analyze the performance of coded systems over block fading channels with Rician and Nakagami distributions is needed. In this chapter, the performance of coded systems over Rician and Nakagami block fading channels is studied using the union bound introduced in Chapter 3. Furthermore, we investigate the effect of channel memory on the system performance and its relation to the parameters of the channel such as the specular-to-diffuse ratio in Rician channels and the fading severity in Nakagami channels. The corresponding pairwise error probabilities are derived for noncoherent detection using a square-law combining and coherent detection with perfect, imperfect and no amplitude SI available at the receiver. Furthermore, the union bound is used with the assumption of imperfect SI at the receiver to investigate the tradeoff between the channel diversity and channel estimation, and the effect of the specular-to-diffuse ratio of the channel on the optimal channel memory.

The chapter is organized as follows. In Sections 4.1, expressions for the pairwise error probability are derived for Rician block fading channels. Noncoherent and coherent receivers are considered. In coherent receivers, different assumptions on the channel SI are assumed and the corresponding pairwise error probability is derived, where square-law combining is used in noncoherent receivers. Block fading channels with Nakagami fading distribution are considered in Section 4.2. Numerical results are presented for the Rician and Nakagami distributions following the derivation of the results.

4.1 Rician Fading

In this section, we derive expressions for the pairwise error probability of coded systems over Rician block fading channels. In Rician block fading channels, the channel gain in each fading block h_f is modeled as a complex Gaussian variable with $\mathcal{CN}(b, 1)$, where b represents the specular component of the channel. Thus the amplitude a_f has a Rician distribution with a normalized density function [12] given by

$$f_a(a) = 2a(1+K)\exp\left[-K - a^2(1+K)\right] I_0\left(2a\sqrt{K(1+K)}\right), \qquad a \geq 0, \tag{4.1}$$

where $K = b^2$ is the energy of the specular component and $I_0(.)$ is the zero-order modified Bessel function of the first kind. In this context, K denotes the ratio of the specular component energy to the diffuse component energy. The unconditional pairwise error probability $P_u(d|\mathbf{f})$ is derived for coherent detection with perfect and imperfect SI as well as the intermediate situation where no amplitude SI is available at the receiver. Also, noncoherent detection using square-law combining is considered.

4.1.1 Coherent Detection - Perfect SI

When the channel SI is known perfectly at the receiver, the ML decoder chooses the codeword that maximizes the metric (2.2). As was shown in Section 3.2, the conditional pairwise error probability is given by (3.11). To find the unconditional pairwise error probability $P_u(d|\mathbf{f})$, (3.11) is averaged over the statistics of the fading amplitudes in (4.1) as in (3.12). The Chernoff upper bound for the unconditional pairwise error probability [12] is given by

$$P_u(d|\mathbf{f}) \leq \frac{1}{2}\prod_{v=1}^{w}\left(\frac{1+K}{1+K+vR_c\gamma_b}\right)^{f_v}\exp\left(-\frac{Kvf_vR_c\gamma_b}{1+K+vR_c\gamma_b}\right)d\theta. \tag{4.2}$$

An exact expression of the pairwise error probability is found by using the integral expression of the Q-function resulting in

$$P_u(d|\mathbf{f}) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{v=1}^{w}\left(\frac{1+K}{1+K+vR_c\gamma_b/\sin^2\theta}\right)^{f_v}\exp\left(-\frac{Kvf_vR_c\gamma_b/\sin^2\theta}{1+K+vR_c\gamma_b/\sin^2\theta}\right)d\theta. \tag{4.3}$$

The union bound was evaluated for a rate-$\frac{1}{2}$ (23,35) convolutional code with a frame size of $N = 2 \times 512$ coded bits. The union bound (3.1) was approximated by including only codewords with distances $d \leq 12$ and using (4.3). In this chapter, we show only the analytical results since the union bound is very close to simulation as was demonstrated in Chapter 3. Figure 4.1 shows the analytical results of the rate-$\frac{1}{2}$ (23,35) convolutional code over Rician fading channels with specular-to-diffuse ratios $K = 1, 10$ dB and perfect SI. We observe that the SNR degradation due to longer channel memory is more severe in the case of $K = 1$ dB than the case of $K = 10$ dB. This is clear from Figure 4.2, where the SNR required for a (23,35) convolutional code to achieve bit error rate of $P_b = 10^{-4}$ is plotted versus the specular-to-diffuse ratio K of the Rician channel. This shows that increasing the energy of the line-of-site component of the channel reduces the effect of the diversity provided by the independent fading blocks. This is expected since increasing K causes the energy of the specular component to increase and the channel approaches the "no fading" scenario, where diversity becomes less important. Hence, we conclude that the smaller the specular-to-diffuse ratio is, the more sensitive the performance becomes to the lack of channel diversity.

4.1.2 Coherent Detection - Imperfect SI

As in Section 3.2.2, when the channel SI is unknown at the receiver, it can be estimated via transmitting pilot signals. In such a system, a pilot signal with energy E_p is transmitted in each fading block. The corresponding received signal is given by (3.15) and the ML estimator for h_f is written as $\hat{h}_f = \frac{y_{f,p}}{\sqrt{E_p}} = h_f + e_f$, where $e_f = \frac{z_{f,p}}{\sqrt{E_p}}$ is the estimation error modeled as $\mathcal{CN}(0, \sigma_e^2)$ with $\sigma_e^2 = \frac{N_0}{E_p}$. The correlation coefficient between the true channel gain and its estimate is given by

$$\mu = \frac{\mathrm{E}[(h_f - b)(\hat{h}_f - b)^*]}{\sqrt{\mathrm{Var}(h_f)\mathrm{Var}(\hat{h}_f)}} = \frac{1}{\sqrt{1 + \sigma_e^2}}. \tag{4.4}$$

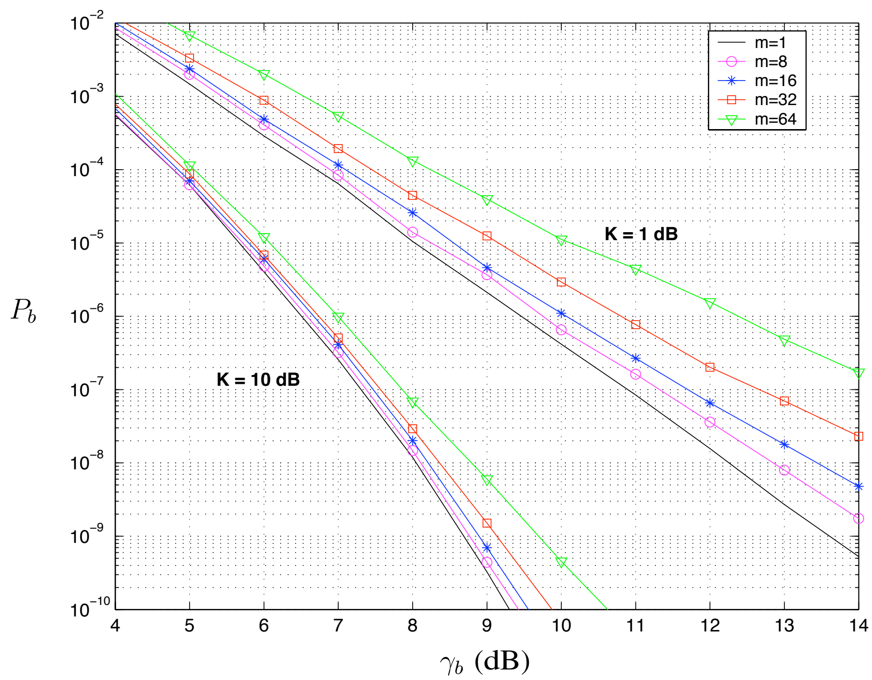

Figure 4.1: Approximation of the bit error probability of a rate-$\frac{1}{2}$ (23,35) convolutional code over a Rician fading channel with $K = 1, 10$ dB, perfect SI and a frame size $N = 1024$ for memory lengths $m = 1, 8, 16, 32, 64$.

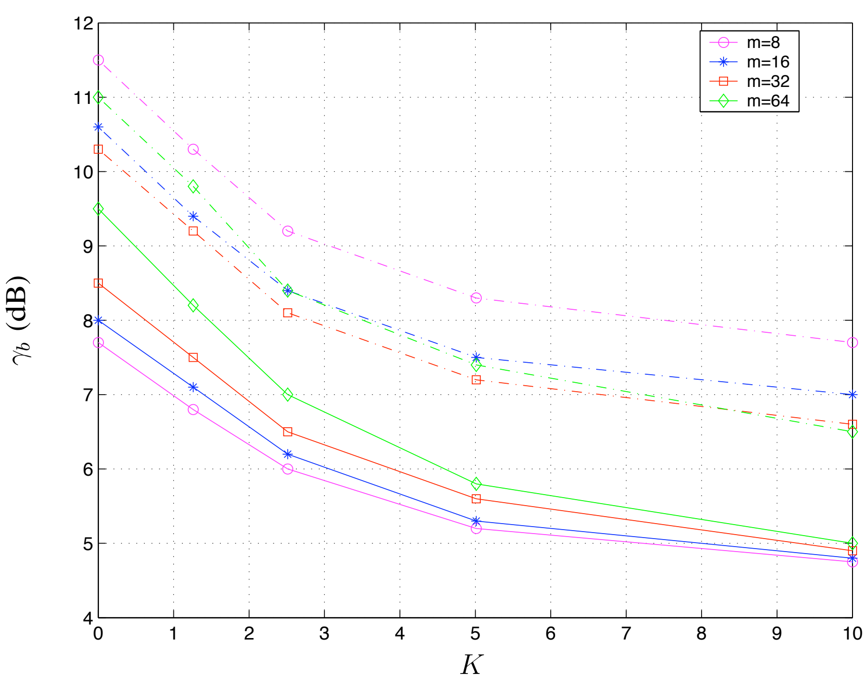

Figure 4.2: SNR required for a (23,35) convolutional code to achieve $P_b = 10^{-4}$ versus the specular-to-diffuse ratio K (linear scale) for memory lengths $m = 8, 16, 32, 64$ (solid: perfect SI, dash: OPE).

The ML decoding rule is found by maximizing the likelihood function (3.18), where the fading density in the Rician fading case is $e^{-|h_f-b|^2}$. The derivation for Rician fading is essentially the same as in (3.18)-(3.20) with the only change being $\beta_f = \sqrt{E_p}y_{f,p} + \sqrt{E_s}\sum_{l=1}^{m} y_{f,l}s^*_{f,l} + bN_0$. Thus based on observing the received and the pilot signals, the ML decoder chooses the codeword $\mathbf{S}$ that maximizes the metric

$$\mathbf{m}(\mathbf{Y},\mathbf{S}) = \sum_{f=1}^{F} E_s|\sum_{l=1}^{m} y_{f,l}s^*_{f,l}|^2 + 2\sqrt{E_s}\sum_{l=1}^{m}\mathrm{Re}\{y^*_{f,l}s_{f,l}(\hat{h}_f + bN_0)\} + C, \tag{4.5}$$

where C is a constant term independent of the transmitted codeword. As discussed in Section 3.2.2, this receiver is difficult to be implemented in a Viterbi decoder. Therefore, a suboptimal decoding metric that maximizes the likelihood function $p(\mathbf{Y}|\hat{\mathbf{H}},\mathbf{S})$ is employed. It is given by

$$\mathbf{m}(\mathbf{Y},\mathbf{S}) = \sum_{f=1}^{F}\sum_{l=1}^{m}\mathrm{Re}\{y^*_{f,l}\hat{h}_f s_{f,l}\}. \tag{4.6}$$

Thus the conditional pairwise error probability for the suboptimal decoder becomes

$$P_c(d|\mathbf{f}) = \Pr\left(\sum_{f=1}^{L}\sum_{l=1}^{m}\mathrm{Re}\{y^*_{f,l}\hat{h}_f\} < 0\middle|\hat{\mathbf{H}},\mathbf{S}\right). \tag{4.7}$$

The received signal $y_{f,l}$ conditioned on $\hat{h}_f$ is a complex Gaussian random variable with mean $\sqrt{E_s}s_{f,l}\mathrm{E}[h|\hat{h}]$ and variance $N_0 + (1-\mu^2)E_s$, where $\mathrm{E}[h|\hat{h}] = \frac{\mu}{\sigma}(\hat{h}_f - b) + b$. Thus the conditional pairwise error probability for the suboptimal decoder is given by

$$P_c(d|\mathbf{f}) = \mathrm{Q}\left(\sqrt{\frac{2E_s\sum_{f=1}^{F} d_f\left|\frac{\mu}{\sigma}(\hat{h}_f - b) + b\right|^2}{N_0 + (1-\mu^2)E_s}}\right), \tag{4.8}$$

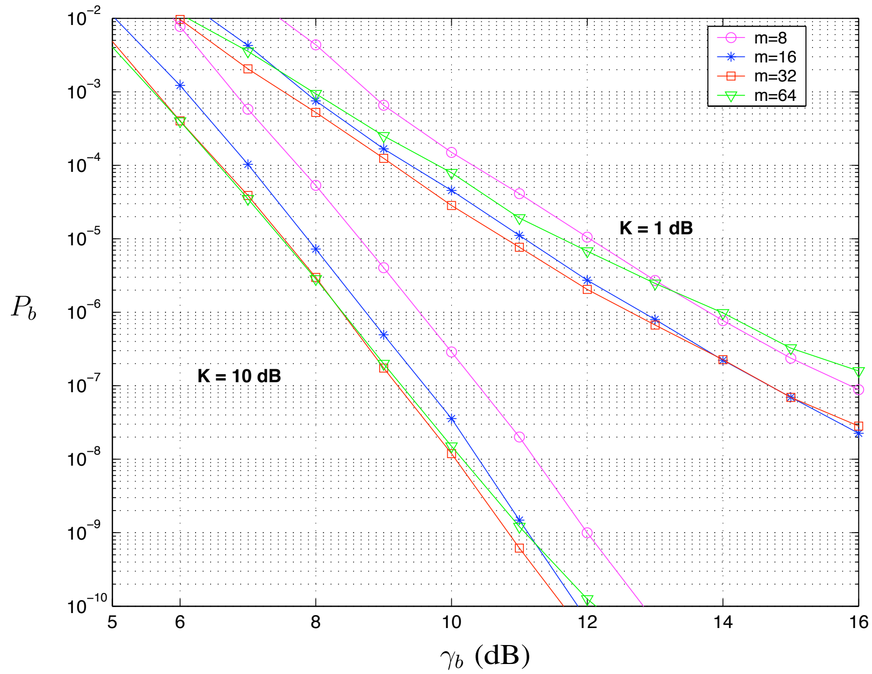

Figure 4.3: Approximation of the bit error probability of a (23,35) convolutional code over a Rician fading channel with $K = 1, 10$ dB, imperfect SI (OPE receiver) and a frame size $N = 1024$ for memory lengths $m = 8, 16, 32, 64$.

where d_f is the number of nonzero error bits in fading block f. Define the normalized complex Gaussian random variable $\zeta_f = \frac{\hat{h}_f - b}{\sigma} + \frac{b}{\mu}$ with distribution $\mathcal{CN}(\frac{b}{\mu}, 1)$. Then the conditional pairwise error probability simplifies to (3.23) with ζ_f is as defined above. Therefore, the pairwise error probability for the case of imperfect SI is the same as that of perfect SI, by replacing γ_b by $\hat{\gamma}_b = \frac{\mu^2 \gamma_b}{1 + R_c \gamma_b (1 - \mu^2)}$ and K by $\frac{K}{\mu^2}$.

In evaluating the union bound for pilot-aided channel estimation, the energy of the pilot is taken into account in the SNR axis and the code is punctured to maintain the same transmission rate for systems with different channel memory. From Figure 4.2, the optimal channel memory value for an OPE receiver with $E_p = E_s$ is $m = 32$ for a Rayleigh fading channel, i.e. $K = 0$, where it is $m = 64$ for a Rician channel with $K = 10$. Also, note that the case of $m = 8$ outperforms the case of $m = 64$ when the channel is more fading, where the reverse occurs for channels that are less faded, i.e., larger values of K. Figures 4.3 and 4.4 show the results of imperfect SI with an OPE receiver

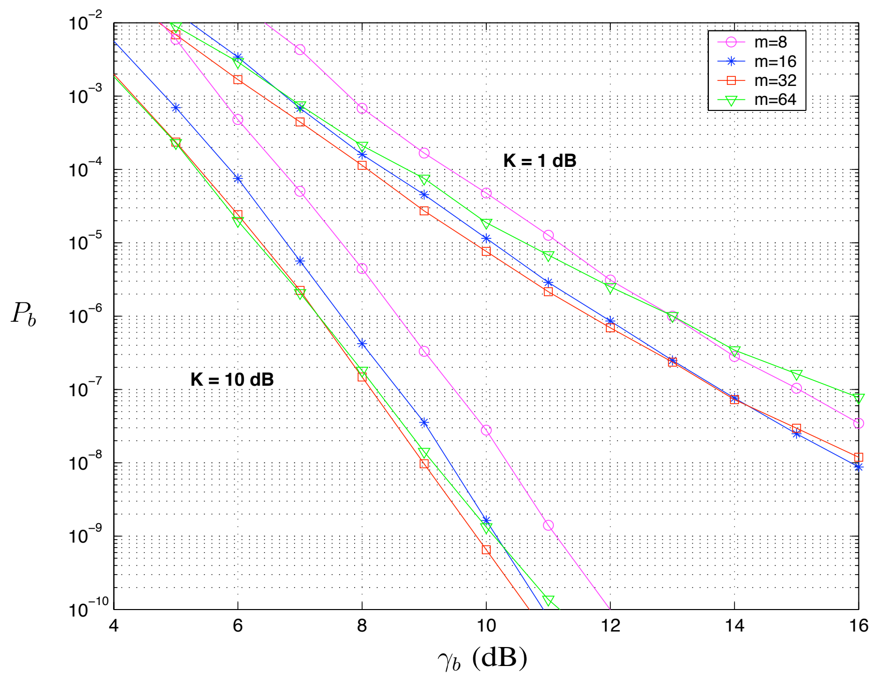

Figure 4.4: Approximation of the bit error probability of a (23,35) convolutional code over a Rician fading channel with $K = 1, 10$ dB, imperfect SI (CDE assumption) with $E_p = E_s$ and a frame size $N = 1024$ for memory lengths $m = 8, 16, 32, 64$.

and the CDE assumptions, respectively. As in Chapter 3, the gain loss in SNR due to the channel memory is less compared to the case of perfect SI. Also, we observe that systems with long channel memory perform better as the energy of the specular component of the channel increases. This is because as K increases the channel becomes less fading which reduces the need for the decoder to average over the statistics of the channel. Therefore, the channel diversity becomes less crucial causing in systems with long channel memory to outperform systems with short memory. Another reason for this is the larger energy fraction spent on pilot signals in systems with short channel memory lengths than in systems with long memory. This is obvious for the case of $K = 10$ dB, where the performance of $m = 64$ is nearly optimal for most of the SNR values. On the other hand, the case of $m = 8$ is the worst every where when $K = 10$ dB, where it outperforms the case of $m = 64$ when $K = 1$ dB.

Figure 4.5 shows a comparison of systems with channel memory lengths

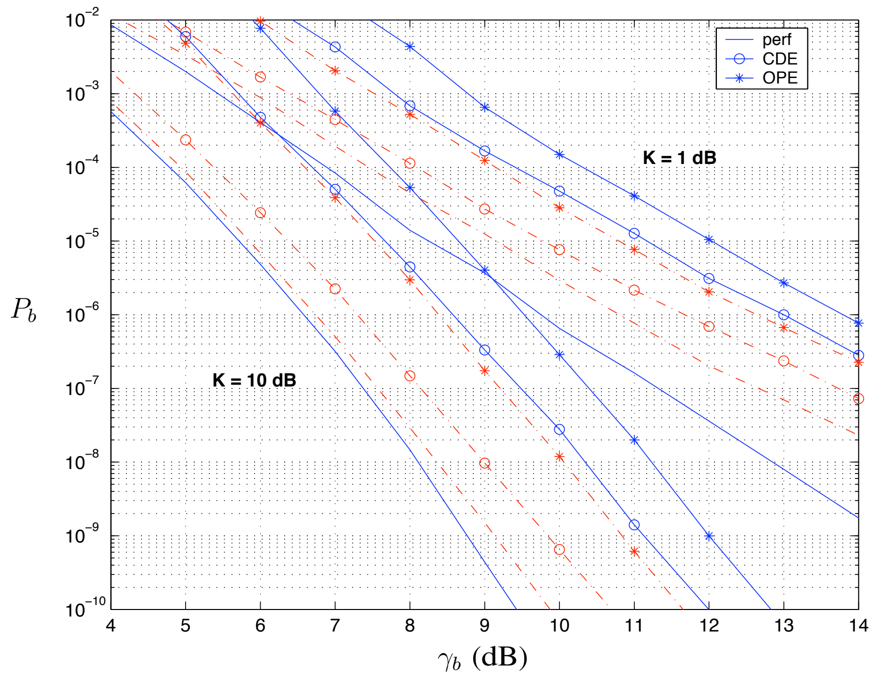

Figure 4.5: Approximation of the bit error probability of a rate-$\frac{1}{2}$ (23,35) convolutional code over a Rician fading channel with $K = 1, 10$ dB, frame size $N = 1024$ and memory lengths $m = 8, 32$ using perfect and imperfect SI with $E_p = E_s$ (solid: $m = 8$, dash: $m = 32$).

