COMMUNICATION OVER THE NONCOHERENT CHANNEL

by

Rza Nuriyev

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:

Assistant Professor Achilleas Anastasopoulos, Chair
Associate Professor Kevin J. Compton
Professor David L. Neuhoff
Professor Wayne E. Stark
Associate Professor Kim A. Winick

To My Parents

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Prof. Achilleas Anastasopoulos, for his constant encouragement, advice and support throughout my research. I have benefited greatly in academic terms from his thorough knowledge of communication theory, as well as in my personal life from his practical mind and wisdom. It is fair to say that this thesis would not be possible without his constant curiosity, gracious patience and without the freedom that I have genuinely felt in pursuing new academic endeavors. I would also like to thank my committee members, Professors David Neuhoff, Wayne Stark, Kim Winick and Kevin Compton for their genuine interest, kind consideration and patience.

My sincere gratitude goes to my friends for their support and encouragement: Maxim Zalota for constantly bringing joy into our lives, Victor Perlin for his unrelenting curiosity and competitiveness, Hua Wang for his, at times funny, skepticism that helped keep my sanity, Dave Rashty for bringing a sober note into the discussions, and so many others whose friendship have touched my life.

Finally I would like to thank my family for their unending support, confidence in me and for constantly inspiring me to pursue my dreams. Among other things, my father has given me the determination in life, my mother has endued me with the sense of fairness and magnanimity, and my sister has supported me in ways that parents can not.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

CHAPTER I

INTRODUCTION

The very first electrical communication system, the telegraph, was introduced by Samuel Morse in 1837 and was digital in nature. However, it was not until 1924, when Nyquist analyzed the maximum signaling rate for the telegraph channel, that digital communication was seriously considered from a scientific point of view. This work was followed by the ideas of Hartley (1928), Wiener (1942) and Kotelnikov (1947). In 1948 Shannon established the mathematical foundation for *Information Theory*, and determined the fundamental limits of digital communication systems. With the work of Hamming in 1950 on error-correcting codes, *Coding Theory* was initiated and since then has developed into a powerful and important field.

Communication engineers are mainly concerned with two issues. The first, and perhaps the foremost, is the design and analysis of versatile coding schemes that provide satisfactorily reliable communication given the limited available resources. The second, a more theoretical aspect, is that of finding the fundamental limits on the maximum achievable transmission rate for the communication system under consideration.

In order to analyze a particular communication system, the physical medium over which information is transmitted is abstractly represented by a channel model.

This model, while not exactly reflecting finer qualitative peculiarities of the physical medium, is intended to capture the basic properties of the latter and yet remain tractable for subsequent mathematical analysis.

Choosing a channel model that is both accurate and simple to analyze is a difficult task. Therefore, some relatively basic channel models have been proposed in the literature and have been verified in practice to adequately model a wide variety of physical media. Of these models, perhaps the most common is the additive noise channel, where the transmitted signal is perturbed by additive noise of known statistics. An extension of this simple model is one where the transmitted signal, apart from the additive noise, is corrupted by a multiplicative term, often referred to as fading.

In any practical communication scenario, there are also constraints dictated by the user and the underlying application. Such constraints include, but are not limited to, transmitter power, bandwidth, complexity, delay, etc. In his pioneering work [79, 80], Shannon showed that for a given channel model, there exists a single parameter, called *channel capacity*, which describes the fundamental tradeoffs between transmitter power, bandwidth, and maximum rate at which information can be reliably conveyed over the channel.

In this thesis we investigate the communication problem over additive noise channels that also introduce a carrier phase rotation unknown to the transmitter and the receiver. The motivation for investigating these channels comes from considering the front-end of a typical wireless communication system as shown in Fig. 1.1. The information signal is modulated using a local oscillator onto the carrier and transmitted over the channel. At the receiver, this operation is reversed by mixing the received signal with a locally generated carrier. Due to the spatial separation between the

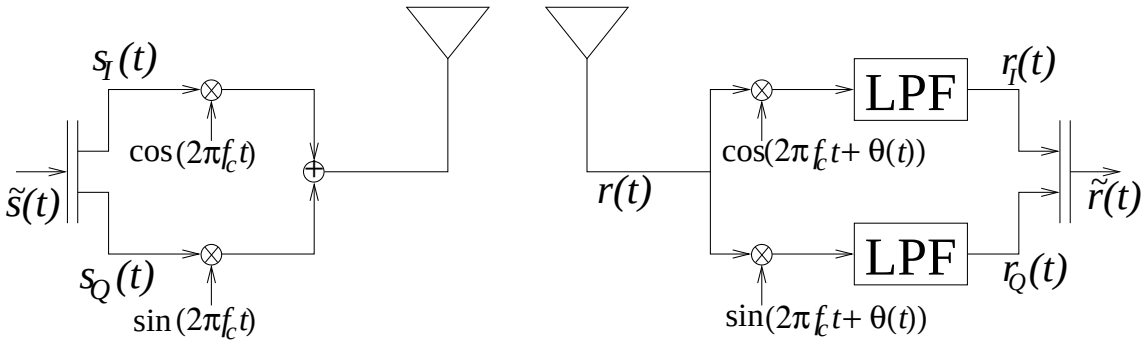


Figure 1.1: Front-end of a bandpass wireless communication system.

transmitter and the receiver, the corresponding oscillator phases differ by a random amount $\theta(t)$. Furthermore, $\theta(t)$ can be time-varying as a direct result of the local oscillator instabilities, the transmitter/receiver mobility, and the propagation medium. Thus, the overall effect on the complex envelope of the transmitted signal can be described by

$$\tilde{r}(t) = \tilde{s}(t)e^{j\theta(t)} + \tilde{n}(t), \tag{1.1}$$

where $\tilde{n}(t)$ is the additive noise.

There are two solutions to this phase mismatch encountered in traditional communication systems. The first solution is to use signaling schemes that are immune to such phase rotations (*e.g.*, frequency shift keying (FSK), or in general orthogonal constellations). The second solution is to eliminate the random phase offset by employing a subsystem that estimates $\theta(t)$ and effectively derotates the observation. In this work we follow a more direct approach, in that we incorporate the phase offset into the overall channel model. Finding the fundamentals limits of communication over this channel and designing practical schemes that come close to these limits is our ultimate goal.

1.1 Channel Model

Several representations have been proposed in the literature for the noncoherent channel, based on how the underlying phase process, $\theta(t)$, is modeled. The simplest such model is one where the phase remains constant, but random, throughout the length of the transmitted codeword.

In order to take into account the channel dynamics, more sophisticated models have been proposed that consider small variations of the phase process between adjacent symbols. A common such model is one where the phase process is a random walk with independent Gaussian increments, where the variance of the increment determines the dynamics of the process. A simplified version of this model has also been used where now the phase process is assumed to be a walk on a grid with finite precision, which results in a Markov process with finite number of states (see [95, 67, 58, 48] and references therein). An even simpler, commonly used model for this channel is one where the unknown phase is considered constant, but random, over a block of N symbols, and independent from block to block. More precisely, the channel is described by the following input/output relationship

$$\mathbf{y}_k = \mathbf{x}_k e^{j\theta_k} + \mathbf{n}_k, \tag{1.2}$$

where $\mathbf{x}_k$, $\mathbf{y}_k$ and $\mathbf{n}_k$ are complex sequences of length N, denoting the k-th block of the transmitted sequence, the observed sequence, and the noise sequence of independent identically distributed (i.i.d.), zero-mean, circular, complex Gaussian random variables with variance $\sigma^2 = N_0/2$ per real dimension, respectively. The unknown phase rotation in each block is represented by the variables θ_k, which are modeled as i.i.d. random variables uniformly distributed in $[-\pi, \pi)$. The parameter N, also referred to as the *channel coherence interval*, is assumed known at both the transmit-

ter and the receiver, as are the positions of the block boundaries. This channel will be referred to as the *block-independent noncoherent AWGN channel*. It adequately captures the phase dynamics with a single parameter N and yet is simple enough for analysis. Furthermore, the assumption of block independence is particularly realistic for frequency hopping systems, where the phase remains constant for the period of a hop and changes arbitrarily from hop to hop. The general modeling of a channel with memory by a block-independent model has been originally introduced in [56].

The block-independent noncoherent AWGN channel resembles two widely used and well-studied channels. The first one is the coherent AWGN channel, where the $e^{j\theta}$ term is absent from (1.2). The second one is the block-independent Rayleigh fading channel where the $e^{j\theta}$ term is multiplied by a Rayleigh distributed random variable, representing the amplitude variation. Despite these similarities, the noncoherent channel is not as well understood as the coherent and the Rayleigh fading counterparts. For example, the capacity of this channel is not known, nor is the structure of the capacity achieving input density. Furthermore, from the more practical point of view, performance of the best-known coding schemes for this channel is still far from the theoretical limits, unlike the case of the coherent AWGN channel.

In the next section we summarize the research topics discussed in this thesis. Throughout this work, with the exception of Chapter V, the aforementioned block-independent noncoherent AWGN channel model is utilized.

1.2 Research Topics and Contributions

Two main problems regarding the noncoherent channel are of interest to us. The first, is the design and analysis of practical codes whose performance is close to the theoretical limit. Although, this topic has been of interest for a long time, most of

the results in the literature lack the systematic approach. Moreover, a great deal of research applies only to the case where the code blocklength is equal to the channel coherence interval N, *i.e.*, the unknown phase is constant over the entire codeword. For small values of N, this assumption implies the use of short codes, and hence results in poor performance. Similarly, for large values of N, the above assumption implies a very slowly varying phase process, which is a useless model in mobile wireless links. We are approaching the problem of coding in a more systematic way, that also allows the design of codes which span over several blocks of length N.

The second, more theoretical problem, is finding the fundamental limits of communication for this channel. More precisely we are interested in the evaluation of the Shannon capacity and the investigation of the structure of the capacity achieving input density of this channel. As mentioned above, the noncoherent channel lies in between the coherent AWGN and the Rayleigh fading channels. However, the information theoretic results for these two channels differ significantly, and it is not clear which of the two extremes the noncoherent channel is closer to. As it turns out, the structure of the capacity achieving density borrows properties from both these extremes.

The sections that follow correspond to subsequent chapters, and are intended as short introductions to the topics discussed therein. The contents of the chapters have in part appeared in several papers. The results of Chapter II can be found in [60, 61, 65], contents of Chapters III and IV are combined in [64, 62, 63], and finally Chapter V corresponds to [59, 66].

1.2.1 Pilot-Symbol-Assisted Coding

A simple and practical approach in the direction of designing codes for the noncoherent channel is the use of pilot-symbol-assisted (PSA) transmission in conjunction with powerful codes designed for the coherent AWGN channel. More precisely, we are interested in coding schemes that transmit a pilot symbol of certain energy in each block of length N, over which the carrier phase remains constant, and the remaining $N-1$ symbols are utilized for transmission of coded bits from a code designed for the coherent AWGN channel. A higher pilot energy translates to better-quality phase estimates, but comes at the expense of reduced coded-bit energy. This tradeoff between the quality of the phase estimate and the effective coded bit signal-to-noise ratio (SNR), implies an optimal pilot-symbol energy allocation, which minimizes the information bit SNR required to achieve good performance. In Chapter II, PSA transmission is investigated in conjunction with low-density parity-check (LDPC) codes. Several low complexity receivers which also lend themselves to analysis are proposed. A powerful technique devised for the analysis of LDPC codes over the AWGN channel, called density evolution, is utilized for performance analysis of these systems. It turns out that optimizing the energy allocation results in an SNR gain of 0.5-2.5 dB, depending on the channel coherence interval N, as well as the particular receiver algorithm used. The presented PSA codes and the corresponding receivers can serve as baseline – yet powerful – transmission schemes, against which more elaborate code designs can be compared.