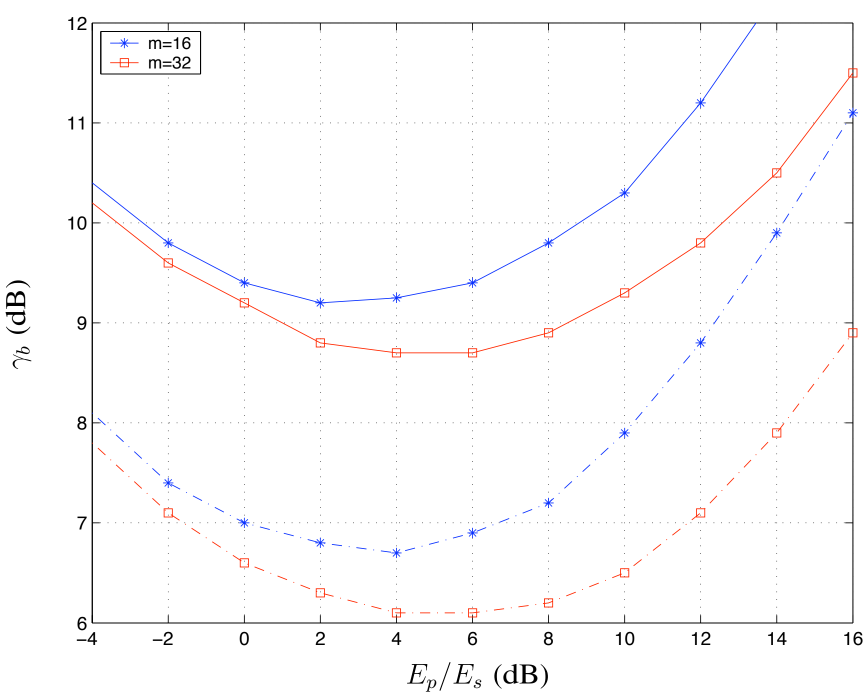

Figure 4.6: SNR required for a rate-$\frac{1}{2}$ (23,35) convolutional code to achieve $P_b = 10^{-4}$ versus E_p/E_s for the OPE receiver with $E_p = E_s$ and memory lengths $m = 16, 32$.

$m = 8$ and $m = 32$. As in Chapter 3, the SNR degradation due to channel estimation reduces as the channel memory increases. Moreover, the SNR loss in OPE receivers with long channel memory increases with increased energy of the specular component of the channel. When the channel is estimated using a pilot signal, the channel estimation error adds a fading component to the channel gain at the decoder. The effect of this new fading component increases as the energy of the specular component increases of the channel, which degrades the performance of OPE receivers more as K increases. The energy allocated for the pilot signal is optimized as shown in Figure 4.6. We observe that the optimal pilot energy allocation is almost independent of the fading nature of the channel, i.e., independent of the energy of the specular component K of the channel. As discussed in Chapter 3, this is because the amount of energy available in each fading block, which can be used in estimating the channel, is the controlling factor of the optimal pilot energy allocation. Clearly, this energy amount is a function of the channel memory length only. Also, the SNR gain resulting from optimizing the pilot energy is almost independent of the channel fading behaviour.

4.1.3 Coherent Detection - No Amplitude SI

For completeness we consider the case where the channel phase is known but the amplitude is unknown. The suboptimal receiver that uses the metric (2.3) in decoding is employed due to its mathematical tractability. Using this metric and going through the derivation in Section 3.2.3, the conditional pairwise error probability is upper bounded using the Chernoff bound as in (3.26). The full derivation is included in Appendix B.1, and the final expression of the pairwise error probability for coherent detection with no amplitude SI at the receiver is given by

$$P_u(d|\mathbf{f}) \lesssim e^{-K} \left(\frac{d(1+K)}{2R_c\gamma_b L} \right)^L \left(\prod_{v=1}^{w} v^{2f_v} \right)^{-1}, \tag{4.9}$$

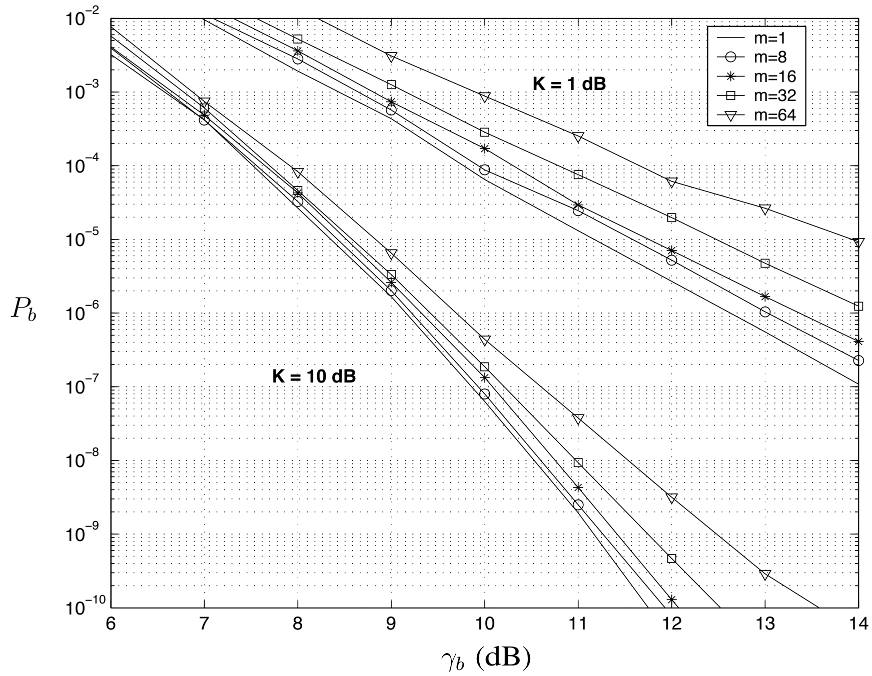

Figure 4.7: Approximation of the bit error probability of a rate-$\frac{1}{2}$ (23,35) convolutional code over a Rician fading channel with $K = 1, 10$ dB, no amplitude SI and a frame size $N = 1024$ for memory lengths $m = 1, 8, 16, 32, 64$.

where $e = \exp(1)$. The union bound is evaluated using (4.9) for coherent detection with no amplitude SI at the receiver and shown in Figure 4.7. We observe that the SNR degradation due to increasing the channel memory is clear from the bound.

4.1.4 Noncoherent Detection

For noncoherent communications over Rician fading channels, where the phase is unknown at the receiver, the optimal detection rule was derived by Jacobs [52]. A suboptimal detection scheme using a square-law combining is used, where its performance was derived in [52]. We use a square-law combining receiver for noncoherent communications over Rician block fading channels. The outputs of the square-law combiner are given by (2.5). The decoder uses these outputs and the suboptimal metric (2.6) for decoding. In this setup, the conditional pairwise error probability is found by substituting the metric (2.6) in (3.9) resulting in (3.28). For Rician fading distribution, the variables $r_f^{(I,0)}, r_f^{(Q,0)}$ are Gaussian random variables with means equal to $\sqrt{E_s}b\cos(\theta_f)$ and $\sqrt{E_s}b\sin(\theta_f)$, respectively and variance of $\frac{1}{2}(E_s + N_0)$. Also, $r_f^{(I,1)}, r_f^{(Q,1)}$ are zero mean Gaussian random variables with variance $\frac{1}{2}N_0$. Let $|r_f^{(c)}|^2 = |r_f^{(I,c)}|^2 + |r_f^{(Q,c)}|^2$ for $c = 0, 1$ and define $\kappa = \sum_{f=1}^{F} d_f(|r_f^{(1)}|^2 - |r_f^{(0)}|^2)$. Then, the unconditional pairwise error probability is found by averaging (3.29) over the density function $p(\kappa)$ and collecting terms having the same distance to arrive to (3.30). The following identity for a random variable x with $\mathcal{CN}(b, \sigma^2)$ distribution [52] is used in averaging (3.29) over the density of κ. The identity is given by

$$\mathrm{E}\left[e^{\omega x^2}\right] = \frac{\exp\left(\omega b^2/(1-2\omega\sigma^2)\right)}{\sqrt{1-2\omega\sigma^2}}, \quad \mathbf{Re}(\omega) < \frac{1}{2\sigma^2}. \tag{4.10}$$

Optimizing the Chernoff parameter λ as in [52], the resultant Chernoff bound on the pairwise error probability is given by

$$P_u(d|\mathbf{f}) \leq \prod_{v=1}^{w} \left[4D_v(1-D_v)\right]^{f_v} \exp\left(-\frac{Kvf_vR_c\gamma_b}{2+vR_c\gamma_b}\right), \tag{4.11}$$

where $D_v = \frac{1}{2+vR_c\gamma_b}$. The union bound is evaluated for convolutionally coded systems with BFSK signaling and square-law combining and results are shown in Figure 4.8. Note that the loss in performance due to the smaller amount of channel diversity is clear. Again this performance loss increases with increasing the energy of the specular component of the channel.

4.2 Nakagami Channels

This section is devoted to deriving the pairwise error probability of coded systems over Nakagami block fading channels. In Nakagami block fading channels, the fading amplitude in each fading block is Nakagami distributed with a

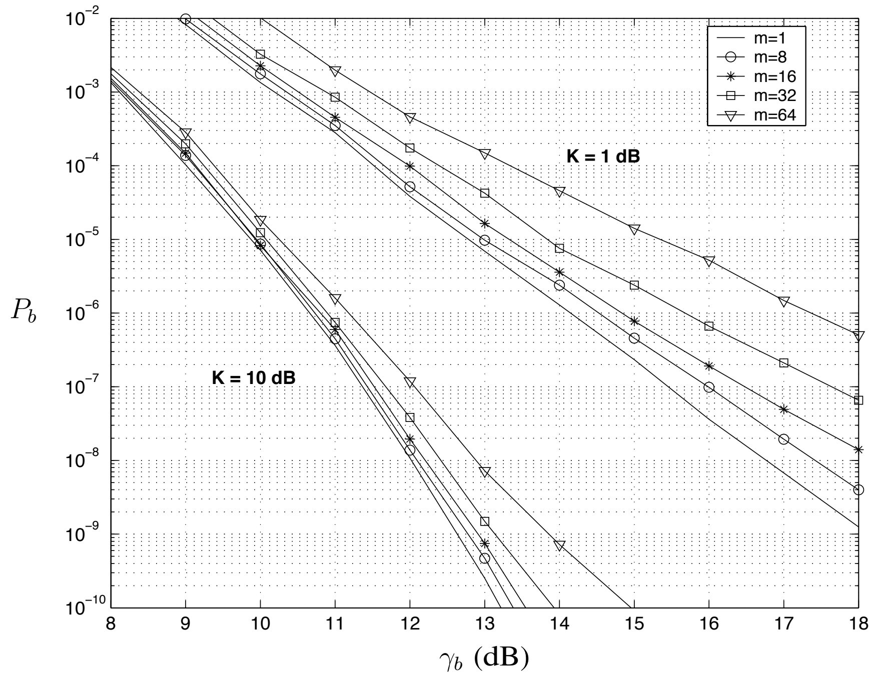

Figure 4.8: Approximation of the bit error probability of a rate-$\frac{1}{2}$ (23,35) convolutional code over Rician fading channel with $K = 1, 10$ dB, noncoherent detection and a frame size $N = 1024$ for memory lengths $m = 1, 8, 16, 32, 64$.

normalized density function [15] given by

$$f_a(a) = \frac{2M^M}{\Gamma(M)A^M} a^{2M-1} \exp\left(-\frac{Ma^2}{A}\right), \quad a > 0, M > 0.5, \tag{4.12}$$

where $A = \mathrm{E}[a^2] = 1$, $M = \frac{A^2}{\mathrm{Var}[a]}$ is the fading parameter and $\Gamma(.)$ is the Gamma function. As M increases, the fading becomes less severe and reaches the non-fading case when $M \to \infty$. The Nakagami distribution covers a wide range of fading scenarios including Rayleigh fading when $M = 1$ and single-sided Gaussian distribution when $M = 0.5$. In the following, we consider coherent detection with perfect SI and no amplitude SI as well as noncoherent detection using square-law combining.

4.2.1 Coherent Detection - Perfect SI

If the channel SI is known perfectly at the receiver, the ML decoder uses the metric (2.2) for decoding. As in Section 3.2, the conditional pairwise error probability is given by (3.11). Averaging over the statistics of the fading amplitude (4.12), an exact expression of $P_u(d|\mathbf{f})$ is found as in [55]

$$P_u(d|\mathbf{f}) = \frac{1}{\pi}\mathrm{E}_{\mathbf{A}}\left[\int_0^{\frac{\pi}{2}} \exp\left(-\frac{R_c\gamma_b}{\sin^2\theta}\sum_{v=1}^{w} v\sum_{i=1}^{f_v} a_i^2\right) d\theta\right]$$

$$= \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \prod_{v=1}^{w}\left(\frac{1}{1+\frac{vR_c\gamma_b}{M\sin^2\theta}}\right)^{Mf_v}. \tag{4.13}$$

In computing the union bound (4.13) is used in (3.6). The results for the rate-$\frac{1}{2}$ (5,7) convolutional code are shown in Figure 4.9 for memory lengths ranging from $m = 1$ to $m = 64$. From the figure, we see that the effect of increasing the channel memory length is more severe in the case of $M = 0.5$. This is expected since as M increases, the fading amount increases and the channel approaches the "no fading" scenario. The effect of the fading amount on the performance of coherent detection with perfect SI is shown in Figure 4.10. We observe that as the fading parameter increases, the SNR loss due to channel memory reduces. This is expected since as the fading parameter increases, the diversity becomes less important since the channel is approaching the AWGN channel.

4.2.2 Coherent Detection - No Amplitude SI

Coherent detection with known phase and unknown amplitude SI at the receiver is considered below. We employ the suboptimal receiver that maximizes the metric (2.3) for its simplicity. Using this metric and going through the derivation in Section 3.2.3, the conditional pairwise error probability is upper bounded using the Chernoff bound as in (3.26). The derivation is given in Appendix B.2. The final expression which uses an approximation at high SNR

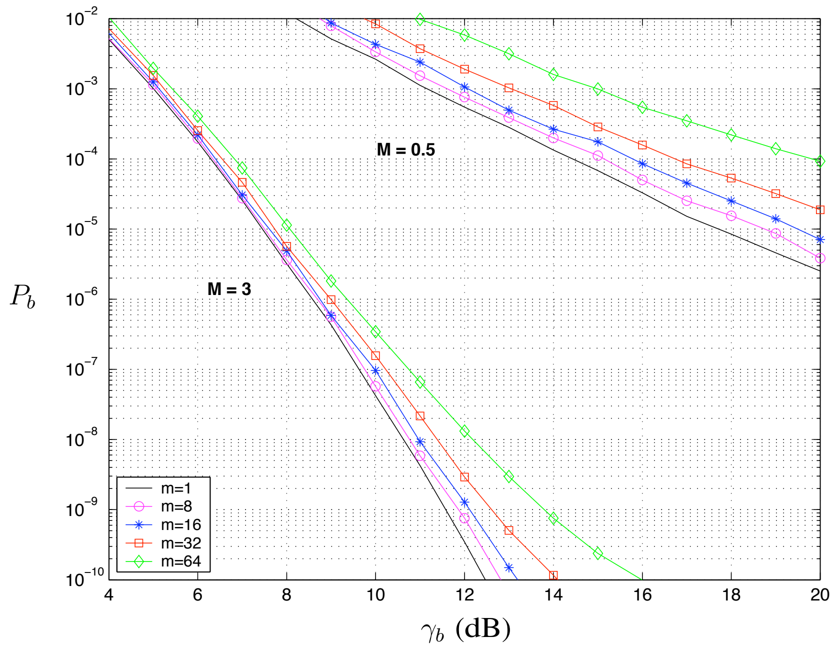

Figure 4.9: Approximation of the bit error probability of a rate-$\frac{1}{2}$ (5,7) convolutional code over Nakagami fading channel with Nakagami parameter $M = 0.5, 3$, perfect SI and a frame size $N = 1024$ for memory lengths $m = 1, 8, 16, 32, 64$.

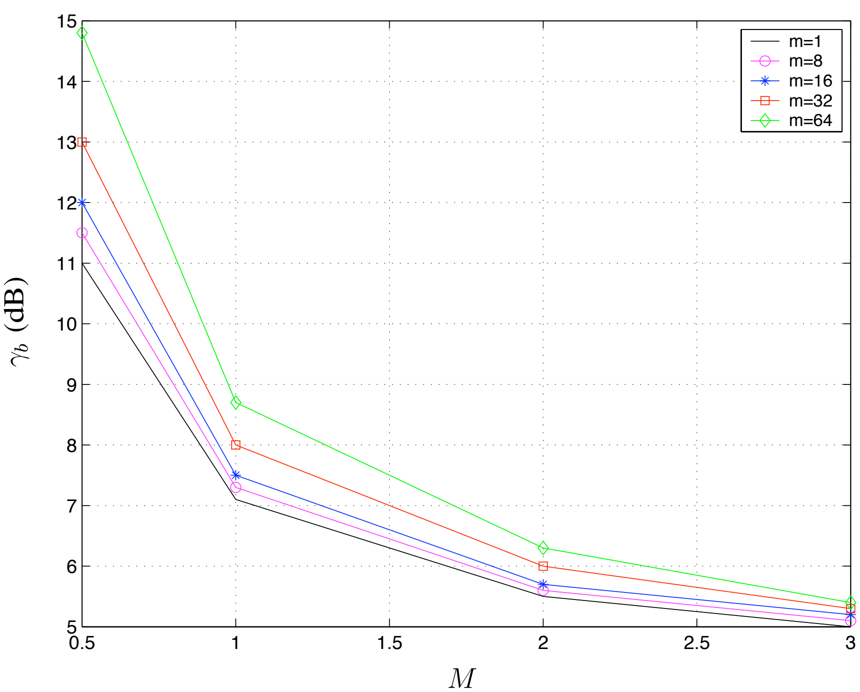

Figure 4.10: SNR required for the (5,7) convolutional code with perfect SI to achieve $P_b = 10^{-3}$ versus the Nakagami parameter M for memory lengths $m = 1, 8, 16, 32, 64$.

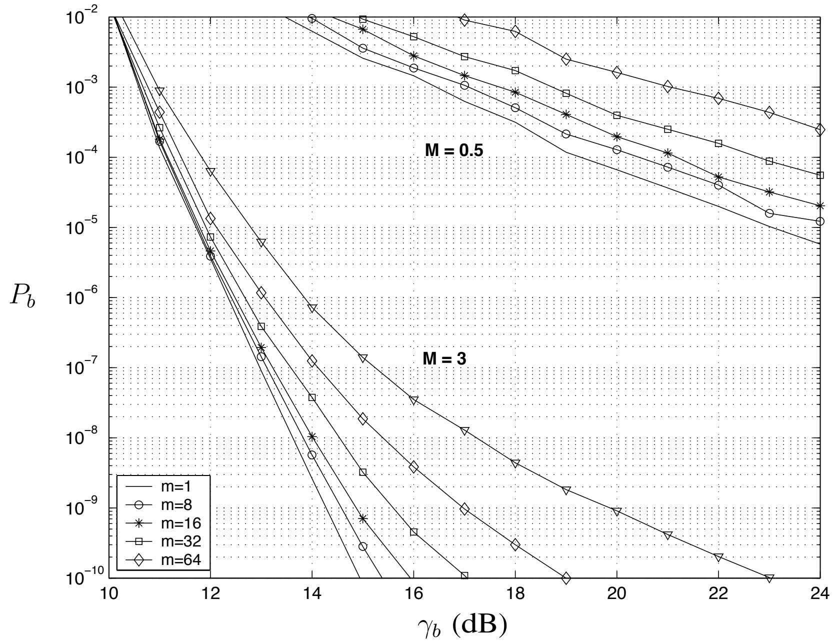

Figure 4.11: Approximation of the bit error probability of a rate-$\frac{1}{2}$ (5,7) convolutional code over Nakagami fading channel with Nakagami parameter $M = 0.5, 3$, no amplitude SI and a frame size $N = 1024$ for memory lengths $m = 1, 8, 16, 32, 64$.

is

$$P_u(d|\mathbf{f}) \lesssim \left(\frac{\Gamma(2M)}{\Gamma(M)2^{M-1}}\right)^L \left(\frac{de}{2R_c\gamma_b L}\right)^{ML} \left(\prod_{v=1}^{w} v^{2Mf_v}\right)^{-1}. \tag{4.14}$$

The results for coherent detection with no amplitude SI are shown in Figure 4.11. Again, as the fading parameter M increases the SNR degradation due to memory reduces because the channel becomes less fading.

4.2.3 Noncoherent Detection

In noncoherent systems, the channel phase is unknown at the receiver. In [53], Charash derived the optimal detection rule for noncoherent diversity receivers over Nakagami fading channels. Also in [53], the performance of the optimal receiver and a suboptimal square-law combining receiver was analyzed. We employ a square-law combining receiver for noncoherent communications over Nakagami block fading channels. The outputs of the square-law combiner are given by (2.5) and the suboptimal decoding metric used is given by (2.6). Thus the conditional pairwise error probability is found by substituting the metric (2.6) in (3.9) resulting in (3.28). For Nakagami fading, the variables $r_f^{(I,0)}$ and $r_f^{(Q,0)}$ conditioned on the fading gains are Gaussian random variables with means $\sqrt{E_s}a_f \cos(\theta_f)$ and $\sqrt{E_s}a_f \sin(\theta_f)$, respectively and variance equal to $\frac{1}{2}N_0$. Similarly, $r_f^{(I,1)}$ and $r_f^{(Q,1)}$ are Gaussian random variables $\mathcal{N}(0, \frac{1}{2}N_0)$. Following the derivation of the conditional Chernoff bound in Section 3.2.4, we arrive at the Chernoff bound (3.30). Using the identity (4.10) with the means and variances of $r_f^{(0)}$ and $r_f^{(1)}$ mentioned above, the Chernoff bound simplifies to

$$P_u(d|\mathbf{f}) \leq \prod_{f=1}^{L} \frac{1}{(1-\lambda d_f N_0)(1+\lambda d_f N_0)} \mathrm{E}_a \left[\exp\left(\frac{-\lambda d_f E_s a_f^2}{1+\lambda d_f N_0} \right) \right]. \tag{4.15}$$

Averaging over the fading gains with density in (4.12) yields

$$P_u(d|\mathbf{f}) \leq \prod_{v=1}^{w} \frac{(1+\lambda v E_s N_0)^{(M-1)f_v}}{(1-\lambda v E_s N_0)^{f_v}} \left(\frac{1}{1+\lambda v (N_0 + E_s/M)} \right)^{M f_v}. \tag{4.16}$$

This bound simplifies to the Chernoff bound for Rayleigh channels when $M = 1$ [35]. Since the expression in (4.16) is difficult to be optimized over λ, numerical optimization is used to evaluate the bound for noncoherent detection of binary codes over block fading channels. The results of a convolutionally encoded system with BFSK signaling and a square-law combining receiver are shown in Figure 4.12. The loss in performance due to the reduced diversity in the channel is clear and increases as the fading parameter M is increased.

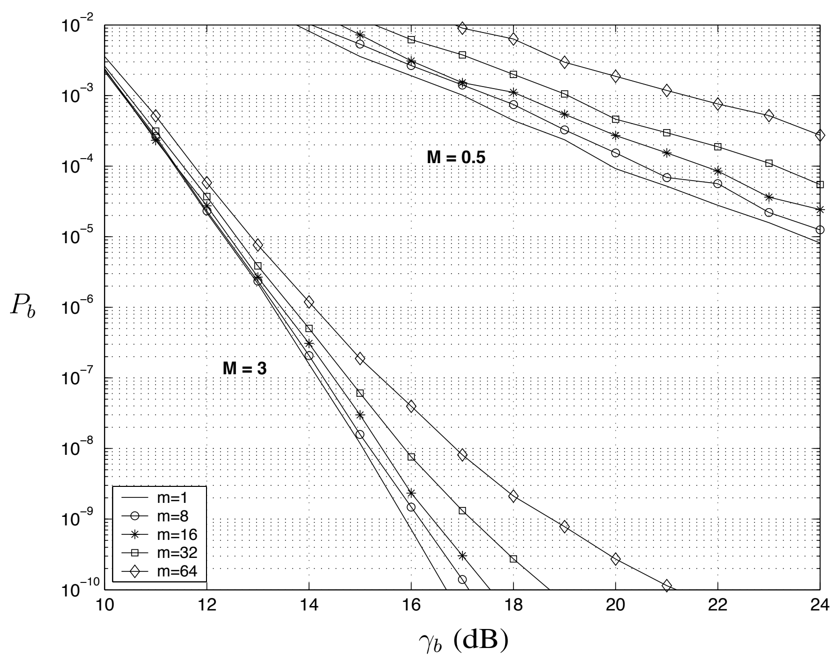

Figure 4.12: Approximation of the bit error probability of a rate-$\frac{1}{2}$ (5,7) convolutional code over Nakagami fading channel with Nakagami parameter $M = 0.5, 3$, noncoherent detection and a frame size $N = 1024$ for memory lengths $m = 1, 8, 16, 32, 64$.

CHAPTER 5

Iterative Joint Decoding and Channel Estimation Receiver for Multi-Antenna Systems

In this chapter, an iterative receiver for joint decoding and channel estimation of multi-antenna systems over block fading channels is derived. As indicated in [23] and illustrated throughout the results of Chapter 3, the performance of multi-antenna systems is severely affected by the quality of channel estimation, especially when large number of transmit antennas are used. The use of orthogonal pilot sequence insertion to estimate the channel was proposed in [23]. It was shown that the estimation quality improves as the length of the pilot sequences is increased, reducing the effective rate and energy efficiency of the system. This emphasizes the need of channel estimation techniques that provide good quality estimation at the least sacrifice in the offered bandwidth and energy.

Iterative receivers that jointly decode and estimate the channel for ST codes have appeared recently. In [28] Li *et al.* implemented a pilot-aided channel estimation scheme for ST coded orthogonal frequency-division multiplexing (OFDM) systems over correlated fading channels. The channel estimation was performed using a least-square estimator with no information exchanged between the decoder and the estimator. An iterative version of the receiver that employs expectation-maximization (EM) algorithm was proposed in [29]. The complexity of the EM algorithm becomes prohibitive for turbo-like codes, and hence a maximum aposteriori version of the EM algorithm was derived in [30]. Grant [31] proposed a pilot-aided iterative receiver for joint decoding and channel estimation of trellis ST codes over quasi-static channels. The receiver was used for turbo ST codes in [32]. Special code structures were utilized to solve the same problem as in [56] for diagonal ST codes and [57] for orthogonal block ST codes. All the above iterative receivers use hard decisions from the decoder to update the channel estimation. However, using the soft information out of the decoder is expected to improve the quality of channel estimation update.

For single-antenna systems, an iterative receiver for decoding and channel estimation of turbo-coded FH systems was proposed in [24]. In this receiver, a quantized representation of the fading process is used and the aposteriori probability of each quantization level is updated iteratively using soft information from the decoder. This algorithm can be extended for the case when the channel is modeled as a Markov process, which was applied in [58] to the phase process.

In this chapter, a pilot-aided iterative receiver for joint decoding and channel estimation of ST codes over block fading channels is proposed. Orthogonal pilot sequence insertion is used to get initial channel estimation for the first iteration. Then, the fading quantization level with the largest aposteriori probability among a "selected" set of levels is found using soft information from the decoder. This fading level is used in subsequent decoding iterations. The receiver iterates between decoding and updating channel SI for a number of iterations. To illustrate the results, trellis and turbo ST codes are used. For these codes, the effect of different parameters of the receiver is studied as well as the convergence behavior of the receiver. Moreover, we find the optimal channel memory via simulation. Also, the affect of the error correcting capability of the code on the optimal channel memory is investigated.

The chapter is organized as follows. In Section 5.1, the proposed iterative receiver is derived. Section 5.2 discusses the results which illustrates the convergence behavior of the iterative receiver as well as the effect of different systems' parameters on the performance of the receiver. Also, the tradeoff between effective diversity and channel estimation is investigated.

5.1 The Iterative Receiver

The proposed iterative receiver uses orthogonal pilot sequence insertion to estimate the channel initially. The use of orthogonal pilot sequence insertion in multi-antenna systems was discussed in Section 3.3.2. From this, a channel estimation is obtained with an error variance $\sigma_e^2 = \frac{N_0}{n_t E_s}$ and a correlation coefficient μ given by (3.40). Thus using longer pilot sequences improves the channel estimation at the cost of reducing the effective energy of the coded system. Therefore, using the soft information of the decoder to aid in channel estimation may reduce the penalty paid in energy for good channel estimation quality.

The block diagram of the proposed iterative receiver is shown in Figure 5.1, where an additional block to update the channel SI was added to the diagram in Figure 2.6. Denote the initial channel estimates obtained from orthogonal pilot sequence insertion by $\hat{\mathbf{H}} = \{\hat{\mathbf{h}}_f\}_{f=1}^{F}$. As in Chapter 3, a suboptimal decoding rule that maximizes the likelihood function $p(\mathbf{Y}|\hat{\mathbf{H}}, \mathbf{S})$ is employed due to its simplicity. It chooses the codeword $\mathbf{S}$ that maximizes the metric

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = -\sum_{f=1}^{F}\sum_{l=1}^{m}\left|y_{f,l} - \frac{\mu}{\sigma}\sqrt{E_s}\sum_{i=1}^{n_t}\hat{h}_f^i s_{f,l}^i\right|^2. \tag{5.1}$$

The distribution of $y_{f,l}$ conditioned on $\hat{\mathbf{h}}_f$ and $\mathbf{s}_{f,l}$ is complex Gaussian with mean $\frac{\mu}{\sigma}\sqrt{E_s}\sum_{i=1}^{n_t}\hat{h}_f^i s_{f,l}^i$ and variance $N_0 + (1-\mu^2)E_s\sum_{i=1}^{n_t}|s_{f,l}^i|^2$. For constant-energy constellations, $\sum_{i=1}^{n_t}|s_{f,l}^i|^2 = n_t$. In the first iteration, initial channel SI obtained from pilot-aided channel estimation is used by the ST decoder to compute the aposteriori probabilities of signal vectors using (2.22). In subsequent iterations, these probabilities are used in the SI update block to improve

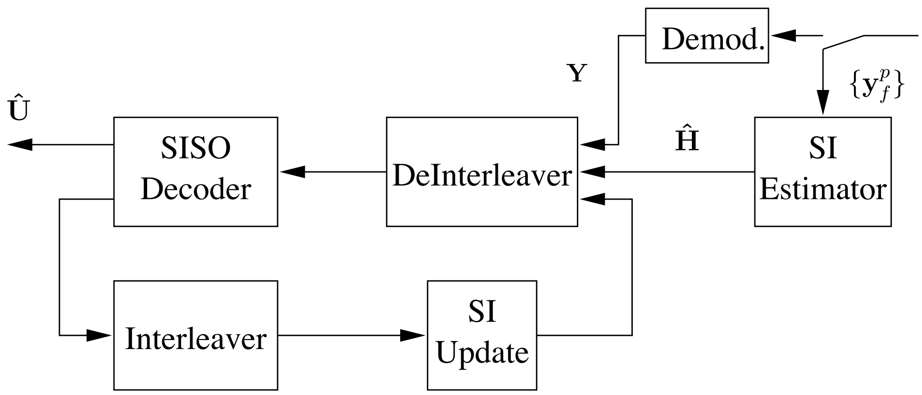

Figure 5.1: The structure of the iterative receiver.

the quality of channel estimation. The new SI are fed to the ST SISO decoder to update the probabilities of the signal vectors in the frame. The process of decoding and updating SI continues for a number of iterations. Updating the channel estimation is accomplished by using a quantized representation of the fading process as follows.

The channel gain from the i^{th} transmit antenna in fading block f is written as $h_f^i = a_f^i \, e^{j\theta_f^i}$, where $j = \sqrt{-1}$, a_f^i and θ_f^i are the amplitude and the phase of h_f^i, which are modeled as Rayleigh and uniform random variables, respectively. In the SI update block, the domain of the amplitude random variable, associated with the channel from each transmit antenna is quantized into L intervals $\{A_j\}_{j=1}^{L}$. The same is performed for the phase random variable resulting in $\{\Theta_j\}_{j=1}^{L}$, and a total of L^{2n_t} possible fading intervals $\{\Phi_j\}_{j=1}^{L^{2n_t}}$ at each receive branch. The center of the quantization interval Φ_j is denoted by g_j. Since phase modulation is used, the phase of the received signal in (2.19) is the cumulative phase of all $s_{f,l}^i$ and h_f^i, making blind estimation impossible in this context.

In order to update the channel estimates, the SI update block computes $p(h_f^i \in \Phi_j|\mathbf{Y})$ for "selected" quantization intervals around the estimated gain. Let $\Phi_f^i = \{\Phi_j\}_{j=1}^{L_c}$ denote the set containing the closest L_c quantization intervals to the channel estimate $\hat{h}_f^i$, i.e., $\Phi_f^i = \{\Phi_j : |g_j - \hat{h}_f^i| \leq d_{L_c}\}$, where d_{L_c} is the distance from the channel estimate to its L_c^{th} nearest neighbor. Thus after each decoding iteration, the SI update block uses the soft information of the decoder to compute the aposteriori probabilities $p(h_f^i \in \Phi_j|\mathbf{Y})$ for quantization levels inside Φ_f^i and updates the channel SI as follows

$$\hat{h}_f^i = g_j, \quad \text{if} \quad p(h_f^i \in \Phi_j|\mathbf{Y}) \geq p(h_f^i \in \Phi_q|\mathbf{Y}), \quad \forall \Phi_q \in \Phi_f^i. \tag{5.2}$$

The soft information of the decoder is used to compute $p(h_f^i \in \Phi_j|\mathbf{Y})$ in the SI update block as follows.

Let $\tilde{\mathbf{Y}}_f$ be the vector containing the channel observations in all fading blocks in the frame except the fading block f. Also, let $\mathbf{Y}_f$ be the vector containing the channel observations in the fading block f. The aposteriori probabilities $p(h_f^i \in \Phi_j|\mathbf{Y})$ are computed as

$$\begin{aligned} p(h_f^i \in \Phi_j|\mathbf{Y}) &= p(\mathbf{Y}|h_f^i \in \Phi_j)\frac{p(h_f^i \in \Phi_j)}{p(\mathbf{Y})} \\ &= p(\mathbf{Y}_f|h_f^i \in \Phi_j, \tilde{\mathbf{Y}}_f)p(\tilde{\mathbf{Y}}_f|h_f^i \in \Phi_j)\frac{p(h_f^i \in \Phi_j)}{p(\mathbf{Y})} \\ &\approx Cp(\mathbf{Y}_f|h_f^i \in \Phi_j)p(h_f^i \in \Phi_j), \end{aligned} \tag{5.3}$$

where $C = p(\tilde{\mathbf{Y}}_f|h_f^i \in \Phi_j)/p(\mathbf{Y})$ is a normalization constant and $p(h_f^i \in \Phi_j)$ is the apriori probability of the fading level. The approximation in (5.3) is due to the fact that channel outputs in different fading blocks are slightly correlated because of the interleaving used in the transmitter. Now, the probability $p(\mathbf{Y}_f|h_f^i \in \Phi_j)$ is written as

$$p(\mathbf{Y}_f|h_f^i \in \Phi_j) \approx \sum_{\mathbf{s}_{f,1},\ldots,\mathbf{s}_{f,m}} p(\mathbf{Y}_f|h_f^i \in \Phi_j, \tilde{\mathbf{h}}_f^i, \mathbf{s}_{f,1}, \ldots, \mathbf{s}_{f,m}).p(\mathbf{s}_{f,1})\ldots p(\mathbf{s}_{f,m}), \tag{5.4}$$

where $\tilde{\mathbf{h}}_f^i$ is the column vector containing the channel estimates of fading block f at all transmit antennas except the i^{th} one. If the size of the signal constellation used at each antenna is M, then computing (5.4) involves summing over M^{mn_t} quantities which is very complex. An iterative algorithm to calculate (5.4) efficiently was derived in [24] for blind estimation in single-antenna FH systems. A modified version of the algorithm that fits the pilot-aided receiver is given in the following

1. Initialization:

$$p(h_f^i \in \Phi_j | y_{f,1}) = p(y_{f,1} | h_f^i \in \Phi_j) \frac{p(h_f^i \in \Phi_j)}{p(y_{f,1})} \tag{5.5}$$

$$p(y_{f,l} | h_f^i \in \Phi_j) \approx \sum_{\forall \mathbf{s}} p(y_{f,l} | h_f^i \in \Phi_j, \tilde{\mathbf{h}}_f^i, \mathbf{s}_{f,l}) p(\mathbf{s}_{f,l}) \tag{5.6}$$

2. Recursion:

$$\begin{aligned} p(h_f^i \in \Phi_j | y_{f,1}, \dots, y_{f,l}) &= \frac{p(y_{f,l} | h_f^i \in \Phi_j, y_{f,1}, \dots, y_{f,l-1})}{p(y_{f,l} | y_{f,1}, \dots, y_{f,l-1})} . p(h_f^i \in \Phi_j | y_{f,1}, \dots, y_{f,l-1}) \\ &\approx \frac{p(y_{f,l} | h_f^i \in \Phi_j) p(h_f^i \in \Phi_j | y_{f,1}, \dots, y_{f,l-1})}{p(y_{f,l} | y_{f,1}, \dots, y_{f,l-1})}. \end{aligned} \tag{5.7}$$

In (5.7), $p(y_{f,l} | h_f^i \in \Phi_j)$ is computed as in (5.6) and the approximation in (5.7) is due to the small correlation between the channel outputs in different fading blocks because of the interleaving used. Also, (5.6) is modified from the original receiver in [24], where the approximation is due to conditioning on $\tilde{\mathbf{h}}_f^i$ instead of averaging over all quantization levels. Note that computing (5.6) requires summing over M^{n_t} signals resulting in exponential complexity in the number of transmit antennas. The results of using the iterative receiver in ST coded systems as well as convolutionally coded STBCs are presented next.

5.2 Results

In this section, the iterative receiver is applied to QPSK trellis and turbo ST codes presented in Chapter 2. The trellis code [33] has 4 states and uses two transmit antennas with a throughput of 2 bits/s. On the other hand, turbo ST codes [39] with two and four transmit antennas are used. The 2-antennas code has 4 states and a throughput of 1 bits/s, where the 4-antenna code has 16 states with a throughput of 2 bits/s. Using these codes, we investigate how the receiver performance is affected by changing its parameters. In the proceeding subsections, we consider the effect of the number of quantization levels L, number of iterations, channel memory length m, frame size N and the number of transmit antennas n_t.

The performance of trellis and turbo ST coded systems is considered with perfect SI, an OPE receiver and joint decoding and channel estimation (JDE) using the iterative receiver described above. In addition, we consider the hypothetical case of feeding the transmitted codeword to the SI update block which results in the CDE assumption. In the cases of CDE, JDE and OPE, the length of the pilot sequences is set to the minimum value needed to preserve orthogonality, i.e., $n_p = n_t$ symbols. As a result, the initial channel estimation has an error with variance $\sigma_e^2 = \frac{N_0}{n_t E_s}$ for JDE and OPE receivers. By assumption, $\sigma_e^2 = \frac{N_0}{mE_s}$ for the CDE case. Note that the pilot energy is taken into account in the cases of the CDE, JDE and OPE, where no pilot sequences are inserted in the case of perfect SI. Throughout the simulation, the amplitude and phase processes are quantized using a Llyod-Max algorithm [7], and the aposteriori probabilities are computed in SI update block for $L_c = 4$ quantization levels. Also, all results are expressed as a function of the SNR per information bit given by $\gamma_b = \frac{E_s}{kN_0}$, where k is the number of information bits transmitted during a transmission interval of length T.

5.2.1 Number of Quantization Levels

In Figure 5.2, the required SNR γ_b for trellis and turbo ST codes to achieve a bit error rate (BER) of 10^{-3} is shown. Also shown in the figures the required SNR for the OPE and CDE cases. We see that for trellis codes, increasing the number of quantization levels from $L = 8$ to $L = 32$ provides a gain of 0.5 dB, and increasing L beyond 32 have a negligible effect. The same observation holds for turbo codes. However, the gain of the iterative receiver over the OPE

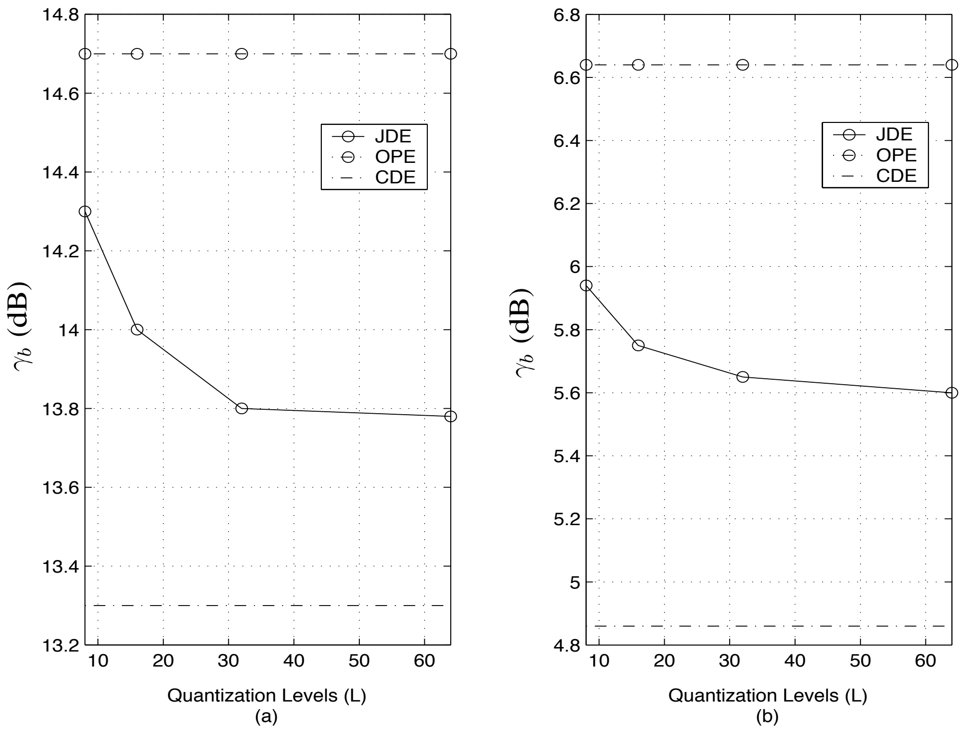

Figure 5.2: SNR required for QPSK ST codes with the iterative receiver to achieve $P_b = 10^{-3}$ versus the number of quantization levels for frame size $N = 1024$, number of transmit antennas $n_t = 2$ and channel memory length $m = 16$. (a) trellis, (b) turbo.

case is larger in turbo codes than in trellis codes. This is expected since turbo codes are more powerful than trellis codes and hence soft information of a turbo decoder is more reliable than that provided by a trellis decoder.