Although PSA transmission allows the use of a vast amount of codes designed for the coherent AWGN channel, and despite the fact that it performs well for moderate-to-large values of the channel coherence interval N, it is expected that for small values of N, such a scheme will be highly suboptimal. This behavior can be explained by

observing that the power loss due to pilots is amplified for small values of N, and is the motivation for considering more direct approaches to code design. As a first step towards this goal, we consider the information-theoretic aspect of the communication problem.

1.2.2 Capacity of the Noncoherent AWGN channel

The problem of finding the capacity and the structure of the capacity-achieving input density for this channel has been of interest for some time, and some partial results have appeared in the literature. In [56, 83, 33] the capacity of the general block-independent channel models has been considered. In [12] the capacity of the noncoherent AWGN channel under FSK modulation was investigated, while capacity evaluations were performed in [68] when phase shift keying (PSK) modulation is used. A numerical lower bound on the capacity was evaluated in [18] by assuming Gaussian distributed input to the channel. Very recently, some new results have been reported for the special case of a memoryless phase process, *i.e.*, for $N = 1$; in [38, 39] it was shown that the capacity achieving input is discrete with infinite number of mass points, while in [44] upper and lower bounds on the capacity were derived. To date, a complete characterization of the capacity-achieving input density is not available for arbitrary phase dynamics, *i.e.*, for arbitrary N.

Our approach aims more at finding the structure of the capacity-achieving density rather than numerically evaluating the capacity. We do this by investigating the necessary and sufficient conditions the optimal input density should satisfy. In general, this problem involves an optimization over densities of N-dimensional complex random vectors. However, we show that by proper parameterization of the space of input random vectors, we can reduce this problem to optimization over densities of

one real variable, namely the density of the amplitude of the N-dimensional complex input vector. Taking this as a starting point we first prove the existence and uniqueness of the optimum amplitude density. Furthermore, we show that the optimum amplitude density is discrete in nature, has infinitely many mass points extending to infinity, and always has a mass point at zero. In other words, the capacity-achieving input signal isotropically occupies the surfaces of infinitely many concentric spheres in the N-dimensional complex space, and is zero with a fixed non-zero probability.

In addition, we derive asymptotic results which show that the mass point probabilities decrease square-exponentially with mass point amplitudes, thus suggesting a reasonable approximation to the optimal density by a discrete density with only finite number of mass points. This result is verified by numerically evaluating the capacity for several values of N and SNR.

To better understand the implications of these results for the code design problem, two practical performance measures are introduced and quantified: "discretization loss" and "shaping gain". The former is the capacity loss due to the use of practical modulation schemes in place of the optimal density. This loss is due to the fact that not all but only a finite number of points on the surface of a sphere are used in a specific modulation scheme. The latter is the gain obtained by augmenting a single-sphere input density with a zero-mass point of certain probability. It is this gain that we try to realize using more direct approaches to code design as described in the following section.

1.2.3 Capacity-Inspired Coding

In this section we discuss code and modulation design based on the capacity results discussed above. Existing coding strategies for the noncoherent channel are

either indirect, or based on very simplistic models for the channel. Early works on signal design were based on minimizing the probability of error using symbol-by-symbol noncoherent detection [78]. Codes with rich algebraic structure were investigated in [40] and good designs were found through computer search. Trellis coded modulation (TCM) was extended to noncoherent transmission in [75] and [2] for the case of equal-energy and unequal-energy signals, respectively. The set partitioning rules for the design of coherent TCM codes in [86] were adopted in [92, 24] to modify these codes for the noncoherent channel. Coding schemes incorporating differential encoding, or more generally rotational invariant encoders [85], were proposed in [70, 75, 52]. High-performance codes were designed in [36] for the noncoherent channel by combining optimized binary irregular LDPC codes and differential encoding. These codes represent the state of the art in code design for the noncoherent channel. It should also be noted that code design for the noncoherent AWGN channel is similar to code design for the block-independent noncoherent fading channel (*i.e.*, where amplitude variation is present in addition to phase rotation), and thus the codes investigated in [88, 69, 34] for the fading channel are also relevant designs.

In order to get some insight on the design criteria for good noncoherent codes, we examine the probability of decoding error. Clues from this investigation combined with the information theoretic results of the previous section will be used to design powerful coding schemes.

We initiate the discussion with the simple case when the unknown phase remains constant over the entire codeword. By virtue of the union bound, the probability of codeword error is dominated – at moderate to high SNR values – by the pairwise probability of error between the "worst" pair of codewords. It can be shown that, in the case of equal-energy codewords, $\|\mathbf{a}\| = \|\mathbf{b}\|$, the pairwise probability of error for

the noncoherent AWGN channel can be bounded as (see [72] for exact expression)

$$P(\mathbf{a} \to \mathbf{b}|\mathbf{a}) = \mathcal{R}(\rho, N\gamma) \leq \frac{3}{2} \exp\Big(\frac{-N\gamma(1-\rho)}{2}\Big), \tag{1.3}$$

where γ is the per-symbol SNR and ρ is the magnitude of the cross correlation between the codewords $\mathbf{a}$ and $\mathbf{b}$, given by

$$\rho = \frac{|\mathbf{a}^H \mathbf{b}|}{\|\mathbf{a}\|\|\mathbf{b}\|}, \tag{1.4}$$

with superscript H denoting conjugate transpose. This bound emphasizes the importance of the cross correlation coefficient: error probability decreases exponentially in $N\gamma(1-\rho)/2$. This suggests the first practical design rule, that is, *minimize the maximum cross correlation between codewords.* It is noted that this problem has a longer history [93, 37] than the problem of noncoherent transmission, and yet no definitive answer has been given to the question of finding the signal constellation with minimum cross correlation for a given number of signals (codewords).