5.2.2 Number of Iterations

The effect of the number of iterations on the performance of the iterative receiver is shown in Figure 5.3. Figure 5.3a shows the required SNR to achieve a BER of 10^{-3}, from which we conclude that 3 iterations are enough for the receiver to converge and iterating more does not improve the performance. On the other hand, the performance of turbo codes is highly affected by the number of iterations. Hence, another parameter should be used to measure the progress of estimation quality with the number of iterations. We chosen the channel gain-to-estimation ratio (CER) to measure the estimation quality at

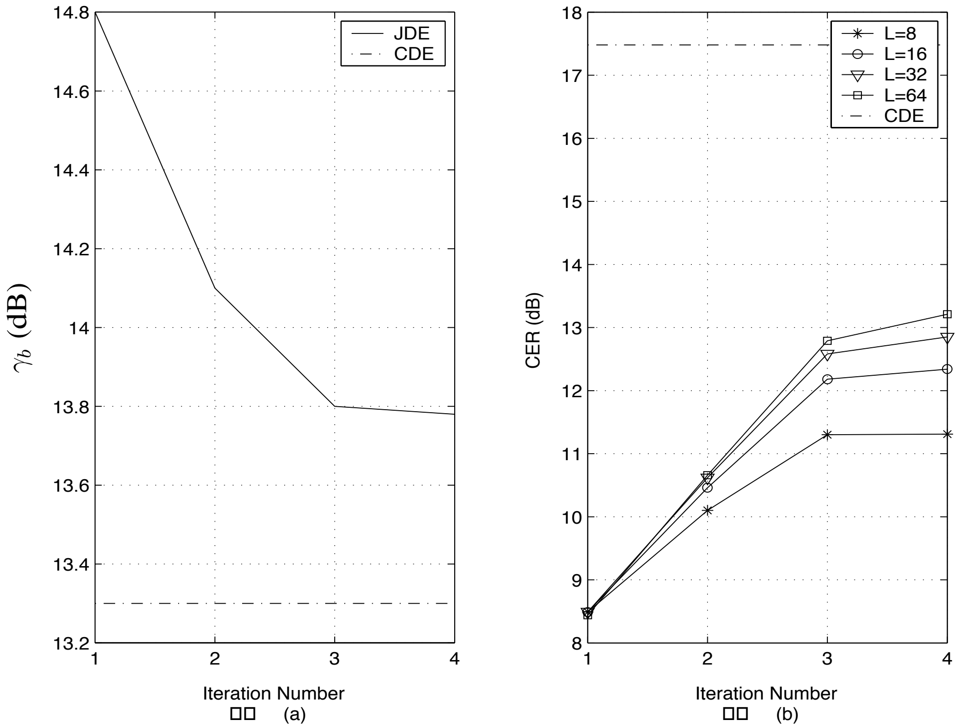

Figure 5.3: Effect of the number of iterations on the performance of the iterative receiver for QPSK ST codes with frame size $N = 1024$, number of transmit antennas $n_t = 2$ and channel memory length $m = 16$. (a) SNR required for the trellis code to achieve $P_b = 10^{-3}$ with $L = 64$, (b) CER = $-\log(\sigma_e^2)$ versus the number of iterations for the turbo code parameterized by the number of quantization levels.

each iteration, which is defined as CER $= -\log(\sigma_e^2)$. In Figure 5.3b, the CER is plotted versus the number of iterations for different quantization levels. Observe that the CER is almost unchanged after 3 iterations showing that it is enough to use the SI update block in 3 iterations only in turbo codes, instead of using it in all turbo iterations.

Observe that the CER for the CDE assumption is 4 dB better than that of the JDE using the iterative receiver. However, from Figure 5.2 the SNR required by the CDE assumption to achieve BER of 10^{-3} is only 0.7 dB less than that required by the iterative receiver with $L = 64$. Thus reducing the channel estimation variance may not provide a proportional gain in the required SNR, which is the important performance measure in communication systems.

5.2.3 Channel Memory Length

Figure 5.4 shows the performance of trellis and turbo ST codes with frame size $N = 1024$, two transmit antennas $n_t = 2$ and channel memory of length $m = 16$. For trellis codes, using the iterative receiver provides less than 1 dB gain over the OPE receiver and is less than 0.5 dB worse than the CDE assumption at BER of 10^{-3}. In turbo codes, the iterative receiver is better than the OPE receiver by slightly more than a dB and is worse than the CDE assumption by almost 0.75 dB. Also, it is observed that turbo ST codes are more sensitive to channel estimation errors than trellis ST codes. This is clear from the SNR loss of the OPE receiver and the CDE assumption with respect to the case of perfect SI. The same information are shown in Figure 5.5 for channel memory length of $m = 128$. In this case, the iterative receiver is provides most of the gain that is achieved by the CDE assumption. This is due mainly to the long channel memory which enhance the iterative estimation quality. However, the iterative receiver is closer to the cases of perfect SI and the CDE assumption in turbo codes than in trellis codes because turbo codes are more powerful codes.

5.2.4 Frame Size

The performance of turbo ST codes with frame size $N = 4096$, two transmit antennas $n_t = 2$ and channel memory of length $m = 64$ is shown in Figure 5.6. In this case, the number of independent fading blocks is the same as the case discussed above, i.e., when $N = 1024$ and $m = 16$. We see that the iterative receiver provides a gain of 2 dB over the OPE receiver and is worse than the CDE assumption by around 0.25 dB. This shows that the performance of the iterative receiver improves as the frame length is increased. Also, note that the code sensitivity to channel estimation errors increases with the frame size. A similar observation was noted in [27], where it was shown that channel estimation becomes more crucial to the performance as the code approaches the

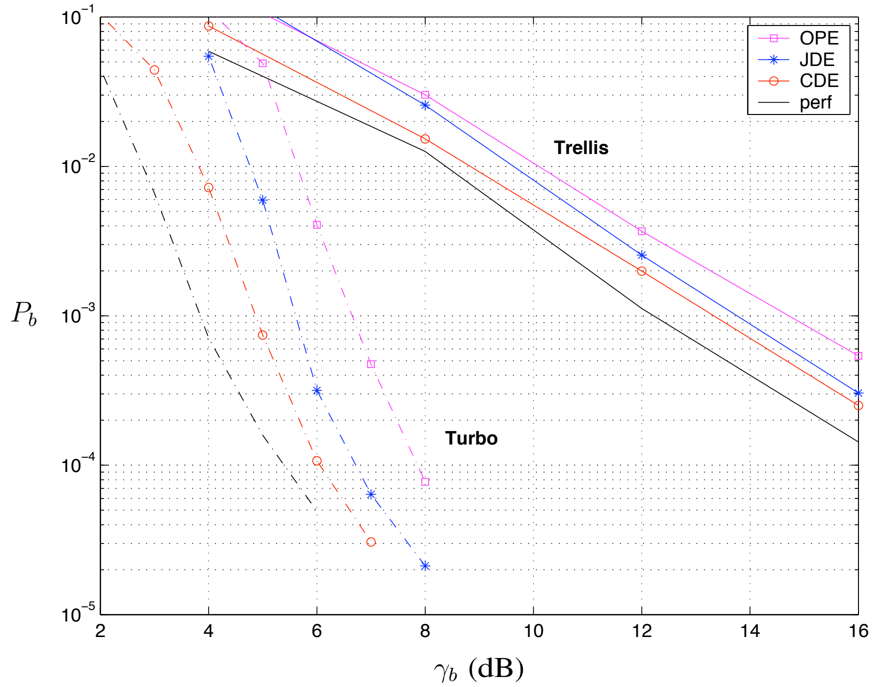

Figure 5.4: Performance of QPSK trellis and turbo ST codes for frame size $N = 1024$, number of transmit antennas $n_t = 2$ and memory length $m = 16$.

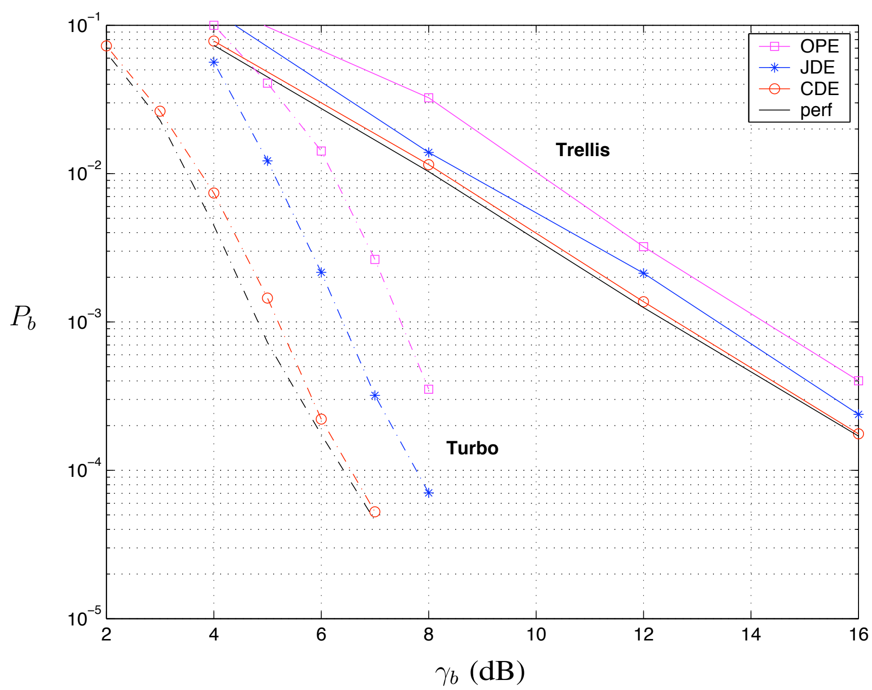

Figure 5.5: Performance of QPSK trellis and turbo ST codes for frame size $N = 1024$, number of transmit antennas $n_t = 2$ and memory length $m = 128$.

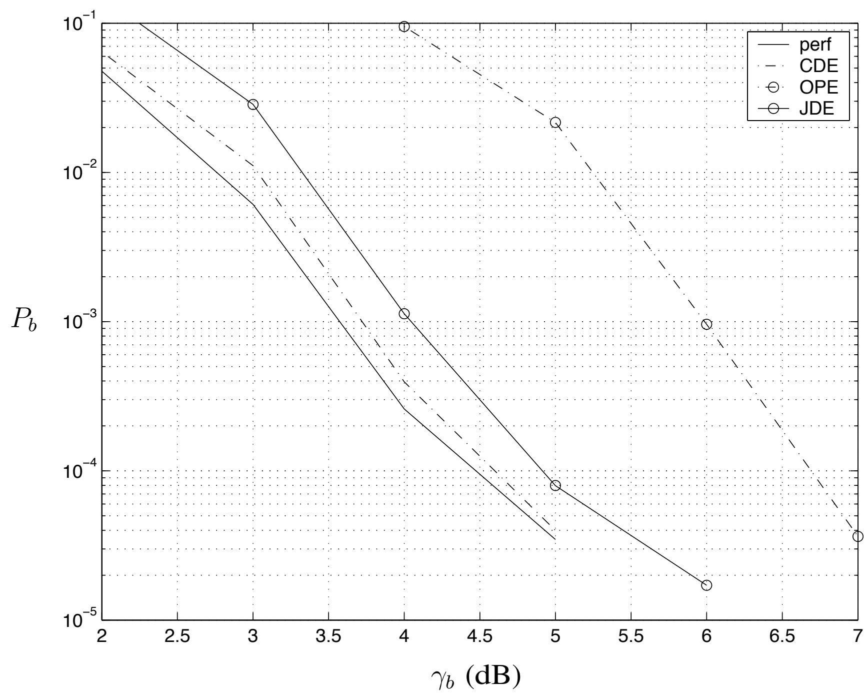

Figure 5.6: Performance of QPSK turbo ST code for frame size $N = 4096$, number of transmit antennas $n_t = 2$ and channel memory length $m = 64$.

capacity, which is the case as the frame size increase in turbo codes.

5.2.5 Number of Transmit Antennas

The performance of turbo ST codes with frame size $N = 1024$, four transmit antennas $n_t = 4$ and channel memory of length $m = 64$ is shown in Figure 5.7. The iterative receiver provides a gain of 1 dB over the OPE receiver and is worse than the CDE assumption by around 1 dB at BER of 10^{-3}. We observe that the code sensitivity to channel estimation errors increases with the number of transmit antennas as was shown in [23]. The performance of the iterative receiver improves and becomes closer to the CDE assumption as the SNR increases. This concludes that the iterative receiver is expected to perform well for large number of transmit antennas.

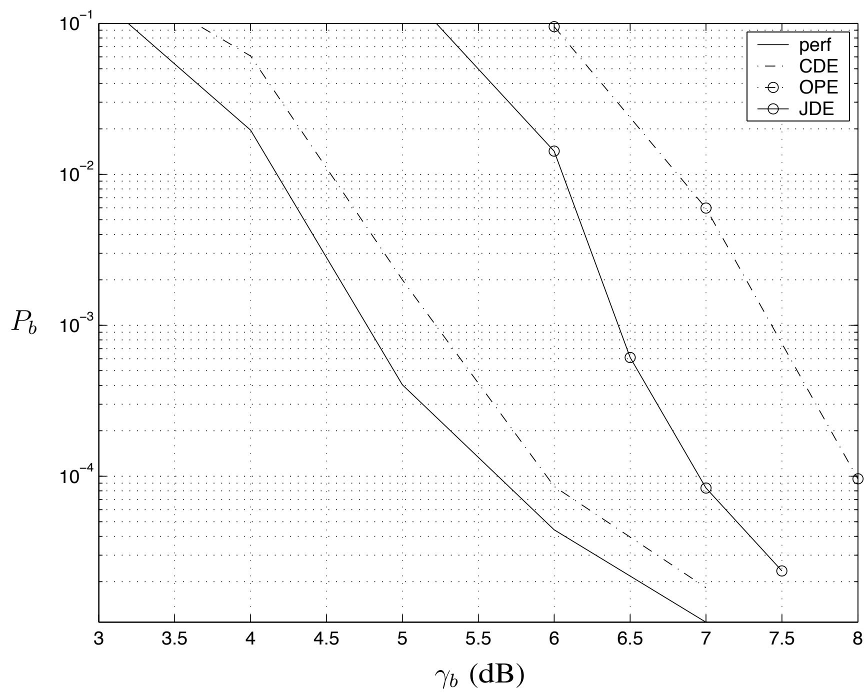

Figure 5.7: Performance of QPSK turbo ST code for frame size $N = 1024$, number of transmit antennas $n_t = 4$ and channel memory length $m = 64$.

5.2.6 Optimal Channel Memory

The SNR required to achieve a BER of 10^{-3} is plotted versus the channel memory length m for trellis and turbo ST codes in Figures 5.8 and 5.9, respectively. Note that the frame size is $N = 1024$ and the number of transmit antennas is $n_t = 2$. The cases of perfect SI, CDE and the use of the iterative receiver are shown in the figure. In the case of perfect SI, effective diversity is the key performance criteria and hence the performance improves as the channel memory length is reduced, i.e., when the effective diversity is increased. However, in the CDE assumption and the iterative receiver, the performance is affected jointly by the effective diversity, fraction of pilot energy and channel estimation quality. From the figures, the optimal memory is around $m = 64$ symbols for the trellis code, and around $m = 32$ symbols for the turbo ST code. Therefore, the turbo code has a shorter optimal memory than the trellis code because the former is more sensitive to diversity, and its performance is more affected by increasing the channel memory length.

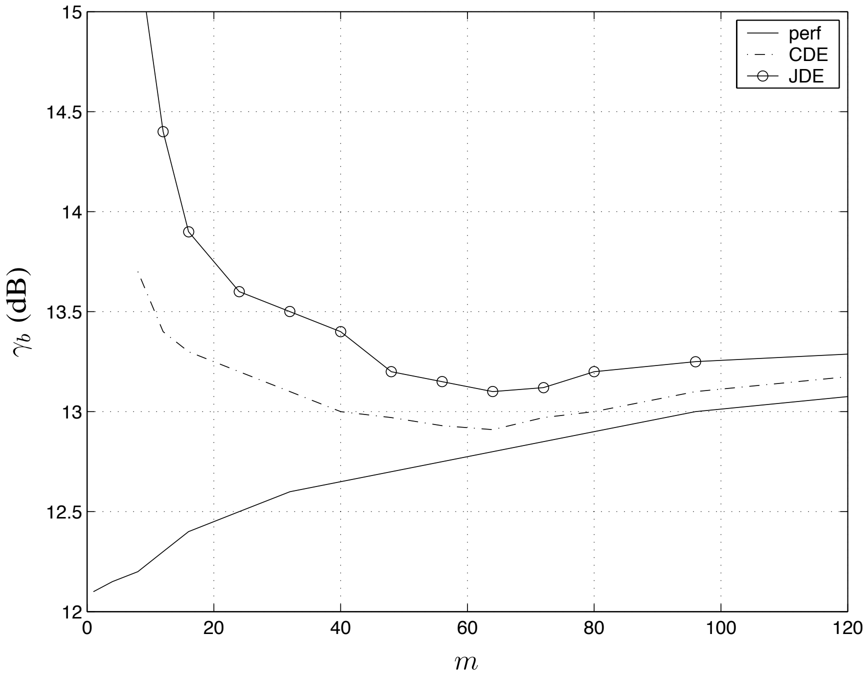

Figure 5.8: SNR required for QPSK ST trellis code to achieve $P_b = 10^{-3}$ versus channel memory m for frame size $N = 1024$ and number of transmit antennas $n_t = 2$.

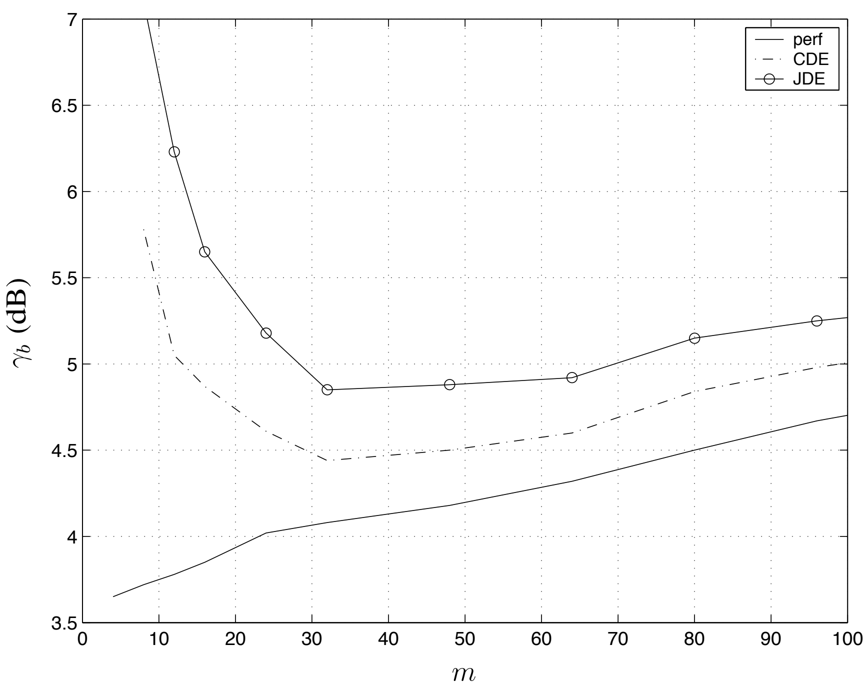

Figure 5.9: SNR required for QPSK ST turbo code to achieve $P_b = 10^{-3}$ versus channel memory m for frame size $N = 1024$ and number of transmit antennas $n_t = 2$.

CHAPTER 6

I-Q Space-Time Coded Systems

In this chapter, the performance of I-Q ST codes is analyzed and two efficient iterative receivers are derived. Trellis codes are good candidates for applications that require low-complexity receivers and short delays. If a trellis code is used over a block fading channel and the coded sequence is interleaved to break up the channel memory, then for low to medium SNR values, the channel can be approximated by a memoryless channel provided that the number of fading blocks is several times larger than the code constraint length. For this observation and due to the difficulty of optimizing trellis codes for block fading channels, various ST trellis codes were optimized for independent fading channels in [33,38,59]. In [60] Tonello presented a bit-interleaved ST coded scheme for independent fading channels. Motivated by the low complexity and short delays of trellis codes, this chapter is devoted to a class of trellis ST codes that provide large time diversity known as I-Q ST codes.

It is known that the time diversity provided by a trellis code is the key design parameter of the code for independent fading channels. Conventionally, increasing the time diversity is achieved either by reducing the number of input bits to the encoder which reduces the throughput, or increasing the memory of the encoder which increases the complexity. In [61], Al-Semari *et al.* showed that using I-Q encoding can increase the time diversity of the code with the same throughput and lower decoding complexity compared to conventional coding techniques. In I-Q encoding, the input stream is encoded using two independent encoders and the output of each encoder is used to determine one dimension of the complex signal constellation, i.e., the I and Q dimensions. This reduces the number of input bits to each encoder to half the original number, which increases the time diversity with less decoding complexity compared to conventional coding techniques. Motivated by the large time diversity of I-Q codes, Zummo *et al.* [33] observed that when I-Q encoding is used in ST coded systems the I and Q coded sequences are faded and super-imposed on top of each other. Therefore, a "super-trellis" corresponding to the product of the trellises of the component codes is necessary for performance evaluation and decoding, which has two drawbacks.

The first drawback of using the super-trellis is that the performance criteria of I-Q ST codes are expressed as functions of the parameters of the super-trellis, which makes it difficult to optimize and design the component codes without the need of the super-trellis. In [19], the pairwise error probability of ST codes was upper bounded, and performance criteria were derived for the case of perfect SI. Similar results for the case of imperfect SI were obtained in [23]. In this chapter, the pairwise error probability as well as the performance criteria of I-Q ST codes are expressed as functions of the parameters of the component codes. Both cases of perfect and imperfect SI are considered.

The second drawback of the super-trellis is its huge complexity for practical code constraint lengths. For example, if each component code had 32 state, the super-trellis would have $32 \times 32 = 1024$ states. A suboptimal decoding algorithm was proposed in [33], which is based on a detection stage prior to the I and Q decoders without exchanging information between the decoders. The performance of the algorithm was not satisfactory especially for non-constant energy signal constellations. In this chapter, two iterative decoding algorithms are proposed. The first algorithm views I-Q ST codes as a concatenation of a channel code and a mapping process, and hence it uses iterative demodulation and decoding (IDD). The second algorithm uses interference cancellation techniques, and is referred to as interference cancellation decoder (ICD). Results show that using 3 iterations in both algorithms provides performance close to optimal decoding. From the complexity point of view, the ICD has a lower complexity than the IDD at the cost of performance degradation, where both algorithms have much lower complexities than optimal decoding.

This chapter is organized as follows. In Section 6.1, the general model of an I-Q ST coded system is described. The performance of I-Q ST codes and their design criteria are derived in Section 6.2. In Section 6.3, the geometrical uniformity of I-Q ST codes is established from the geometrical uniformity component codes. Two iterative receivers are presented in 6.4 as well as their performance and complexity issues.

6.1 System Description

In this section, we describe the transmitter, channel and receiver of I-Q ST coded systems. The general I-Q ST transmitter is shown in Figure 6.1. During a frame transmission period of length NT, the input to the transmitter is a length-N sequence of binary vectors $\{\mathbf{u}_l\}_{l=1}^{N}$ each of length k bits. At time lT the input to the transmitter is $\mathbf{u}_l$ and the corresponding output is the signal vector $\mathbf{s}_l$ of length n_t. The signal vector is transmitted using n_t antennas, one for each component of $\mathbf{s}_l$. The resulting throughput is k/T bits/s. The transmitter first splits each input vector into two equal-length vectors $\mathbf{u}_{I,l}$ and $\mathbf{u}_{Q,l}$. The vector $\mathbf{u}_{I,l}$ is encoded by the I encoder into a signal vector $\mathbf{s}_{I,l}$ of length n_t, whose elements are drawn from an alphabet $\mathcal{A}_I$ which is a 1-D constellation such as M-PAM. The 1-D signals in $\mathbf{s}_{I,l}$ constitute the I components of the 2-D signals to be transmitted over the n_t transmit antennas. The same applies to the Q branch, resulting in two length-N sequences of 1-D signal vectors $\mathbf{S}_I = \{\mathbf{s}_{I,l}\}_{l=1}^{N}$ and $\mathbf{S}_Q = \{\mathbf{s}_{Q,l}\}_{l=1}^{N}$. After that, the I and Q codewords $\mathbf{S}_I$ and $\mathbf{S}_Q$ are interleaved. After interleaving, the 2-D signal s_l^i to be transmitted over the i^{th} antenna is drawn from the alphabet $\mathcal{A} = \mathcal{A}_I \times \mathcal{A}_Q$ according to: $s_l^i = s_{I,l}^i + js_{Q,l}^i$, where

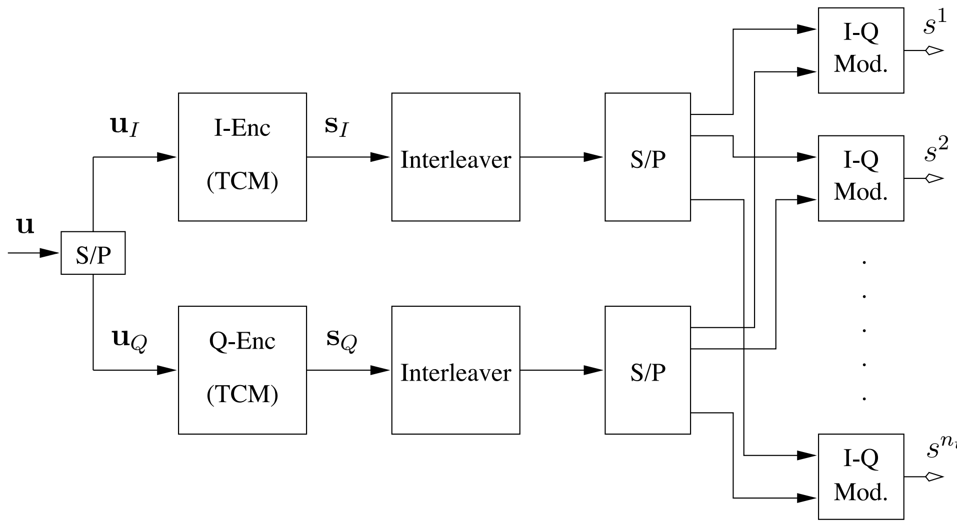

Figure 6.1: I-Q ST transmitter structure

$j = \sqrt{-1}$, resulting in a transmitted codeword $\mathbf{S} = \mathbf{S}_I + j\mathbf{S}_Q$.

The channel considered here is an independent fading channel, where the received signal at time l is given by

$$y_l = \sqrt{E_s} \sum_{i=1}^{n_t} h_l^i s_l^i + z_l, \tag{6.1}$$

where in this case h_l^i is the channel gain from the i^{th} transmit antenna at time l and is modeled as $\mathcal{CN}(0, 1)$. Here, the fading affecting each signal in the frame is independent from that affecting the other signals, which results in the independent fading channel model. The receiver employs the ML decoding rule that minimizes the frame error probability by maximizing the metric

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = -\sum_{l=1}^{N} \left| y_l - \sqrt{E_s} \sum_{i=1}^{n_t} h_l^i s_l^i \right|^2. \tag{6.2}$$

The error probability of I-Q ST codes over independent fading channels is discussed in the next section.

6.2 Performance Analysis

In this section, the pairwise error probability and the transfer function of ST codes are reviewed. Next, the bit error probability of I-Q ST codes is derived for the cases of perfect and imperfect SI. Then, the geometrical uniformity of I-Q ST codes is proved from the geometrical uniformity of the component codes. The section concludes with analytical results for some I-Q ST codes in the literature. First consider the case of perfect SI available at the receiver. The conditional pairwise error probability [19] is defined as the probability of decoding a received sequence as $\hat{\mathbf{S}}$ given that S was transmitted conditioned on the fading variables $\mathbf{H} = \{\mathbf{h}_l\}_{l=1}^{N}$. It is given by (3.9) where the metrics are from (6.2). Substituting the metrics in (3.9) yields

$$P_c(\mathbf{S}, \hat{\mathbf{S}}, \mathbf{H}) = \Pr\left(\kappa > d_E^2(\mathbf{S}, \hat{\mathbf{S}}) \big| \mathbf{S}, \mathbf{H}\right), \tag{6.3}$$

where

$$d_E^2(\mathbf{S}, \hat{\mathbf{S}}) = E_s \sum_{l=1}^{N} \Big| \sum_{i=1}^{n_t} h_l^i e_l^i \Big|^2, \tag{6.4}$$

$$\kappa = 2\sqrt{E_s}\ \mathrm{Re}\left\{ \sum_{l=1}^{N} z_l^* \sum_{i=1}^{n_t} h_l^i e_l^i \right\}, \tag{6.5}$$

where $e_l^i = s_l^i - \hat{s}_l^i$. In (6.5) the random variable κ has a complex Gaussian $\mathcal{CN}(0, \sigma_\kappa^2)$ distribution where $\sigma_\kappa^2 = 2N_0 E_s \sum_{l=1}^{N} | \sum_{i=1}^{n_t} h_l^i e_l^i |^2$. The unconditional pairwise error probability is found by averaging (6.3) over the channel statistics

$$P_u(\mathbf{S}, \hat{\mathbf{S}}) = \mathrm{E}_{\mathbf{H}}\left[Q\left(\sqrt{\frac{d_E^2(\mathbf{S}, \hat{\mathbf{S}})}{2N_0}} \right) \right] \le \frac{1}{2} \prod_{l=1}^{L} \left(\frac{1}{1 + \frac{d_l}{4N_0}} \right) \sim \frac{1}{2} \prod_{l=1}^{L} \frac{d_l}{4N_0}, \tag{6.6}$$

where $d_l = E_s \sum_{i=1}^{n_t} |e_l^i|^2$, $L = \min \big| \{l : \mathbf{s}_l \ne \hat{\mathbf{s}}_l\} \big|$ is called the *minimum time diversity* of the ST code. Note that the Chernoff bound was used in (6.6). Using the integral expression of the Q-function [45], the unconditional pairwise error probability is written as

$$P_u(\mathbf{S}, \hat{\mathbf{S}}) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{l=1}^{L} \left(\frac{1}{1 + \frac{d_l}{4N_0 \sin^2 \theta}} \right) d\theta. \tag{6.7}$$

6.2.1 Design Parameters

From (6.6) important design parameters for ST codes over independent fading channels are:

1. The *diversity* gain, L.
2. The *coding* gain defined by the *squared product distance*, i.e., $d_p^2 = \prod_{l=1}^{L} d_l$.

The main advantage of I-Q encoding technique appears in increasing the diversity gain L of the ST code. In Table 6.1, a comparison between ST trellis codes employing single encoder and the I-Q encoding scheme from the complexity and diversity gain points of view. The complexity is defined as the total number of trellis branches divided by the number of information bits [61]. We see that for the same throughput and decoding complexity, the diversity gain provided by I-Q encoding is larger than those provided by single-encoder ST codes. Furthermore, as the constellation size increases, diversity gains of I-Q ST codes increases significantly, resulting in huge gains over single-encoder ST codes.

By employing the union bound and averaging over all transmitted codewords we obtain an upper bound on the error event probability $P_{e,u}$

$$P_{e,u} \leq \sum_{\mathbf{S}} \sum_{\hat{\mathbf{S}}} P(\mathbf{S}) P_u(\mathbf{S}, \hat{\mathbf{S}}), \tag{6.8}$$

where $P(\mathbf{S})$ is the probability that codeword S is transmitted. If the trellis code is geometrically uniform, then without loss of generality the all-zero codeword

Table 6.1: Comparison of ST codes employing single encoder and I-Q encoding technique.

Constellation	(bits/s)	Code	# States	Diversity gain L	Complexity
QPSK	2	Single	4	2	8
		I-Q	4	3	8
		Single	8	3	16
		I-Q	8	4	16
16-QAM	4	Single	16	2	64
		I-Q	16	4	32
		I-Q	32	5	64
		Single	32	3	128
		I-Q	64	6	128
64-QAM	6	Single	64	2	683
		I-Q	64	5	85
		I-Q	256	7	683

$\mathbf{S_0} = \{\mathrm{s}_0\}_{l=1}^{N}$ can be assumed to be sent and $P_{e,u}$ becomes

$$P_{e,u} \leq \sum_{\hat{\mathrm{S}}} P_u(\mathbf{S_0}, \hat{\mathbf{S}}). \tag{6.9}$$

In order to evaluate (6.9), the transfer function of the code is required. The transfer function of a trellis code enumerates the number of codewords at every input weight and output distance and can be calculated using the error state diagram [62]. A description of the transfer function of a ST code is included in Appendix C.1. Computing (C.2) requires the construction of the state diagram of the super-trellis of the composite code, which may be too complex in general. In addition, designing good I-Q ST codes requires optimizing the *coding* and *diversity* gains of the I-Q ST code by choosing the signal labels of the component codes rather than choosing those of the super-trellis. Hence, it is of practical importance to express the performance of I-Q ST codes as a function of the transfer functions of the component codes. This is discussed in the remainder of this section.

6.2.2 Perfect SI

Define $\mathcal{E} = \{\mathbf{S} \rightarrow \hat{\mathbf{S}}\}$ to be the event of decoding a received sequence as a codeword $\hat{\mathbf{S}}$. The conditional probability of $\mathcal{E}$ given that S was transmitted is denoted by $P_c(\mathcal{E})$ and given by (6.3). Similarly, $\mathcal{I} = \{\mathbf{S}_I \rightarrow \hat{\mathbf{S}}_I\}$ and $\mathcal{Q} = \{\mathbf{S}_Q \rightarrow \hat{\mathbf{S}}_Q\}$ refer to the events that the I and Q decoders choose erroneously the 1-D codewords $\hat{\mathbf{S}}_I$ and $\hat{\mathbf{S}}_Q$, respectively. Since $\mathcal{E}$ occurs if either $\mathcal{I}$ or $\mathcal{Q}$ occurs or both of them, then using the union bound

$$P_c(\mathcal{E}) \leq P_c(\mathcal{I}) + P_c(\mathcal{Q}). \tag{6.10}$$

Our goal is to bound the pairwise error probability in (6.3) as a function of the parameters of the I and Q codes explicitly. Modifying (6.4) by using the complex representation $e_l^i = e_{I,l}^i + je_{Q,l}^i$, then

$$\begin{aligned} d_E^2(\mathbf{S}, \hat{\mathbf{S}}) &= E_s \sum_{l=1}^{N} \Big| \sum_{i=1}^{n_t} h_l^i \big(e_{I,l}^i + je_{Q,l}^i\big) \Big|^2 \\ &= E_s \sum_{l=1}^{N} \left(\Big| \sum_{i=1}^{n_t} h_l^i e_{I,l}^i \Big|^2 + \Big| \sum_{i=1}^{n_t} h_l^i e_{Q,l}^i \Big|^2 \right). \end{aligned} \tag{6.11}$$

Following the derivation in [19] to simplify the distances in (6.11), we have

$$\begin{aligned} d_E^2(\mathbf{S}, \hat{\mathbf{S}}) &= E_s \sum_{l=1}^{N} |\beta_l|^2 \left(d_{I,l} + d_{Q,l}\right) \\ &= d_E^2(\mathbf{S}_I, \hat{\mathbf{S}}_I) + d_E^2(\mathbf{S}_Q, \hat{\mathbf{S}}_Q), \end{aligned} \tag{6.12}$$

where $d_{I,l} = E_s \sum_{i=1}^{n_t} |e_{I,l}|^2$, $d_{Q,l} = E_s \sum_{i=1}^{n_t} |e_{Q,l}|^2$ and $\beta_{j,l}$ follows $\mathcal{CN}(0,1)$ distribution. In (6.12), the squared distance is split into two parts: a part due to the error signal along the I direction and another part due to the error signal along the Q direction. Substituting (6.12) in (6.3), the conditional pairwise error probability becomes

$$P_c(\mathcal{E}) = \Pr\left(\kappa > d_E^2(\mathbf{S}_I, \hat{\mathbf{S}}_I) + d_E^2(\mathbf{S}_Q, \hat{\mathbf{S}}_Q)\Big|\mathbf{S}, \mathbf{H}\right). \tag{6.13}$$

When no error event occurs in the I code, $d_E^2(\mathbf{S}_I, \hat{\mathbf{S}}_I) = 0$, and (6.13) is the probability of an error event in the Q code, i.e., $P_c(\mathcal{Q}) = P_c\left(\kappa_Q > d_E^2(\mathbf{S}_Q, \hat{\mathbf{S}}_Q)|\mathbf{S}_Q, \mathbf{H}\right)$. Here, κ_Q is the noise affecting the Q direction only, which results after removing the contribution of the I error signal $e_{I,l}^i$ from κ. Similarly, $d_E^2(\mathbf{S}_Q, \hat{\mathbf{S}}_Q) = 0$ when no error event occurs in the Q code and the probability in (6.13) is the probability of an error event in the I code, i.e., $P_c(\mathcal{I}) = P_c\left(\kappa_I > d_E^2(\mathbf{S}_I, \hat{\mathbf{S}}_I)|\mathbf{S}_I, \mathbf{H}\right)$. Using (6.10) and averaging over the channel statistics, the unconditional pairwise error probability becomes

$$P_u(\mathcal{E}) \leq P_u(\mathcal{I}) + P_u(\mathcal{Q}). \tag{6.14}$$

Using the Chernoff bound results in

$$P_u(\mathcal{I}) \leq \frac{1}{2}\prod_{l=1}^{L_I}\left(\frac{1}{1+\frac{d_{I,l}}{4N_0}}\right) \sim \frac{1}{2}\prod_{l=1}^{L_I}\frac{d_{I,l}}{4N_0}, \tag{6.15}$$

where $L_I = \min\left|\{l : \mathrm{s}_{I,l} \neq \hat{\mathrm{s}}_{I,l}\}\right|$ is the *minimum time diversity* of the I code. The *coding* gain for the I code is defined as $d_{P,I}^2 = \prod_{l=1}^{L_I} d_{I,l}$. Similar expressions for $P_u(\mathcal{Q})$, L_Q and $d_{P,Q}^2$ hold for the Q code. From (6.15) and (6.14), the design criteria for the component codes of I-Q ST codes for independent Rayleigh fading channels are:

1. In order to have a *diversity* gain of L, then L should satisfy $L = \min(L_I, L_Q)$. This means that the minimum of L_I and L_Q should be set equal to the target *minimum time diversity* of the composite code, i.e., L.

2. In order to have a *coding* gain of d_P^2, then it should satisfy $d_P^2 = \min(d_{P,I}^2, d_{P,Q}^2)$. So, the minimum of $d_{P,I}^2$ and $d_{P,Q}^2$ should be set equal to d_P^2.

The above design criteria are for arbitrary I and Q codes, which in general can be different if unequal error protection or variable transmission rates are desired for the I and Q parts of the signal. However, in the case of identical codes, then the design parameters become $L = L_I = L_Q$ and $d_P^2 = d_{P,I}^2 = d_{P,Q}^2$. Exact expressions for $P_u(\mathcal{I})$ and $P_u(\mathcal{Q})$ are given by (C.2) by replacing $T(J, D_1, .., D_m)$ by $T_I(J, D_1, .., D_m)$ and $T_Q(J, D_1, .., D_m)$ for the I and Q codes, respectively. Also, L and d_l are replaced by the corresponding parameters of the I and Q codes. From (6.14), the bit error probability is bounded by

$$P_b \leq P_{b,I} + P_{b,Q}, \tag{6.16}$$

where $P_{b,I}$ and $P_{b,Q}$ are the bit error probabilities of the I and Q codes, respectively.

Design Example: The I-Q ST code with 16-QAM constellation, 4 bits/s throughput and $n_t = 2$ transmit antennas [33]. The component codes employ a 4-PAM constellation and the 2-D 16-QAM signal space is partitioned as shown in Figure 6.2. Set partitioning is performed such that the normalized squared Euclidean distance $d = \sum_{i=1}^{n_t} |e|^2$ and the squared product distance d_P^2 of the generated subsets are higher each time the partitioning is performed. The trellis diagram of the 4-state code is shown in Figure 6.3b. From the figure, the labels of branches remerging to the same state are chosen from different subsets in the last partitioning level. The resultant *diversity* and *coding* gains are 2 and 3.2, respectively with corresponding error event [0 3;2 3]. Note that this code is not geometrically uniform.

6.2.3 Imperfect SI

The use of orthogonal pilot sequence insertion for channel estimation in ST coded systems was presented in Section 3.3.2. From this, channel estimation is obtained with an error variance $\sigma_e^2 = \frac{N_0}{n_t E_s}$ and a correlation coefficient μ given by (3.40). The suboptimal decoding rule that maximizes the likelihood function

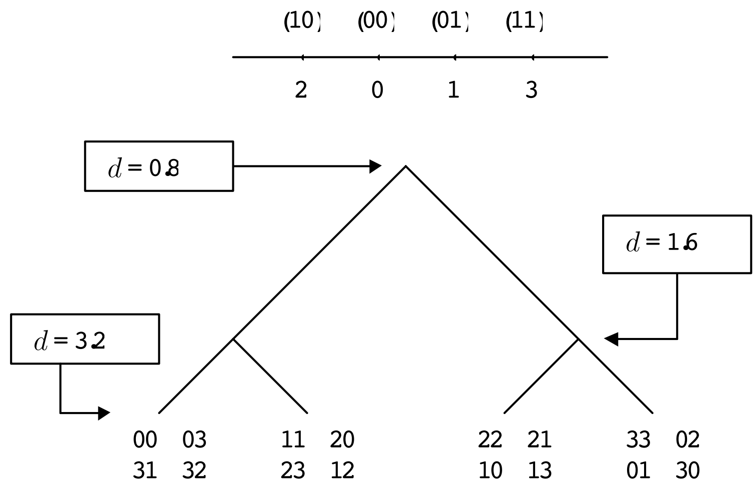

Figure 6.2: 2-D 4-PAM signal space partitioning.