The case when codewords extend over several, say L, independent blocks of constant but otherwise unknown phase presents an additional difficulty. Although there is no simple expression for the pairwise error probability, the following bounds hold:

$$\prod_{\substack{k=1 \\ k:\rho_k \neq 1}}^{L} \mathcal{R}(\rho_k, N\gamma) \leq P(\mathbf{a} \to \mathbf{b}|\mathbf{a}) \leq \sum_{\substack{k=1 \\ k:\rho_k \neq 1}}^{L} \mathcal{R}(\rho_k, N\gamma), \tag{1.5}$$

where ρ_k is the cross correlation between the k-th blocks of codewords (sequences of length N). Combining these bounds with (1.3), we see that the error probability is dominated by the block with the highest ρ_k. Furthermore, the Hamming weight of an error event, *i.e.*, number of terms in the above sum, is also an important parameter. However, it is not clear how these two parameters are combined to form the pairwise probability of error. In view of these two observations, code design for this channel becomes more complicated.

We now outline a specific strategy that we will use to design powerful coding schemes. We follow a systematic approach by separating the code into two parts: a binary outer code and a modulation code. The task of the outer code is to introduce large memory into the overall code. This can be done by utilizing powerful error-correcting codes. In this work we concentrate on serially concatenated turbo TCM (SCTCM), and LDPC codes (the basic properties of these two families of powerful binary error-correcting codes are described in Appendix C). On the other hand, the purpose of the modulation code is to guarantee small cross correlations between the codewords, as well as to imitate the signaling structure suggested by the capacity results of the previous section. More specifically, the modulation code is responsible for introducing the zero-mass point, *i.e.*, no transmission in certain blocks. This layered approach is motivated by the conjecture reached in [55] that general turbo-like codes can achieve capacity for a wide range of channels if suitable modulation schemes are designed to interface with the channel. Within this context, two classes of codes are proposed:

- In the first one, a highly optimized modulation code is used which results in relatively high complexity (exponential with respect to N). Therefore, the overall design is attractive only for small values of N. However, it is demonstrated through a specific example that this modulation scheme, together with a simple outer code (SCTCM code with small complexity), results in performance close to capacity.

- In the second class, a novel modulation scheme is proposed which incorporates the transmission of the zero mass point and can be demodulated with a factor-graph-based algorithm with linear complexity with respect to N. This scheme

is well-suited for higher values of N and, as demonstrated through a specific example, when paired with irregular LDPC codes outperforms all existing designs.

1.2.4 Rotationally Invariant and Rotationally Robust Codes

Regardless of the particular structure of the coding scheme, a potential problem can occur when communicating over the noncoherent channel. This problem can be explained by observing that if two sequences of length N differ by a mere phase rotation, they will be indistinguishable at the receiver side. Therefore, such pairs should be avoided during the transmission as they will lead to a catastrophic behavior. On the other hand, if all such sequences are included in the codebook and are assigned the same input sequence, then the catastrophic behavior is avoided. This is true since no input errors will occur as a result of a random phase rotation. Coding schemes that resolve the inherent phase uncertainty in the transmitted sequence in this particular way are called rotationally invariant (RI), and are essentially generalizations of the differential phase encoding to higher signaling alphabets.

In Chapter V this property is investigated in more detail. Specifically, in the first part of Chapter V, we outline the design guidelines for RI-SCTCM codes. Following these guidelines, several powerful RI-SCTCM codes are designed and simulated. As it is illustrated through these examples, in some cases the RI property does not incur any performance loss, while for others it is obtained at the cost of reduced performance. In the second part, we introduce the notion of rotationally robust (RR) codes, namely, codes that are immune to arbitrary phase rotations that affect part of the codeword. In particular, RR codes are required to suffer only a finite number of input bit errors as a result of such rotation. Furthermore, we require the

input error pattern to be localized around the position of the phase jump and to be independent of the codeword length. We consider this property both in coherent and noncoherent settings, and prove that under certain conditions RI codes actually satisfy this stronger property. To extend this result to SCTCM codes we propose a simple modification of the decoding algorithm, that makes RI-SCTCM codes robust to phase jumps as well.

1.3 Dissertation Outline

The rest of the thesis is organized as follows. In Chapter II we design PSA schemes and analyze their performance. Chapter III investigates the information capacity and the structure of the capacity-achieving input density for the block-independent noncoherent channel. Inspired by the results of this chapter, in Chapter IV low-complexity coding and modulation schemes are proposed that are matched to the noncoherent channel characteristics. Chapter V discusses rotationally invariant codes, with emphasis on their design and robustness properties under different conditions. A summary of the current results as well as several possible avenues for future research are outlined in Chapter VI. Some technically involved parts of the thesis are collected in several appendices at the end. Appendix A derives the expressions for distribution functions required for analysis of the receivers proposed in Chapter II, while most of the proofs of Chapter III are collected in Appendix B. Appendix C gives a short introduction to powerful binary error-correcting codes, that are used throughout the thesis in the code design process. Finally, Appendix D derives a simple lower bound on the bit-error-rate performance of SCTCM codes, the results of which are used in Chapter V.

CHAPTER II

PILOT-SYMBOL-ASSISTED CODED TRANSMISSION

2.1 Introduction

In this chapter, pilot-symbol-assisted (PSA) coded transmission over the noncoherent channel is investigated. We are interested in the performance of a system that uses pilot symbols to aid the phase estimation in conjunction with codes primarily designed for the coherent AWGN channel. This approach can be motivated by observing that in the case of slow phase dynamics, a noncoherent code can be thought of as the combination of a training sequence (or pilot symbols) and a pure AWGN code. The role of the former is to facilitate phase estimation and effectively translate the noncoherent channel to a coherent AWGN channel.

In particular, powerful AWGN codes are utilized, resulting in an operating SNR close to the capacity of the coherent AWGN channel. They can be either parallel or serially concatenated codes [9, 5], collectively known as turbo codes, or low-density parity-check (LDPC) codes [29, 76]. These codes are augmented by inserting a single pilot symbol of specified power in the beginning of each block of length N, in order to aid the joint phase estimation and decoding process. As expected, the presence of a pilot symbol implies an inevitable loss in the transmitted power, resulting in a

trade-off between the power allocated between the pilots and the coded bits, and the quality of the joint phase estimation and decoding process.