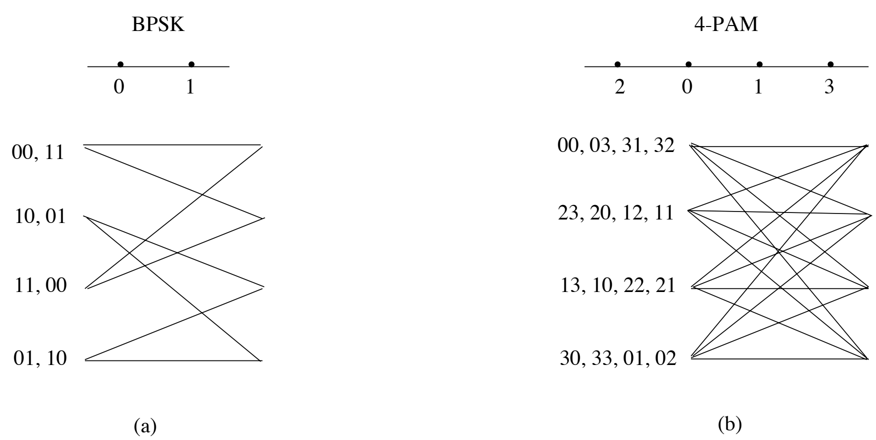

Figure 6.3: The 1-D constellations and trellis diagrams of the 4-state component codes used as an I-Q ST codes with n_t=2 (a) QPSK (2 bits/s) (b) 16-QAM (4 bits/s).

$p(\mathbf{Y}|\hat{\mathbf{H}}, \mathbf{S})$ is employed. it chooses the codeword that maximizes

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = -\sum_{l=1}^{N} \left| y_l - \frac{\mu}{\sigma}\sqrt{E_s}\sum_{i=1}^{n_t} h_l^i s_l^i \right|^2. \tag{6.17}$$

The conditional pairwise error probability given the estimated channel gains is given by substituting $\mathbf{m}(\mathbf{Y}, \mathbf{S})$ in (3.9). Thus, the conditional pairwise error probability simplifies to (6.3) with $d_E^2(\mathbf{S}, \hat{\mathbf{S}})$ is given by (6.4) by replacing h_l^i by ξ_l^i for all l, i, where $\xi_l^i = \hat{h}_l^i/\sigma$ is a random variable with $\mathcal{CN}(0, 1)$ distribution. Here, $\sigma^2 = 1 + \sigma_e^2$. In this case, κ is a complex Gaussian with zero-mean and variance $2\left(N_0 + n_t E_s(1-\mu^2)\right)\mu^2 E_s \sum_{l=1}^{N} \left| \sum_{i=1}^{n_t} \xi_{f,l}^i e_{f,l}^i \right|^2$. Thus the conditional pairwise error probability is given by

$$P_c(\mathbf{S}, \hat{\mathbf{S}}) = Q\left(\sqrt{\frac{d_E^2(\mathbf{S}, \hat{\mathbf{S}})}{2N_0 + 2n_t E_s(1-\mu^2)}}\right) = Q\left(\sqrt{\frac{d_E^2(\mathbf{S}, \hat{\mathbf{S}})}{2\nu}}\right), \tag{6.18}$$

where

$$\nu^{-1} = \frac{\mu^2}{N_0 + n_t E_s(1-\mu^2)}. \tag{6.19}$$

The pairwise error probabilities of the I and Q codes for the case of imperfect SI follows from the case of perfect SI, which are given by (6.15) with ν in (6.19) replacing N_0.

6.2.4 Analytical Results

Throughout the chapter, we consider only the case of two transmit and one receive antennas. The code used is a 4-state QPSK I-Q code whose component codes have the trellis shown in Figure 6.3a. The throughput of this code is 2 bit/s and the code is clearly geometrically uniform. The bound on the bit error probability and simulation results are plotted in Figure 6.4 versus the average SNR per information bit $\gamma_b = \frac{E_b}{N_0}$. The figure shows the performance of the code with different channel estimation quality parameterized by the channel

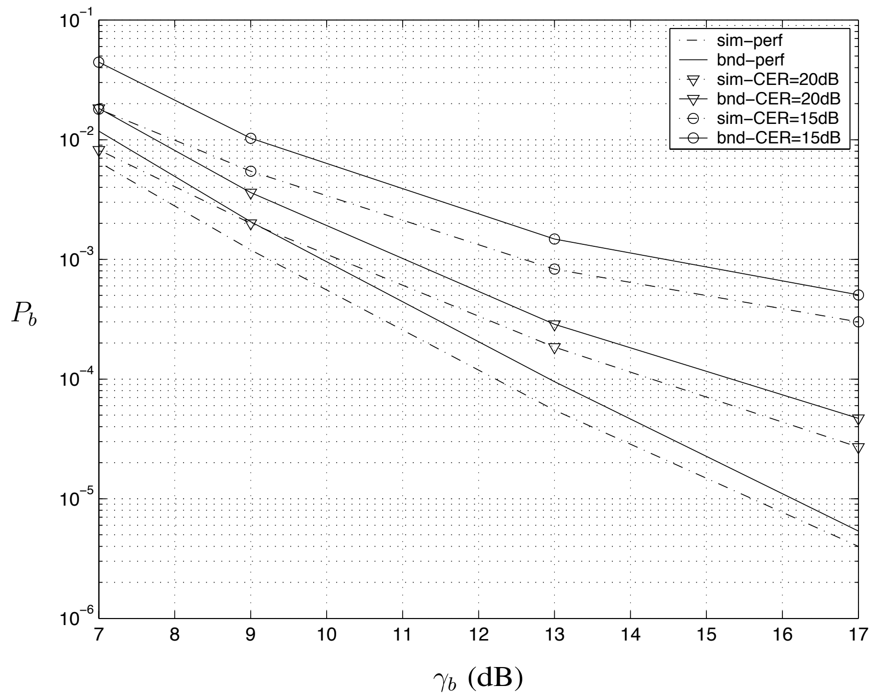

Figure 6.4: Performance of I-Q QPSK code using optimal decoding with perfect and imperfect SI parameterized by CER$= -\log \sigma_e^2$. (bnd: bound, sim: simulation).

gain-to-estimation error ratio: CER $= -\log(\sigma_e^2)$. The bound on the error probability is relatively tight at high SNR values as expected. Also the performance improvement with improved estimation quality exhibits a diminishing returns effect. Increasing the CER above 20 dB yields small improvement in performance except at high SNR. The decoder uses a 16-state super-trellis to decode. Larger time diversity could be achieved using codes with a larger number of states but the complexity of decoding grows rapidly. Thus we need to consider suboptimal decoding algorithms as in the following.

6.3 Geometrical Uniformity

In the following, we prove that if the component codes of an I-Q ST code are geometrically uniform then the overall I-Q ST code is geometrically uniform. A trellis code is said to be *geometrically uniform* [63] if the distance spectrum of the code relative to any codeword is the same as that taken relative to the all-zero codeword. In [64], Biglieri *et al.* derived sufficient conditions for geometrical uniformity of trellis codes. Consider a trellis code whose output signal s is given by a mapping of a binary code vector $\mathbf{c}$ onto a signal constellation point, $s = f(\mathbf{c})$. The code space $\mathbf{C}$ consisting of all possible codewords is partitioned into subsets $\mathcal{C}$ and $\tilde{\mathcal{C}}$. In the trellis, codewords from $\mathcal{C}$ are permitted at a subset of trellis states $\mathcal{S}$, where the other state subset $\tilde{\mathcal{S}}$ permits codewords from $\tilde{\mathcal{C}}$. Define the partition of the signal space corresponding to $\mathcal{C}$ as $\mathcal{A} = \{s : s = f(\mathbf{c}), \forall \mathbf{c} \in \mathcal{C}\}$. Similarly, $\tilde{\mathcal{A}}$ is defined resulting in partitioning the signal space into $\mathcal{A}$ and $\tilde{\mathcal{A}}$. In [64], it was shown that a trellis code is geometrically uniform if:

1. The subset $\tilde{\mathcal{C}}$ is a coset of $\mathcal{C}$, i.e., $\tilde{\mathcal{C}} = \mathcal{C} + \tilde{\mathbf{c}}$, where $\tilde{\mathbf{c}}$ is the coset representative of $\tilde{\mathcal{C}}$ in bits and addition is performed bitwise for each codeword in $\mathcal{C}$.

2. The signal partitions $\mathcal{A}$ and $\tilde{\mathcal{A}}$ are isometrics, i.e., they have the same distance spectrum,

$$d_E^2[f(\mathbf{c}), f(\mathbf{c}+\mathbf{e})] = d_E^2[f(\mathbf{c}+\tilde{\mathbf{c}}), f(\mathbf{c}+\tilde{\mathbf{c}}+\mathbf{e})], \qquad \forall \mathbf{c} \in \mathcal{C}, \mathbf{e} \in \mathbf{C}, \quad (6.20)$$

where d_E^2 represents the squared Euclidean distance between two signal points.

Proposition 6.1.

An I-Q ST code is geometrically uniform if its component I and Q codes are geometrically uniform.

Proof. Consider an I-Q ST code with geometrically uniform component codes, i.e., the code spaces of the I and Q encoders are partitioned into $\mathcal{C}_I$, $\tilde{\mathcal{C}}_I = \mathcal{C}_I + \tilde{c}_I$ and $\mathcal{C}_Q$, $\tilde{\mathcal{C}}_Q = \mathcal{C}_Q + \tilde{c}_Q$, respectively. The output signal vector of an I-Q ST encoder is given by a mapping $\mathbf{s} = f(\mathbf{c})$, where $\mathbf{c} = (\mathbf{c}_I, \mathbf{c}_Q)$ is the concatenation of the I and Q code vectors. Therefore, the code space of the I-Q ST encoder is partitioned into four sets $(\mathcal{C}_I, \mathcal{C}_Q), (\tilde{\mathcal{C}}_I, \mathcal{C}_Q), (\mathcal{C}_I, \tilde{\mathcal{C}}_Q)$ and $(\tilde{\mathcal{C}}_I, \tilde{\mathcal{C}}_Q)$, resulting in partitioning the complex signal into $(\mathcal{A}_I, \mathcal{A}_Q), (\tilde{\mathcal{A}}_I, \mathcal{A}_Q), (\mathcal{A}_I, \tilde{\mathcal{A}}_Q)$ and $(\tilde{\mathcal{A}}_I, \tilde{\mathcal{A}}_Q)$, where $\mathcal{A}_I$ and $\mathcal{A}_Q$ are the signal partitions corresponding to $\mathcal{C}_I$ and $\mathcal{C}_Q$, respectively. Now we have:

1. The code sets $(\mathcal{C}_I, \mathcal{C}_Q), (\tilde{\mathcal{C}}_I, \mathcal{C}_Q), (\mathcal{C}_I, \tilde{\mathcal{C}}_Q)$ and $(\tilde{\mathcal{C}}_I, \tilde{\mathcal{C}}_Q)$ are cosets of each other since their I and Q elements are cosets.
2. The complex signal partitions $(\mathcal{A}_I, \mathcal{A}_Q), (\tilde{\mathcal{A}}_I, \mathcal{A}_Q), (\mathcal{A}_I, \tilde{\mathcal{A}}_Q)$ and $(\tilde{\mathcal{A}}_I, \tilde{\mathcal{A}}_Q)$ are isometrics since their I and Q components are isometrics.

Thus I-Q ST codes that employ geometrically uniform component codes are also geometrically uniform. $\square$

6.4 Iterative Decoding

This section is devoted to suboptimal decoding of I-Q ST codes. The section starts with reviewing the decoding problem, and then two iterative decoders are described. The received signal at each receive antenna in (2.1) is expanded as

$$y_l = \sum_{i=1}^{n_t} (h^i_{I,l} s^i_{I,l} - h^i_{Q,l} s^i_{Q,l}) + j(h^i_{I,l} s^i_{Q,l} + h^i_{Q,l} s^i_{I,l}) + z_l, \tag{6.21}$$

where $h^i = h^i_I + jh^i_Q$. The ML decoding rule at the I and Q decoders requires the computation of the likelihood functions $p(\mathbf{Y}|\mathbf{S}_I, \mathbf{H})$ and $p(\mathbf{Y}|\mathbf{S}_Q, \mathbf{H})$, respectively. However, $\mathbf{Y}$ depends on $\mathbf{S}_I$ and $\mathbf{S}_Q$ which complicates the implementation of the optimal decoder. Therefore, low complexity decoders are of interest. Next we propose two iterative decoders.

6.4.1 Iterative Demodulation-Decoding (IDD)

The first algorithm views the I-Q encoding process as a concatenation of two independent stages: the encoding using the I and Q codes and the I-Q mapping of the 1-D I and Q signals to the 2-D signal constellation. Therefore, the overall system is viewed as a serial concatenation of a convolutional code and a block code (mapping), and hence iterative demodulation and decoding in [60, 65] can be used to demodulate and decode the information. The block diagram of this decoder is shown in Figure 6.5. It consists of a detection stage and two SISO modules for the I and Q codes. The detection stage receives the channel output for a frame and computes the following probabilities

$$p(y_l|\mathbf{s}_{I,l},\mathbf{h}_l) = Kp(\mathbf{s}_{I,l})\sum_{\forall \mathbf{s}_Q} p(y_l|\mathbf{s}_{I,l},\mathbf{s}_{Q,l},\mathbf{h}_l)p(\mathbf{s}_{Q,l}), \qquad l = 1,\ldots,N \tag{6.22}$$

$$p(y_l|\mathbf{s}_{Q,l},\mathbf{h}_l) = Kp(\mathbf{s}_{Q,l})\sum_{\forall \mathbf{s}_I} p(y_l|\mathbf{s}_{I,l},\mathbf{s}_{Q,l},\mathbf{h}_l)p(\mathbf{s}_{I,l}), \qquad l = 1,\ldots,N, \tag{6.23}$$

where $\mathbf{h}_l = \{h_l^i\}_{i=1}^{n_t}$, K is a normalization constant and $p(y_l|\mathbf{s}_{I,l},\mathbf{s}_{Q,l},\mathbf{h}_l)$ is the channel transition probability

$$p(y_l|\mathbf{s}_{I,l},\mathbf{s}_{Q,l},\mathbf{h}_l) = \exp\left(-\frac{\mathbf{m}(y_l,\mathbf{s})}{2\nu}\right)\Bigg|_{\mathbf{s}=(\mathbf{s}_{I,l},\mathbf{s}_{Q,l})}, \tag{6.24}$$

where $\mathbf{m}(y_l,\mathbf{s})$ is given by

$$\mathbf{m}(y_l,\mathbf{s}) = \begin{cases} \left|y_l - \sqrt{E_s}\sum_{i=1}^{n_t} h_l^i s^i\right|^2, & \text{perfect SI,} \\ \left|y_l - \frac{\mu}{\sigma}\sqrt{E_s}\sum_{i=1}^{n_t} \hat{h}_l^i s^i\right|^2, & \text{imperfect SI,} \end{cases} \tag{6.25}$$

and $\nu = N_0$ or given by (6.19) for perfect and imperfect SI, respectively. The probabilities $p(\mathbf{s}_{I,l})$, $p(\mathbf{s}_{Q,l})$ are the apriori information about the I and Q signal vectors at time l, which are assumed to be equally likely at the initialization of the algorithm. To avoid positive feedback of apriori information, only extrinsic information is passed to the I and Q decoders. The extrinsic information is defined as the probabilities in (6.22) and (6.23) after removing the contribution of the apriori information [2], and will be denoted as $p(\mathbf{s}_{I,l})$ and $p(\mathbf{s}_{Q,l})$.

The SISO module is a maximum aposteriori decoder that accepts soft information about signal vectors and updates them using the BCJR algorithm in [37]. The I and Q SISO decoders use $\{p(\mathbf{s}_{I,l})\}_{l=1}^{N}$ and $\{p(\mathbf{s}_{Q,l})\}_{l=1}^{N}$ as their

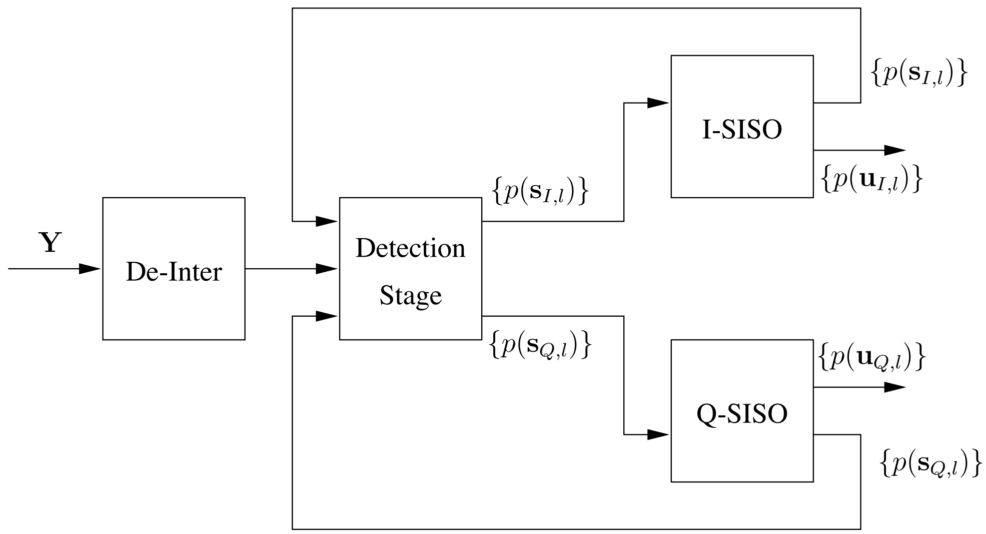

Figure 6.5: The structure of the IDD receiver.

observation vectors and compute soft information about signal vectors using

$$p\left(\mathbf{s}_{I,l}\middle|\mathbf{Y},\mathbf{H}\right) = K \sum_{(m,m^{'}):\mathbf{s}_{I,l}} \gamma_l(m,m^{'})h_{l-1}(m^{'})\beta_l(m), \qquad l = 1,\ldots,N, \tag{6.26}$$

where $\gamma_l(m,m^{'}) = p(y_l|\mathbf{s}_{I,l},\mathbf{h}_l)$ is the branch metric for a transition in the I code from state m at time l to state $m^{'}$ at time $l+1$, which is computed in the detection stage using (6.22). The variables h_l and β_l are the standard forward and backward recursions in the BCJR algorithm [37]. The same computation is performed in the Q-SISO decoder, and extrinsic information about the signal vectors are passed to the detection stage for the next iteration. The algorithm continues for a certain number of iterations and decision is made in the last iteration.

The detection stage needs to compute M^{n_t} metrics given by (6.25). The complexity of the I and Q SISO modules is linear in the number of states of the component codes, which reduces the decoding complexity compared to super-trellis significantly. Another decoding strategy with lower complexity is proposed in the following.

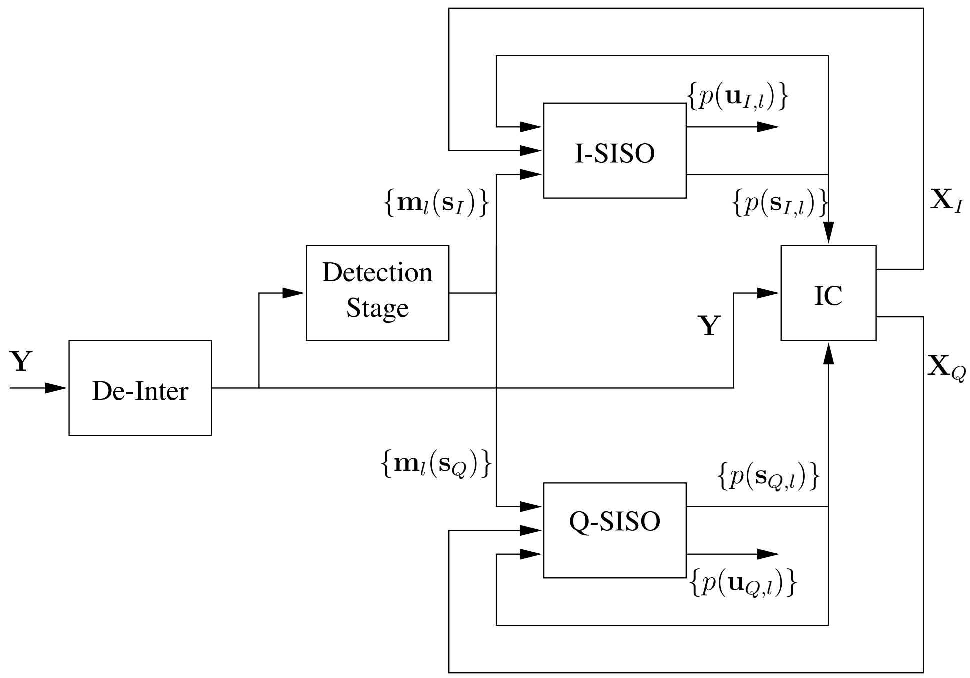

Figure 6.6: The structure of the ICD receiver.

6.4.2 Interference Cancellation Decoder (ICD)

In this algorithm, the decoding problem of I-Q ST codes is viewed as a multiuser detection problem. The block diagram of this decoder is shown in Figure 6.6, which is similar in spirit to parallel interference cancellation in [66, 67]. It consists of a detection stage, I and Q SISO modules and an interference cancellation (IC) stage. In the detection stage, metrics corresponding to all possible pairs of signal vectors $\{(\mathbf{s}_I, \mathbf{s}_Q) : \mathbf{s}_I \in \mathcal{A}_I,\, \mathbf{s}_Q \in \mathcal{A}_Q\}$ in the frame are computed using (6.25). The metric corresponding to each of the signal vectors $\mathbf{s}_I$, $\mathbf{s}_Q$ is

$$\mathbf{m}_l(\mathbf{s}_I) = \min_{\mathbf{s}_Q \in \mathcal{A}_Q} \mathbf{m}(y_l, \mathbf{s})\big|_{\mathbf{s}=(\mathbf{s}_I,\mathbf{s}_Q)}, \qquad \mathbf{s}_I \in \mathcal{A}_I,\, l = 1, \ldots, N \tag{6.27}$$

$$\mathbf{m}_l(\mathbf{s}_Q) = \min_{\mathbf{s}_I \in \mathcal{C}_I} \mathbf{m}(y_l, \mathbf{s})\big|_{\mathbf{s}=(\mathbf{s}_I,\mathbf{s}_Q)}, \qquad \mathbf{s}_Q \in \mathcal{A}_Q,\, l = 1, \ldots, N. \tag{6.28}$$

These metrics are fed to the I and Q SISO modules, which employ the BCJR algorithm. As in IDD, the SISO modules compute aposteriori probabilities of signal vectors using (6.26), with $\gamma_l(m, m')$ for the I and Q decoders is given, respectively by

$$\gamma_{I,l}(m, m^{'}) = \exp\left(-\frac{\mathbf{m}_l(\mathbf{s}_I)}{2\nu}\right)\bigg|_{\mathbf{s}_I:(m,m')} \tag{6.29}$$

$$\gamma_{Q,l}(m, m^{'}) = \exp\left(-\frac{\mathbf{m}_l(\mathbf{s}_Q)}{2\nu}\right)\bigg|_{\mathbf{s}_Q:(m,m')} \tag{6.30}$$

Once the I and Q SISO modules finish one frame, they pass the extrinsic information signal vectors to the IC stage. The IC stage forms new estimates of the I and Q faded signals by looking at (6.21) and using

$$x_{I,l} = y_l - \sqrt{E_s}\sum_{\forall \mathbf{s}_Q} p(\mathbf{s}_{Q,l}) \sum_{i=1}^{n_t} s^i_{Q,l}(-h^i_{Q,l} + jh^i_{I,l}), \qquad l = 1, \ldots, N \tag{6.31}$$

$$x_{Q,l} = y_l - \sqrt{E_s}\sum_{\forall \mathbf{s}_I} p(\mathbf{s}_{I,l}) \sum_{i=1}^{n_t} s^i_{I,l}(h^i_{I,l} + jh^i_{Q,l}), \qquad l = 1, \ldots, N. \tag{6.32}$$

These new observations along with the extrinsic information about signal vectors are passed to the SISO modules in the next iteration. In the proceeding iterations, the SISO modules operate on the vectors $\mathbf{X}_I = \{x_{I,l}\}_{l=1}^N$ and $\mathbf{x}_Q = \{x_{Q,l}\}_{l=1}^N$ as their new observations and update the soft information of the I and Q signal vectors. Note that the branch metric in the I and Q SISO modules are computed using the likelihood functions of the new observations obtained from the IC stage, i.e., $p(x_{I,l}|\mathbf{s}_{I,l}, \mathbf{h}_l)$ and $p(x_{Q,l}|\mathbf{s}_{Q,l}, \mathbf{h}_l)$. The algorithm keeps exchanging extrinsic information between the SISO modules and the IC stage for a number of iteration and decision is made in the last iteration.

The complexities of the detection stage and the SISO blocks are the same as in the IDD receiver. However, the detection stage in the ICD is used at the initialization of the algorithm only. In the rest of the iterations, the IC stage computes $M^{(n_t/2)}$ metrics given by (6.31) and (6.32). Therefore, the ICD is less complex than IDD at the cost of performance degradation as discussed below.

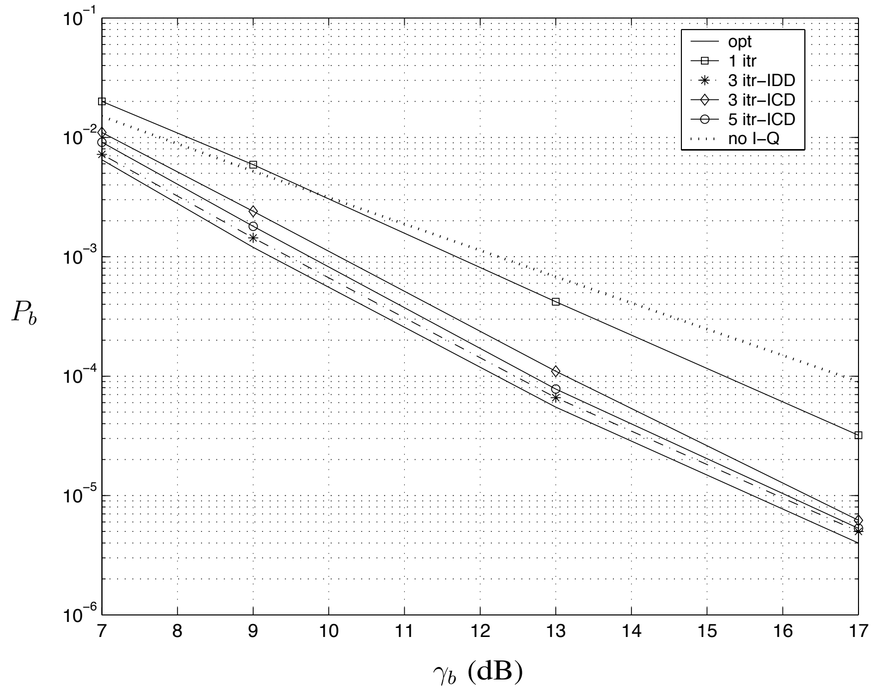

Figure 6.7: Simulation of the I-Q QPSK code using IDD and ICD with perfect SI. (itr: iterations, opt: optimal, no I-Q: QPSK code with single encoder).

6.4.3 Simulation Results

Figure 6.7 shows the performance of the 4-state I-Q QPSK code with IDD and ICD. The frame size is $N = 500$. Note that for one iteration the two algorithms are identical to the suboptimal algorithm in [33]. It can be seen that using both algorithms with 3 iterations perform very close to the optimal decoding. Also, the figure shows the performance of a single-encoder ST QPSK code that is optimized for independent fading channels [33]. This code uses a 4-state encoder and has double the complexity of the corresponding I-Q code with the same throughput of 2 bits/s. From the figure, it is clear that decoding the I-Q code using IDD with 3 iterations provides a coding gain of 5 dB at 10^{-4} bit error rate over the QPSK code with single encoder.

Figure 6.8 shows the performance of the 4-state 16-QAM I-Q ST code presented in Section 6.2.2. From the figure, the IDD algorithm with 3 iterations performs only 0.5 dB worse than optimal decoding. Also, ICD algorithm with 3 iterations provides a gain of 1.5 dB over the case of one iteration. For the

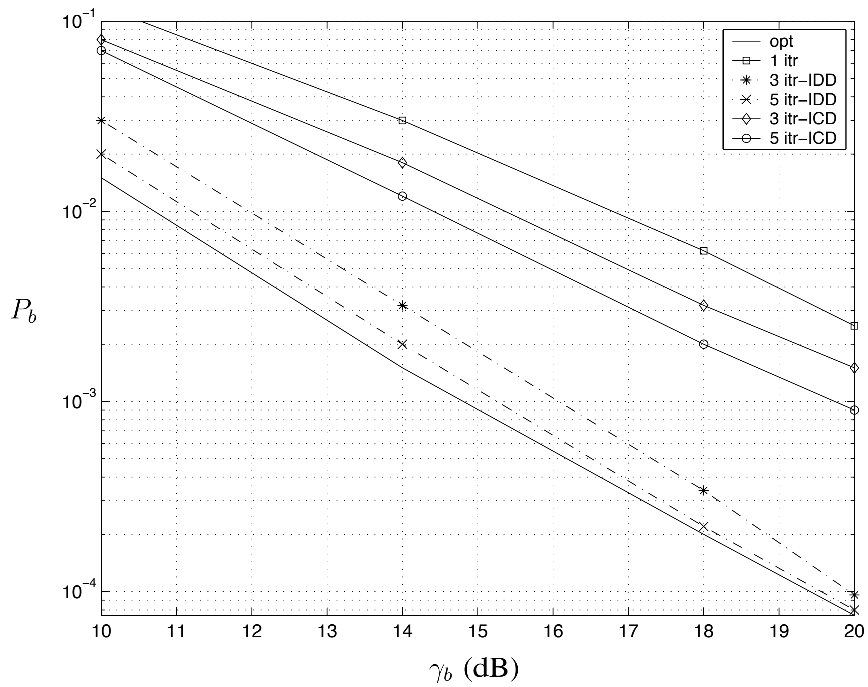

Figure 6.8: Simulation of the I-Q 16-QAM code using IDD and ICD with perfect SI. (itr: iterations, opt: optimal).

16-QAM case, the gains obtained from the ICD algorithm is less than the gains in the QPSK code. This is due to the reliability of the detection stage. As discussed in [33], the detection stage reliability is smaller for non-constant energy constellations. It is clear that decoding the 16-QAM code using the ICD algorithm does not converge to the optimum decoding as the number of iterations is increased, unlike the IDD algorithm. In general, ICD is less complex than IDD but at the cost of performance degradation.

Figure 6.9 shows the performance of the QPSK code with IDD and imperfect SI with CER=20 and 15 dB. We can see that 3 iterations provide performance close to the performance of the optimal decoder, which indicates that IDD is robust to channel estimation errors. Note that the performance is degraded by 1 dB and 2 dB for the cases of CER = 20 db and 15 dB, respectively. The performance degradation due to channel estimation errors is a common problem in ST codes and it increases as the number of transmit antennas increases as appears in (6.19). The effect of the frame size N on the performance of IDD and

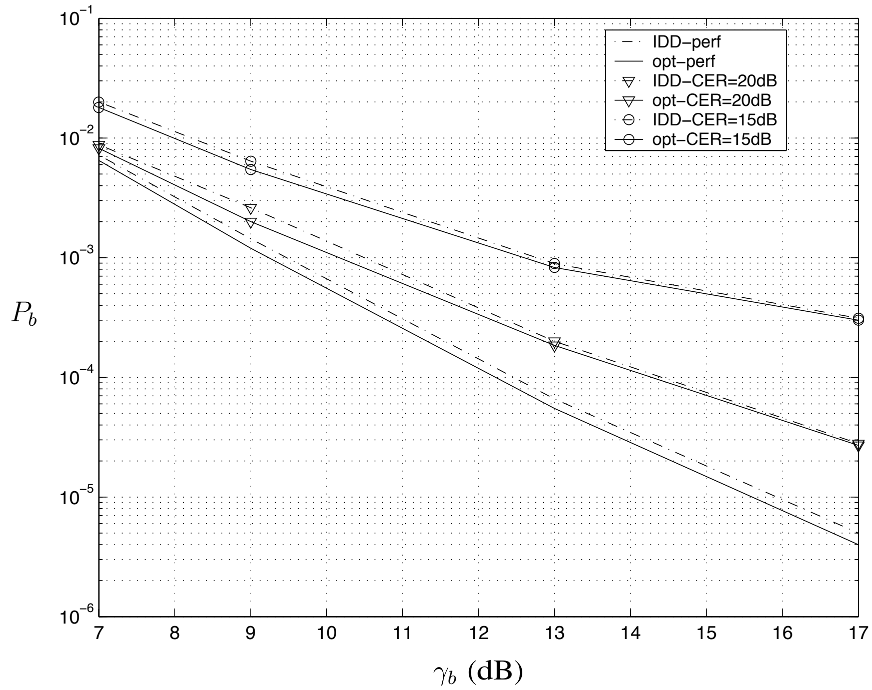

Figure 6.9: Simulation of the I-Q QPSK code with IDD using 3 iterations and imperfect SI parameterized by CER$= -\log \sigma_e^2$. (itr: iterations, opt: optimal).

ICD receivers for the QPSK and 16-QAM codes is shown in Figure 6.10. From the figure, the effect of reducing the frame size from $N = 500$ to $N = 200$ is negligible. Hence, both IDD and ICD algorithms are not sensitive to decoding delays and using them does not impose any delay constraint on the system.

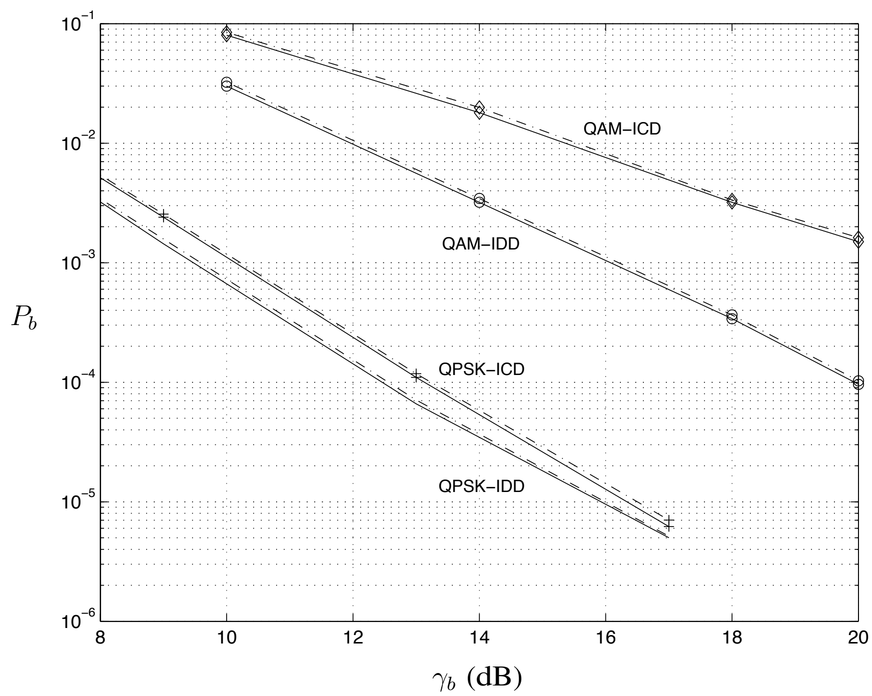

Figure 6.10: Simulation of the I-Q QPSK and 16-QAM codes using IDD and ICD with 3 iterations and perfect SI for different frame sizes (solid: $N = 500$**, dash:** $N = 200$**).**

CHAPTER 7

Conclusions and Future Research

In this chapter, we conclude this thesis by summarizing the content of the thesis and discussing possible future research directions.

7.1 Summary of Contributions

The most important contribution of this thesis is the performance analysis of binary coded systems over block fading channels using the union bound approach. In deriving the bound, we considered different receivers with different assumptions on the channel side information at the receiver. Expressions for the pairwise error probability for single and multi-antenna systems were derived. The tradeoff between the channel diversity and channel estimation was investigated assuming pilot-aided channel estimation. We introduced two assumptions to asses the performance of pilot-aided channel estimation; namely, the only pilot estimation (OPE) and the correct data estimation (CDE). Using the CDE assumption, a lower bound on the performance of iterative receivers employing joint decoding and channel estimation was evaluated. It was observed that the optimal channel memory tends to increase as under the CDE assumption because the channel estimation improves with increasing channel memory. Moreover, the optimal channel memory increases as the number of transmit antennas is increased.

The union bound of coded systems over block fading channels was extended to the cases when the fading in each fading block is distributed according to Rician and Nakagami distributions. The performance loss due to channel memory was clear from the analytical results. We observed that this loss increases as the channel becomes more fading, i.e., by reducing the specular-to-diffuse ratio in Rician fading and increasing the fading amount in Nakagami fading. The optimal channel memory was investigated for systems with different values of the specular-to-diffuse ratio of the channel. It was shown that the optimal channel memory increases as the specular-to-diffuse ratio of a Rician channel is increased.

As an effort to improve channel estimation in systems employing multi-antenna transmitters over block fading channels, an iterative receiver for decoding and channel estimation was presented. The iterative receiver performs very close to the best performance dictated by the system, i.e., under the CDE assumption. The convergence properties of the receiver were studied and it was shown via simulations that updating the channel estimation in more than 3 iterations increases the complexity of the system with no significant performance improvement. Also, the performance of the iterative receiver was tested for a large frame size and a large number of transmit antennas, and results show that the receiver performs well under these conditions. The tradeoff between channel estimation and channel diversity was investigated and the optimal channel memory was found via simulations. Moreover, the effect of the code error correcting capabilities on the optimal channel memory was investigated.

The performance of I-Q space-time codes was analyzed using the transfer functions of the component codes. From this, code design criteria for I-Q ST codes over independent fading channels were derived as functions of the parameters of the component codes rather than the super-trellis. The geometrical uniformity of I-Q space-time codes was established from the geometrical uniformity of the component codes. The decoding problem of I-Q codes was reviewed and two iterative decoding receivers were proposed. The performance and complexities of the proposed receivers were compared. Results showed that using the iterative receivers with 3 iterations results in performance that is very close to the optimal decoder.

7.2 Future Research

Since block fading channels approximate the fading behaviour in many important communication systems, different future research directions are possible. First, the union bound approach can be generalized to analyze the performance of coded systems over block fading channels with correlated fading blocks. An example of systems that encounter correlated fading blocks is multi-carrier transmission systems. In general, the fading processes at different carriers in a multi-carrier system may be correlated according the frequency selectivity of the channel. Another example is a system employing time-division multiplexing with a very slow varying channel. In this case, the fading affecting different time slots in the frame forms a correlated random process with a certain correlation function. Thus the effect of the number of carriers as well as the correlation function on the performance can be found from the analysis. Furthermore, the tradeoff between the channel estimation and channel diversity can be investigated. It would be interesting to investigate the effect of the channel correlation function on the optimal channel memory.

Another research avenue is to generalize the union bound to coded systems with arbitrary constellation size. By doing this, the performance of systems employing trellis coded modulation, space-time trellis codes and space-time block codes with arbitrary complex signal constellations can be analyzed. The effect of channel estimation and optimal channel memory can be investigated for constellations with constant and nonconstant-energy signals. Furthermore, the effect of the size and shape of the constellation on the optimal channel memory can be investigated, which may help in code and constellation designs for block fading channels.

The research in iterative decoding and channel estimation for multi-antenna systems can be directed towards designing low complexity receivers. As was demonstrated, the complexity of the channel update block is exponential in the number of transmit antennas. Thus iterative receivers with lower complexities are necessary for systems employing large number of transmit antennas.

APPENDICES

APPENDIX A

Appendix for Chapter 3

A.1 Derivation of (3.27) (for Section 3.2.3)

In the following, a complete derivation of (3.27) is presented which follows [12] in spirit. Expanding (3.26),

$$P_c(d|\mathbf{f}) \leq \prod_{f=1}^{L} \exp\Big(- 2\lambda \sum_{l=1}^{m} a_f s_{f,l} |s_{f,l} - \hat{s}_{f,l}| \Big)$$

$$\times \mathrm{E}_z \Big[\exp\Big(- 2\lambda \sum_{l=1}^{m} z_{f,l} |s_{f,l} - \hat{s}_{f,l}| \Big) \Big]. \tag{A.1}$$

For constant-energy signal constellations, $2s_{f,l}|s_{f,l} - \hat{s}_{f,l}| = |s_{f,l} - \hat{s}_{f,l}|^2$. As in [7], it can be shown that

$$\mathrm{E}_z \Big[e^{\left(-2\lambda \sum_{l=1}^{m} z_l |s_l - \hat{s}_l|\right)} \Big] = e^{\left(\lambda^2 N_0 \sum_{l=1}^{m} |s_l - \hat{s}_l|^2\right)}. \tag{A.2}$$

Substituting back in (A.1), the conditional pairwise error probability simplifies to

$$P_c(d|\mathbf{f}) \leq \prod_{f=1}^{L} \exp\Big(- \lambda a_f \sum_{l=1}^{m} |s_{f,l} - \hat{s}_{f,l}|^2 + \lambda^2 N_0 \sum_{l=1}^{m} |s_{f,l} - \hat{s}_{f,l}|^2 \Big). \tag{A.3}$$

Since $\sum_{l=1}^{m} |s_{f,l} - \hat{s}_{f,l}|^2 = 4E_s d_f$, where d_f is the number of nonzero locations in block f, (A.3) simplifies to

$$P_c(d|\mathbf{f}) \leq \prod_{f=1}^{L} \exp\left(-4E_s \lambda a_f d_f + 4E_s \lambda^2 N_0 d_f\right). \tag{A.4}$$

For the sake of simplifying presentation, we define $\gamma = 2R_c\gamma_b$, where $\gamma_b = \frac{E_s}{R_c N_0}$. Substituting $\tilde{\lambda} = \lambda N_0$,

$$P_c(d|\mathbf{f}) \leq \prod_{f=1}^{L} \exp\left(-2\tilde{\lambda} d_f a_f \gamma + 2\tilde{\lambda}^2 d_f \gamma\right). \tag{A.5}$$

Averaging over the fading gains $\{a_f\}_{f=1}^{F}$ as

$$\begin{aligned} P_u(d|\mathbf{f}) &\leq \prod_{f=1}^{L} \exp\left(2\tilde{\lambda}^2 d_f \gamma\right) \mathrm{E}_a\left[\exp\left(-2\tilde{\lambda} d_f a_f \gamma\right)\right] \\ &= \prod_{f=1}^{F} \exp\left(2\tilde{\lambda}^2 \gamma\right) \left[1 - 2\sqrt{\pi}\beta_f \exp(\beta_f^2) \mathrm{Q}(\sqrt{2}\beta_f)\right], \end{aligned} \tag{A.6}$$

where $\beta_f = \tilde{\lambda} d_f \gamma$. To find $\tilde{\lambda}$ that minimizes the bound, the approximation $\mathrm{Q}(x) \approx \frac{1}{2\sqrt{\pi}} \exp(-\frac{x^2}{2})(1 - \frac{1}{x^2})$ is used and (A.6) becomes

$$P_u(d|\mathbf{f}) \lesssim \prod_{v=1}^{w} \left(\frac{1}{2\beta_v^2}\right)^{f_v} \exp\left(2\tilde{\lambda}^2 f_v \gamma\right), \tag{A.7}$$

where in this case $\beta_v = \tilde{\lambda} v \gamma$. Recall that $d = \sum_{v=1}^{w} v f_v$, then the unconditional pairwise error probability simplifies to

$$P_u(d|\mathbf{f}) \lesssim \frac{\exp\left(2\tilde{\lambda}^2 d\gamma\right)}{2^L (\tilde{\lambda}\gamma)^{2L} \prod_{v=1}^{w} v^{2f_v}}. \tag{A.8}$$

This can be minimized over $\tilde{\lambda}$ resulting in an optimum value $\lambda_{\text{opt}}^2 = L/(2d\gamma)$. Substituting for λ_{opt} and γ, (3.27) follows directly.