Optimal, maximum aposteriori probability (MAP) receivers for the above described, high-performance codes are impractical due to their exponential complexity, even for the coherent AWGN channel. Although the presence of the unknown phase further complicates the design, several powerful iterative joint decoding and phase estimation algorithms have been suggested in the literature [31, 16, 1, 57]. Motivated by the need for simple, yet powerful, receivers, several iterative decoding algorithms are proposed in this work. These algorithms are broadly classified in two categories: (i) algorithms that implicitly estimate the phase by pre-processing the observation related to the pilot symbol only at the beginning of the iterative decoding process, and (ii) algorithms that perform joint decoding and phase estimation in an iterative fashion.

In this chapter we discuss several algorithms that fall into these categories, the main advantage of which is that they are simple enough to lend themselves to analysis. In particular, using density evolution, a recently developed technique for the analysis of LDPC codes over the AWGN channel, the performance of the proposed iterative receivers operating in the block-independent noncoherent channel is characterized. Specifically, assuming infinite iterations, and arbitrarily long codes, the minimum required information bit SNR, E_b/N_0, for achieving error-free communication is evaluated. In addition, the optimal power allocation to the pilot symbol is investigated, thus providing design guidelines for the PSA transmission scheme. It is found that this optimal power allocation depends highly on the channel coherence interval N, and the particular algorithm used for phase estimation/decoding. Furthermore, optimizing the power allocation results in performance close to that of the

coherent AWGN channel. This last result is obtained by deriving a simple bound on the performance of the MAP receiver for the block-independent noncoherent channel.

Since most of the codes proposed in the literature for this channel are based on differential encoding, the study of PSA codes can be motivated by the following three observations. First, for the particular model that we are interested in, *i.e.*, the block-independent noncoherent channel, it can be shown that differential encoding and PSA codes are exactly equivalent in terms of code rate and performance, both for maximum likelihood sequence detection and symbol-by-symbol maximum aposteriori probability detection. Second, for channels affected by a continuously varying phase process, it is not clear whether the $N/(N-1)$ rate advantage of differential encoding—when translated to energy requirements—is greater than the advantage of the proposed PSA schemes with optimized pilot symbol energy (assuming similar receiver complexity). And third, the results presented herein for the PSA schemes, suggest that a differential encoding scheme that periodically transmits a symbol with higher energy might result in increased performance over conventional differential encoding.

For simplicity of the presentation of the analytical tools used herein, binary phase-shift keying (BPSK) modulation is assumed in the first part of the chapter. However, numerical results presented in Section 2.5 consider other modulation schemes, and illustrate the applicability of the general ideas in this chapter to more elaborate modulation schemes, as well as to other generic receivers.

The rest of the chapter is structured as follows. Section 2.2 provides a description of the PSA transmission, and the structure of the optimal receiver. In Section 2.3 several approximate receivers are proposed, while their performance is evaluated in Section 2.4, using density evolution. In Section 2.5 numerical results for the

performance of the proposed receivers are presented using a concrete example of an LDPC code, while the concluding remarks are summarized in Section 2.6. Most of the derivations are collected in Appendix A for smoothness of presentation.

2.2 Pilot-symbol-assisted Transmission

Recall the channel input/output relationship

$$\mathbf{z}_k = \mathbf{y}_k e^{j\theta_k} + \mathbf{n}_k, \tag{2.1}$$

where $\mathbf{y}_k$, $\mathbf{z}_k$ and $\mathbf{n}_k$ are N-dimensional complex vectors, denoting the k-th block ($k = 1, 2, \ldots, L$) of the transmitted, the received and the noise sequence of independent identically distributed (i.i.d.), zero-mean, circular, complex Gaussian random variables with variance $\sigma^2 = N_0/2$ per real dimension, respectively. The unknown phase rotation in each block is represented by the variables θ_k, which are modeled as i.i.d. random variables uniformly distributed in $[-\pi, \pi)$.

Pilot-symbol-assisted transmission is considered; specifically, the first input symbol in each block is the pilot symbol and the remaining $N - 1$ symbols are coded bits. More precisely, we have

$$\mathbf{y}_k = [\sqrt{E_p}, \sqrt{E_s}\mathbf{x}_k^T]^T, \tag{2.2}$$

where superscript T denotes transpose, with $\mathbf{x}_k \in \{+1, -1\}^{N-1}$ being the k-th block of the codeword, and E_p and E_s denote the pilot and coded bit energy, respectively. Insertion of a pilot symbol decreases the throughput to $R_t = (N-1)R/N$ (information bits per complex dimension), where R is the code rate, and increases the required information bit SNR to

$$\frac{E_b}{N_0} = \frac{1}{R}\left(\frac{E_s}{N_0} + \frac{1}{N-1}\frac{E_p}{N_0}\right). \tag{2.3}$$

This simple relation shows the trade off associated with the pilot energy, namely, quality of phase estimation is gained at the expense of increased information bit SNR.

The optimal detection rule, in the sense of minimizing the symbol error rate for this system is the symbol-by-symbol MAP detection rule, which assuming equally likely codewords can be written as

$$\hat{a}_j = \arg\max_{a_j} p(a_j|\mathbf{z}_1, \ldots, \mathbf{z}_L) \tag{2.4a}$$

$$= \arg\max_{a_j} \sum_{\substack{\mathbf{x}:a_j \\ \mathbf{x}\in\mathcal{C}}} \prod_{k=1}^{L} \int_0^{2\pi} \exp\Big(\frac{\mathrm{Re}(\mathbf{y}_k^T \mathbf{z}_k e^{-j\theta_k})}{\sigma^2}\Big) d\theta_k \tag{2.4b}$$

$$= \arg\max_{a_j} \sum_{\substack{\mathbf{x}:a_j \\ \mathbf{x}\in\mathcal{C}}} \prod_{k=1}^{L} I_0\left(\frac{|\mathbf{y}_k^T \mathbf{z}_k|}{\sigma^2}\right), \tag{2.4c}$$

where $\mathcal{C}$ is the overall codebook, $I_0(\cdot)$ represents the 0th order modified Bessel function of the first kind, and the expression $\mathbf{x} : a_j$ denotes all codewords $\mathbf{x}$ whose corresponding input sequence has j-th element a_j. Since, we are interested in using and analyzing PSA modulation in conjunction with LDPC codes, and since the latter are conveniently described on factor graphs [28], it is helpful to derive a graphical representation of this transmission scheme (refer to Appendix C for a brief introduction of LDPC codes and factor graphs). This system can be represented by factor graphs in two different ways. In the first model, considered before in [94, 96], the unknown phase variables θ_k are included explicitly in the factor graph, depicted in Fig. 2.1(a), which corresponds to the expression in (2.4b). Here, variables C and X represent check and variable nodes, respectively and the function nodes f_i correspond to the factors $f_i(x_i, \theta) = \exp[\mathrm{Re}(\sqrt{E_s} x_i z_i e^{-j\theta})/\sigma^2]$. When such explicit modeling is used, the message alphabet is infinite, and thus, in practice, some quantization of these messages is necessary [96] (for additional discussion of factor-graph codes over the

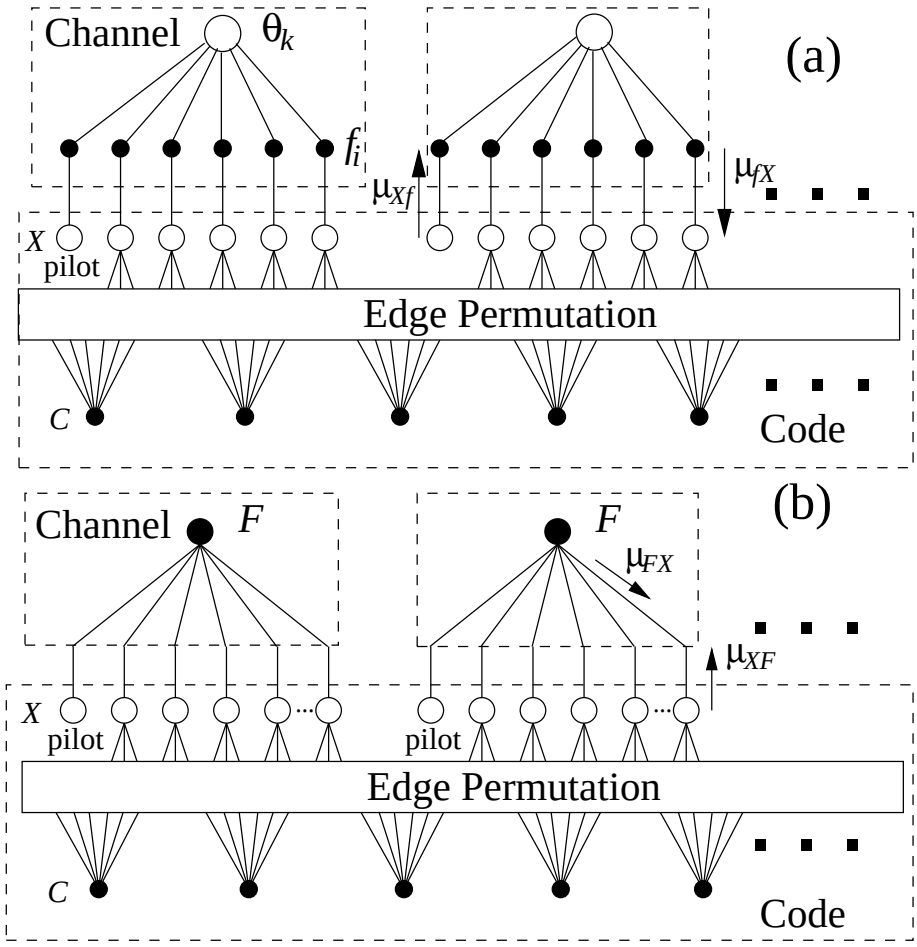


Figure 2.1: Two factor graphs corresponding to the optimal receiver: (a) unknown phase variables are explicitly modeled; (b) unknown phase variables have been averaged out. The "edge permutation" block describes the connections between variable and check nodes.

Rayleigh fading channel please refer to [13] and to Appendix C). The model assumed herein corresponds to the expression in (2.4c), where averaging with respect to θ_k has been explicitly performed, evidenced by the presence of the $I_0(\cdot)$ function, rather than having it done iteratively as part of the sum-product algorithm. The factor graph corresponding to this model is shown in Fig. 2.1(b), where the functional node F represents the channel constraint, *i.e.*, the $I_0(\cdot)$ term in (2.4c).

In the absence of cycles in the graph of Fig. 2.1(b), the expression in (2.4c) can be exactly evaluated by applying the sum-product algorithm on the corresponding

factor graph [94]. The messages exchanged by the sum-product algorithm at the variable and check nodes are given, in the logarithmic domain, by (see [77])

$$\mu_{XC} = \sum_{i=1}^{d_v-1} \mu_{C_iX} + \mu_{FX} \tag{2.5a}$$

$$\mu_{CX} = 2\tanh^{-1}\left(\prod_{j=1}^{d_c-1} \tanh\left(\frac{\mu_{X_jC}}{2}\right)\right), \tag{2.5b}$$

where μ_{XC} and μ_{CX} represent variable-to-check and check-to-variable node messages, with d_v+1 and d_c being the number of edges connected to variable and check nodes, respectively, and $\{C_i\}_{i=1}^{d_v-1}$ ($\{X_j\}_{j=1}^{d_c-1}$) is the set of check (variable) nodes connected to variable node X (check node C), other than the check node C (variable node X)[1]. To simplify notation the index k is dropped in the rest of the chapter with the understanding that all the involved variables belong to the same block.