APPENDIX B

Appendix for Chapter 4

B.1 Derivation of (4.9) (for Section 4.1.3)

In the following, complete derivation of (4.9) is presented which follows in general [12]. By averaging (A.5) over the density in (4.1),

$$P_u(d|\mathbf{f}) \leq \prod_{f=1}^{L} \exp\left(2\tilde{\lambda}^2 d_f \gamma\right) \mathrm{E}_a \left[\exp\left(-2\tilde{\lambda} d_f a_f \gamma\right)\right]$$

$$= \prod_{f=1}^{L} e^{(2\tilde{\lambda}^2\gamma)} e^{-K} \left[1 - \frac{1}{\sqrt{\pi}} \int_0^{\pi} \psi(\tau) e^{\psi^2(\tau)} \mathrm{Q}(\sqrt{2}\psi(\tau)) d\tau\right], \tag{B.1}$$

where $\gamma = 2R_c\gamma_b$ and the second line was taken from [12], where there was no proof provided.

$$\psi(\tau) = \frac{\tilde{\lambda} d_f \gamma}{\sqrt{1+K}} - \sqrt{K}\cos\tau. \tag{B.2}$$

For sufficiently large γ, $\psi(\tau)$ is dominated by its first term and it becomes independent of τ. In this case, (B.1) reduces to

$$P_u(d|\mathbf{f}) \lesssim \prod_{f=1}^{L} \exp\left(2\tilde{\lambda}^2 d_f \gamma\right) e^{-K} \left[1 - \frac{2\sqrt{\pi}\tilde{\lambda} d_f \gamma}{\sqrt{1+K}} \exp\left(\frac{\tilde{\lambda}^2 d_f^2 \gamma^2}{\sqrt{1+K}}\right) \mathrm{Q}\left(\frac{\sqrt{2}\tilde{\lambda} d_f \gamma}{\sqrt{1+K}}\right)\right]. \tag{B.3}$$

To find $\tilde{\lambda}$ that minimizes the bound, the approximation $\mathrm{Q}(x) \approx \frac{1}{2\sqrt{\pi}x}\exp(-\frac{x^2}{2})(1-\frac{1}{x^2})$ is used

$$P_u(d|\mathbf{f}) \lesssim \prod_{v=1}^{w} e^{-K}(1+K)\left(\frac{1}{2\beta_v^2}\right)^{f_v}\exp\left(2\tilde{\lambda}^2 f_v\gamma\right), \tag{B.4}$$

where $\beta_v = \tilde{\lambda}v\gamma$. Recall that $d = \sum_{v=1}^{w} vf_v$, then the pairwise error probability simplifies to

$$P_u(d|\mathbf{f}) \lesssim \frac{e^{-K}(1+K)^L\exp\left(2\tilde{\lambda}^2 d\gamma\right)}{2^L(\tilde{\lambda}\gamma)^{2L}\prod_{v=1}^{w} v^{2f_v}}. \tag{B.5}$$

This can be minimized over $\tilde{\lambda}$, resulting in an optimum value $\lambda_{\text{opt}}^2 = L/(2d\gamma)$. Substituting for λ_{opt} and γ, (4.9) follows directly.

B.2 Derivation of (4.14) (for Section 4.2.2)

In the following, (4.14) is derived in details. By averaging (A.5) over the density in (4.12) and using Eq. (3.462) from [68]

$$P_u(d|\mathbf{f}) \leq \prod_{f=1}^{L}\exp(2\tilde{\lambda}^2 d_f\gamma)\mathrm{E}_a\left[\exp\left(-2\tilde{\lambda}d_f a_f\gamma\right)\right]\prod_{f=1}^{F}\exp(2\tilde{\lambda}^2 d_f\gamma)\frac{\Gamma(2M)}{\Gamma(M)2^{M-1}} \tag{B.6}$$

$$\times\exp\left(\frac{\tilde{\lambda}^2 d_f^2\gamma^2}{2M}\right)D_{-2M}\left(\frac{\sqrt{2}\tilde{\lambda}d_f\gamma}{M}\right), \tag{B.7}$$

where $D_p(z)$ is the parabolic cylindrical function. Using Eq. (9.246) from [68], $D_p(z) \approx z^p e^{-z^2/4}$ for $z \gg 1, |z| \gg |p|$. Thus, for sufficiently large γ (i.e., for $\tilde{\lambda}d_f\gamma >> \frac{M}{\sqrt{2}}$ and $\tilde{\lambda}d_f\gamma >> \sqrt{2}M^{3/2}$)

$$D_{-2M}\left(\frac{\sqrt{2}\tilde{\lambda}d_f\gamma}{M}\right) \approx \left(\frac{\sqrt{2}\tilde{\lambda}d_f\gamma}{M}\right)^{-2M}\exp\left(-\frac{\tilde{\lambda}^2 d_f^2\gamma^2}{2M}\right). \tag{B.8}$$

Substituting (B.8) in (B.7), the unconditional error probability is approximately upper bounded as

$$P_u(d|\mathbf{f}) \lesssim \prod_{f=1}^{L} \exp(2\tilde{\lambda}^2 d\gamma) \frac{\Gamma(2M)}{\Gamma(M)2^{M-1}} \left(\frac{2\tilde{\lambda}^2 d_f^2 \gamma^2}{M} \right)^{-M}$$

$$= \left(\frac{\Gamma(2M)}{\Gamma(M)2^{M-1}} \right)^{L} \exp(2\tilde{\lambda}^2 d\gamma) \left(\frac{2\tilde{\lambda}^2 \gamma^2}{m} \right)^{-ML} \left(\prod_{v=1}^{w} v^{2Mf_v} \right)^{-1}, \qquad \text{(B.9)}$$

where $d = \sum_{v=1}^{w} v f_v$ by definition. The expression (B.9) can be minimized over $\tilde{\lambda}$, resulting in an optimum value $\lambda_{\text{opt}}^2 = ML/(2d\gamma)$. Substituting for λ_{opt} and γ in (B.9), the final expression of the pairwise error probability over Nakagami block fading channel with no amplitude SI is given in (4.13)

APPENDIX C

Appendix for Chapter 6

C.1 The Transfer Function of ST codes (for Section 6.2)

In this appendix, the transfer function of a ST trellis code is reviewed. It is a function that enumerates the number of codewords at every input weight and output distance and can be calculated using the error state diagram [64]. The label associated with each branch connecting two states in the error state diagram depends on the error vector between the signal associated with that branch and the zero signal vector s_0, i.e., $\mathrm{e} = \mathrm{s} - \mathrm{s}_0$. If an M-ary signal constellation is used at each transmit antenna, there are $M^{n_t} - 1$ different error vectors. Denote the distinct squared Euclidean distances from s_0 as $\{\xi_1, \xi_2, \ldots, \xi_n\} = \{\xi : \xi = E_s||\mathrm{s} - \mathrm{s}_0||^2, \mathrm{s} \in \mathcal{A}^{n_t}\}$, where $||\mathbf{x}||^2$ denotes the norm of a vector $\mathbf{x}$. In the error state diagram, each branch is labeled by $J^u D_1^{v_1} \ldots D_n^{v_n}$ where $v_l = 1$ if the corresponding signal vector has distance ξ_l. Also, the exponent of the variable J is the weight of the input vector causing the transition. For example, a state transition with input weight u and signal vector with distance ξ_l is represented by $J^u D_l$. The reader is referred to [64] for the complete details of computing the transfer function of trellis codes. The transfer function of the ST code is written as

$$T(J, D_1, \ldots, D_n) = \sum_u \sum_{v_1,\ldots,v_n} a(u, v_1, \ldots, v_n) J^u D_1^{v_1} \ldots D_n^{v_n}, \tag{C.1}$$

where $a(u, v_1, \ldots, v_n)$ is the number of codewords with input weight u and v_i error vectors with distance ξ_i from s_0, for $i = 1, \ldots, n$. Comparing the expressions in (6.9) and (C.1) and using the integral expression of the Q-function [45], the bit error probability can be written as

$$P_b \leq \frac{1}{\pi k} \int_0^{\pi/2} \left. \frac{\partial T(J, D_1, \ldots, D_n)}{\partial J} \right|_{J=1, D_v = \left(1 + \frac{\xi_v}{4N_0 \sin^2 \theta}\right)^{-n_r}, v=1,\ldots,n} d\theta. \tag{C.2}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] C. Shannon, "A Mathematical Theory of Communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October 1948.

[2] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo Codes," *IEEE Transactions on Communications*, vol. 44, pp. 1261–1271, October 1996.

[3] R. Gallager, *Low-Density Parity-Check Codes*, Ph.D. thesis, Massachusetts Institute of Technology (MIT), Cambridge, MA, 1963.

[4] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes," *IEEE Transactions on Information Theory*, vol. 47, pp. 619–637, February 2001.

[5] E. Hall and S. Wilson, "Design and Analysis of Turbo Codes on Rayleigh Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 160–174, February 1998.

[6] J. Hou, P. Siegel, and L. Milstein, "Performance Analysis and Code Optimization of Low-Density Parity-Check Codes on Rayleigh Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 924–934, May 2001.

[7] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, USA, 4th edition, 2000.

[8] R. J. McEliece and W. E. Stark, "Channels with Block Iinterference," *IEEE Transactions on Information Theory*, vol. 30, pp. 44–53, January 1984.

[9] E. Biglieri, J. Proakis, and S. Shamai, "Fading Channels: Information-Theoretic and Communications Aspects," *IEEE Transactions on Information Theory*, vol. 44, pp. 2619–2692, October 1998.

[10] W. C. Jakes, *Microwave Mobile Communications*, IEEE Press, New Jersy, USA, 1974.

[11] A. Viterbi, “Convolutional Codes and Their Performance in Communication Systems,” *IEEE Transactions on Communications*, vol. COM-19, pp. 751–772, October 1971.

[12] D. Divsalar and M. Simon, “Trellis Coded Modulation for 4800-9600 bits/s Transmission Over a Fading Satellite Channel,” *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, pp. 162–175, February 1987.

[13] A. Worthen and W. Stark, *On the Channel Memory-Diversity Tradeoff in Communication Systems*, in Information, Coding and Mathematics. Kluwer Academic Publisher, Netherlands, 2002.

[14] S. Rice, “Statistical Properties of a Sine Wave Plus Random Noise,” *Bell Systems Technical Journal*, vol. 27, pp. 109–157, January 1948.

[15] M. Nakagami, *“The m-Distribution- A General Formula of Intensity Distribution of Fading”*, in Statistical Methods in Radio Wave Propagation. Pergamon Press, London, 1960.

[16] E. Al-Hussaini and A. Al-Bassiouni, “Performance of MRC Diversity Systems for the Detection of Signals with Nakagami Fading,” *IEEE Transactions on Communications*, vol. COM-33, pp. 1315–1319, December 1985.

[17] H. Hashemi, “The Indoor Radio Propagation Channel,” *Proceedings of the IEEE*, vol. 81, pp. 943–968, July 1993.

[18] J. Guey, M. Fitz, M. Bell, and K. Wen-Yi, “Signal Design for Transmitter Diversity Wireless Communication Systems Over Rayleigh Fading Channels,” *Wireless Personal Communications*, vol. 47, pp. 527–537, April 1999.

[19] V. Tarokh, N. Seshadri, and A. Calderbank, “Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction,” *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, March 1998.

[20] S. Alamouti, “A Simple Transmit Diversity Technique for Wireless Communications,” *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451–1458, October 1998.

[21] V. Tarokh, H. Jafarkhani, and A. Calderbank, “Space-Time Block Codes from Orthogonal Designs,” *IEEE Transactions on Information Theory*, vol. 45, pp. 1456–1467, July 1999.

[22] V. Tarokh and H. Jafarkhani, “A Differential Detection Scheme for Transmit Diversity,” *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 1169–1174, July 2000.

[23] V. Tarokh, A. Naguib, N. Seshadri, and A. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion in the Presence of Channel Estimation Errors, Mobility, and Multiple Paths," *IEEE Transactions on Communications*, vol. 47, pp. 199–207, February 1999.

[24] J. Kang and W. Stark, "Iterative Estimation and Decoding for FH-SS with Slow Rayleigh Fading," *IEEE Transactions on Communications*, vol. 48, pp. 2014–2023, December 2000.

[25] K. Chugg, A. Anastasopoulos, and X. Chen, *Adaptive Detection: Adaptivity, Complexity Reduction, and Applications*, Kluwer Academic Publisher, Netherlands, 2001.

[26] A. Anastasopoulos and K. Chugg, "Adaptive Soft-Input Soft-Ouput Algorithms for Iterative Detection with Parametric Uncertainty," *IEEE Transactions on Communications*, vol. 48, pp. 1638–1649, October 2000.

[27] A. P. Worthen, *Codes and Iterative Receivers for Wireless Communication Systems*, Ph.D. thesis, University of Michigan, Ann Arbor, MI, May 2001.

[28] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel Estimation for OFDM Systems with Transmitter Diversity in Mobile Wireless Channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 461–471, March 1999.

[29] Y. Li, C. Georghiades, and G. Huang, "Iterative Maximum-Likelihood Sequence Estimation for Space-Time Coded System," *IEEE Transactions on Communications*, vol. 49, pp. 948–951, June 2001.

[30] B. Lu, X. Wang, and K. Narayanan, "LDPC-Based Space-Time Coded OFDM Systems Over Correlated Fading Channels: Performance Analysis and Receiver Design," *IEEE Transactions on Communications*, vol. 50, pp. 74–88, January 2002.

[31] A. Grant, "Joint Decoding and Channel Estimation for Space-Time Codes," *IEEE Vehicular Technology Conference, VTC/Fall*, pp. 416–420, 2000.

[32] A. Grant, "Joint Channel Estimation and Decoding for Space-Time Turbo Codes," *IEEE International Symposium on Information Theory, ISIT*, p. 161, 2002.

[33] S. Zummo and S. Al-Semari, "Design of Space-Time Trellis Codes for Rapid Rayleigh Fading Channels," *To appear in the Journal of Communications and Networking (JNC)*, 2003.

[34] D. Divsalar and M. Simon, "The Design of Trellis Coded MPSK for Fading Channels: Set Partitioning for Optimum Code Design," *IEEE Transactions on Communications*, vol. 36, pp. 1013–1021, September 1988.

[35] J. Wozencraft and I. Jacobs, *Principles of Communication Engineers*, John Wiley & Sons, Inc., New York, USA, 1965.

[36] A. Naguib, V. Tarokh, N. Seshadri, and A. Calderbank, "Approaching Shannon's Capacity Limit by 0.2 dB Using Simple Hamming Codes," *IEEE Communications Letters*, vol. 1, pp. 130–132, September 1997.

[37] 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. IT-20, pp. 284–287, March 1974.

[38] W. Firmanto, B. Vucetic, and J. Yuan, "Space-Time TCM with Improved Performance on Fast Fading Channels," *IEEE Communications Letters*, vol. 5, no. 4, pp. 154–156, April 2001.

[39] J. Yuan, B. Vucetic, C. Zhuo, and W. Firmanto, "Design of Space-Time Turbo TCM on Fading Channels," *Information Theory Workshop*, pp. 123–125, 2001.

[40] Y. Leung, S. Wilson, and J. Ketchum, "Multifrequency Trellis Coding with Low Delay for Fading Channels," *IEEE Transactions on Communications*, vol. 41, pp. 1450–1459, October 1993.

[41] E. Malkamaki and H. Leib, "Coded Diversity on Block-Fading Channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 771–781, March 1999.

[42] R. Knopp and P. Humblet, "On Coding for Block-Fading Channels," *IEEE Transactions on Information Theory*, vol. 46, pp. 189–205, January 2000.

[43] E. Malkamaki and H. Leib, "Evaluating the Performance of Convolutional Codes Over Block Fading Channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 1643–1646, July 1999.

[44] D. Divsalar and R. Pollara, "Performance Analysis of Turbo Codes," *IEEE Military Communications Conference, MILCOM*, pp. 91–96, 1995.

[45] M. Simon and D. Divsalar, "Some New Twists to Problems Involving the Gaussian Probability Integral," *IEEE Transactions on Communications*, vol. 46, pp. 200–210, February 1998.

[46] R. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, Inc., New York, USA, 1968.

[47] K. Fu and A. Anastasopoulos, "Performance Analysis of LDPC Codes for Time-Selective Complex Fading Channels," *Global Telecommunications Conference, GLOBECOM*, pp. 1279–1283, 2002.

[48] R. Nuriyev and A. Anastasopoulos, "Analysis and Design of Pilot-Symbol-Assisted Codes, for the Noncoherent AWGN Channel, Using Density Evolution," *IEEE International Conference on Communication, ICC*, pp. 1511–1515, 2002.

[49] R. Nuriyev and A. Anastasopoulos, "Analysis of Joint Iterative Decoding and Phase Estimation for the Noncoherent AWGN Channel, Using Density Evolution," *IEEE International Symposium on Information Theory, ISIT*, p. 168, 2002.

[50] J. Salz and J. Winters, "Effect of Fading Correlation on Adaptive Arrays in Digital Wireless Communications," *IEEE International Conference on Communication, ICC*, pp. 1768–1774, 1993.

[51] P. Shamain and L. Milstein, "Detection With Spatial Diversity Using Noisy Channel Estimates in A Correlated Fading Channel," *IEEE Military Communications Conference, MILCOM*, pp. 691–696, 2002.

[52] I. Jacobs, "Probability of Error Bounds for Binary Transmission on the Slowly Fading Rician Channel," *IEEE Transactions on Information Theory*, vol. IT-12, pp. 431–441, October 1966.

[53] U. Charash, "Reception Through Nakagami Fading Multipath Channels with Random Delays," *IEEE Transactions on Communications*, vol. COM-27, pp. 657–670, April 1979.

[54] K. Noga, "The Performance of Binary Transmission in Slow Nakagami-Fading Channels with MRC Diversity," *IEEE Transactions on Communications*, vol. 46, pp. 863–865, July 1998.

[55] M. Alouini and M. Simon, "A Unified Approach to the Performance Analysis of Digital Communication Over Generalized Fading Channels," *IEE Proceedings*, vol. 86, pp. 1860–1877, September 1998.

[56] A. Swindlehurst, "Simultaneous Channel Estimation and Decoding for Diagonal Space-Time Codes," *Workshop on Sensor Array and Multichannel Signal Processing*, pp. 173–177, 2000.

[57] C. Budianu and L. Tong, "Channel Estimation for Space-Time Orthogonal Block Codes," *IEEE International Conference on Communication, ICC*, pp. 1127–1131, 2001.

[58] C. Komninakis and R. Wesel, "Joint Iterative Channel Estimation and Decoding in Flat Correlated Rayleigh Fading," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 1706–1717, September 2001.

[59] S. Zummo and S. Al-Semari, "Space-Time Coded 8-PSK Schemes for Rapid Rayleigh Fading Channels," *EURASIP Journal on Applied Signal Processing*, vol. 5, pp. 482–486, May 2002.

[60] A. Tonello, "Space-Time Bit-Interleaved Coded Modulation with Iterative Decoding Strategy," *IEEE Vehicular Technology Conference, VTC/Fall*, pp. 473–478, 2000.

[61] S. Al-Semari and T. Fuja, "I-Q TCM Reliable Communication Over Rayleigh Fading Channel Close to the Cutoff Rate," *IEEE Transactions on Information Theory*, vol. 43, pp. 250–262, January 1997.

[62] S. H. Jamali and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*, Kluwer Academic Publisher, Netherlands, 1994.

[63] G. Forney, "Geometrically Uniform Codes," *IEEE Transactions on Information Theory*, vol. 5, no. 37, pp. 1241–1260, September 1991.

[64] E. Biglieri, D. Divsalar, P. McLane, and M. Simon, *Inroduction to Trellis-Coded Modulation with Applications*, Macmillan Publishing, New York, USA, 1991.

[65] G. Bauch, "Concatenation of Space-Time Block Codes and "Turbo"-TCM," *IEEE International Conference on Communication, ICC*, pp. 1202–1206, 1999.

[66] D. Divsalar, M. Simon, and D. Raphaeli, "Improved Parallel Interference Cancellation for CDMA," *IEEE Transactions on Communications*, vol. 46, pp. 258–268, February 1998.

[67] S. Ariyavisitakul, "Turbo Space-Time Processing to Improve Wireless Channel Capacity," *IEEE Transactions on Communications*, vol. 48, pp. 1347–1359, August 2000.

[68] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press Inc., London, UK, 4th edition, 1965.