The messages to and from the node F have the form

$$\mu_{X_iF} = \sum_{j=1}^{d_v} \mu_{C_jX_i} \tag{2.5c}$$

$$\mu_{FX_i} = \log\frac{L_i(+1)}{L_i(-1)} \tag{2.5d}$$

$$L_i(a) = \sum_{\mathbf{x}:x_i=a} I_0\Big(\frac{|\mathbf{y}^T\mathbf{z}|}{\sigma^2}\Big) \prod_{\substack{j=1,j\neq i\\ j:x_j=1}}^{N-1} e^{\mu_{X_jF}}, \tag{2.5e}$$

where, μ_{FX_i} and μ_{X_iF} are F-to-variable and variable-to-F node messages (refer to [42] for a detailed discussion of the sum-product algorithm). In the following the code length is assumed big enough so that the length of the smallest cycle present in the overall factor graph is at least twice as much as the number of iterations required, thus the application of the sum-product algorithm results in the exact evaluation of (2.4c).

---

[1]For notational simplicity a regular LPDC code is considered in the analysis part of the chapter. The generalization to irregular LDPC codes and/or other iteratively decoded codes is straightforward. Irregular LDPC codes are considered in the examples.

Since both *pilot and data* (PD) symbols are utilized for the generation of messages, and since observation model, this optimal receiver, referred to as M-PAD, will be our benchmark receiver for all subsequent comparisons. However, the benchmark receiver for all subsequent comparisons. However, the exponential (in N) complexity associated with the implementation of (2.5e), resulting from the summation over all sequences $\mathbf{x} : x_i = a$, renders this algorithm unattractive even for moderate values of N. This obstacle is the main motivation for seeking suboptimal receiver schemes.

2.3 Proposed Receivers

Several practical suboptimal receivers are proposed in this section. These receivers are based on the sum-product algorithm, and are only modifying the phase estimation part of the receiver, *i.e.*, the messages μ_{FX_i}. We start by describing receivers that implicitly estimate the phase by pre-processing the observation related to the pilot symbol only at the beginning of the iterative process, and then proceed with decoding; thus phase estimation is performed in a non-iterative manner. We then propose an iterative receiver, which performs adaptive phase estimation and decoding, and yet is simple enough to lend itself to analysis using density evolution.

2.3.1 Receivers with non-iterative phase estimation

Receivers in this section initialize the messages μ_{FX_i} based only on the observation z_i and the pilot variable z_0. Since the messages μ_{XF} are ignored by these receivers, the iterative processing is performed only in the lower part of the corresponding factor graph, which is depicted in Fig. 2.2. This is equivalent to applying the sum-product algorithm to a hypothesized memoryless channel which takes as an input the coded bit x_i and outputs two symbols: the observation due to the pilot symbol, z_0, and the observation due to the data bit, z_i. All three receivers discussed in this

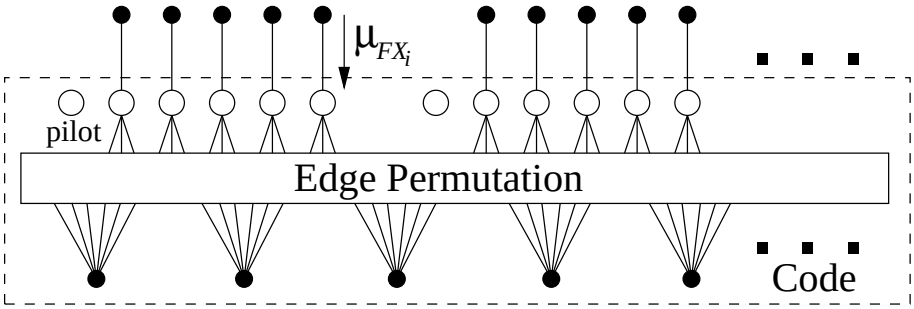


Figure 2.2: Factor graph corresponding to the PO receivers. Observe that this receiver model is matched to a memoryless channel with input x_i and outputs z_0 and z_i only.

section differ only in the initial message μ_{FX_i}, and will be referred to as *pilot only* (PO) receivers.

The initial message implied by the sum-product algorithm when the only variables observed are z_0 and z_i is given by

$$\mu_{FX_i} = \log \frac{p(z_i, z_0|x_i = +1)}{p(z_i, z_0|x_i = -1)} = \log \frac{I_0(|z_0\sqrt{E_p} + z_i\sqrt{E_s}|/\sigma^2)}{I_0(|z_0\sqrt{E_p} - z_i\sqrt{E_s}|/\sigma^2)}, \tag{2.6}$$

where $p(z_i, z_0|x_i = a)$ is the joint density of z_0 and z_i conditioned that the i-th transmitted symbol is a (see expression (A.2) in the Appendix). Because this receiver scheme is *matched* to the observation model, it will be referred to as M-PO receiver. It is worth noting that for large values of E_p, the message in (2.6) approaches $2\mathrm{Re}(\sqrt{E_s}z_i e^{-j\theta})/\sigma^2$, which is exactly the initial message in the coherent AWGN channel.[2] As a result, it is expected that the asymptotic behavior of this suboptimal receiver approaches that of the coherent AWGN channel for large values of pilot energy.

A slightly modified version of the above receiver is one that pre-processes the pilot observation z_0 and generates an estimate $\hat{\theta}$ on θ. This estimate is then used

[2]For large of values of E_p, the approximations $I_0(x) \cong e^x/\sqrt{2\pi x}$ and $\angle z_0 \cong \theta$ can be used. The desired result is a consequence of the above two approximations together with the fact that $|A + z| - |A - z| \cong 2\mathrm{Re}\{ze^{-\angle A}\}$ for two complex numbers A, z with $|A| \gg |z|$.

for generating the μ_{FX_i} messages. This approach, which essentially derives from generalized likelihood [71, ch. II] results in the message

$$\mu_{FX_i} = \log \frac{p(z_i, \hat{\theta}|x_i = +1)}{p(z_i, \hat{\theta}|x_i = -1)} = \log \frac{\int_{-\pi}^{\pi} e^{\mathrm{Re}(\sqrt{E_s} z_i e^{-j\theta})/\sigma^2} p(\hat{\theta}|\theta) d\theta}{\int_{-\pi}^{\pi} e^{-\mathrm{Re}(\sqrt{E_s} z_i e^{-j\theta})/\sigma^2} p(\hat{\theta}|\theta) d\theta}. \tag{2.7}$$

where $p(\hat{\theta}|\theta)$ is the conditional density of the estimated phase $\hat{\theta}$ given the random channel phase θ. For maximum-likelihood phase estimation, $\hat{\theta}$ is simply the angle of z_0. This particular receiver that uses an *estimate* of the unknown parameter based solely on the pilot symbol, will be referred to as E-PO. Since in this initialization, the amplitude of the pilot symbol, and hence the reliability information of the pilot, is disregarded, it is expected that its performance will be inferior to the performance of the scheme in (2.6). However, one advantage of this scheme is that it can be easily extended to the case when the pre-processing step utilizes not only the pilot symbol, but the entire observation $\mathbf{z}$. Indeed, we observe that (2.7) holds for any estimate $\hat{\theta}$ of θ, as long as it is conditionally independent of z_i given θ. Therefore, if $\hat{\theta}$ is any estimate that depends on all the symbols of the sequence $\mathbf{z}$ except z_i, then (2.7) can be used, resulting in the E-PD receiver, as long as the conditional distribution $p(\hat{\theta}|\theta)$ for the phase estimator can be derived.

Finally, we also present a receiver similar to E-PO which simply derotates the observation by the phase estimate, and forms the message corresponding to the coherent AWGN channel, *i.e.*,

$$\mu_{FX_i} = \frac{2\sqrt{E_s}}{\sigma^2} \mathrm{Re}(z_i e^{-j\hat{\theta}}), \tag{2.8}$$

where again $\hat{\theta}$ can be any estimate of θ which does not depend on z_i. To emphasize the fact that the observation is only *derotated* in this case, the corresponding receivers are referred to as EDR-PO and EDR-PD, when the phase estimate is formed by the pilot only, or the pilot and data, respectively.

2.3.2 Receivers with Iterative Estimation

The natural extension of the PO schemes described above is a receiver that performs both phase estimation and decoding in an iterative fashion. However, it is desirable that such receiver has only linear complexity with the channel coherence interval N. In this section, a suboptimal receiver is described, that not only is simple to implement, but also lends itself to analysis using density evolution. The basic idea is that the μ_{XF} messages with large positive or negative values indicate an almost certain decision on the corresponding coded bit. This hard decision can be utilized in the phase estimator, by essentially treating this bit as an additional pilot symbol, resulting in higher quality μ_{FX} messages. Furthermore, in the proposed algorithm, the coded bits corresponding to messages with small values are not processed in the generation of the μ_{FX} messages. Similar decision feedback techniques for improving noncoherent decoding were proposed in [47, 22, 43].

The above idea is made precise in the following. Let $\mathcal{Q}(\cdot)$ be the *quantized decision* mapping

$$\mathcal{Q}(x) = \begin{cases} +1 & , x > A^{(l)} \\ -1 & , x < -A^{(l)} \\ 0 & , \text{otherwise} \end{cases} \tag{2.9}$$

for some predefined sequence of thresholds $A^{(l)}$, where the superscript (l) denotes possible iteration-dependent thresholds, and let $\mathbf{v} = [\sqrt{E_p}, v_1, \ldots v_{N-1}]^T$, with $v_j = \sqrt{E_s}\mathcal{Q}(\mu_{X_jF})$, be the quantized decision vector. The message μ_{FX_i} is defined as[3]

$$\mu_{FX_i} = \log \frac{I_0(|w_i a^{(l)} + z_i\sqrt{E_s}|/\sigma^2)}{I_0(|w_i a^{(l)} - z_i\sqrt{E_s}|/\sigma^2)}, \tag{2.10}$$

[3]A more "reasonable" message would be the one similar to (2.5e) with messages μ_{X_iF} replaced with their hard quantized versions. However, this choice would be considerably harder to analyze.

where

$$w_i = \frac{(\mathbf{v}^i)^T \mathbf{z}^i}{\|\mathbf{v}^i\|}, \tag{2.11}$$

and $a^{(l)}$ is some predefined sequence of "weights", signifying how much the receiver relies on the information provided by the variable w_i. Superscript i over a vector variable denotes that the i-th variable in that vector is absent. The message in (2.10) is a direct extension of the message used in the M-PO receiver, where now coded symbols with significant bias are treated as pilots as well. In fact, if the initial value for $a^{(0)}$ is $\sqrt{E_p}$, the exact equation (2.6) for the M-PO receiver is obtained. Although one might choose to optimize the values of the thresholds $A^{(l)}$, and weights $a^{(l)}$, numerical results indicated an insignificant performance loss, when these values were assumed constant over iterations[4]. For this reason, constant values will be assumed in the rest of the chapter, with $a^{(l)} = \sqrt{E_p}$, $A^{(l)} = A$, and the value of A is chosen so that $P(\mu_{Fx_i} < 0)$ is minimized after the first iteration. Because this receiver performs *quantized decision feedback*, it is referred to as M-QDF receiver. Although not shown here, this same idea could be applied to equations (2.7) and (2.8) by taking the phase estimate $\hat{\theta}$ to be the angle of w_i, to obtain E-QDF and EDR-QDF receivers, respectively.

It is noted that the M-QDF receiver can be further generalized to the case when the quantization has more than three levels, thus obtaining a family of receivers with increasing performance; this direction is not pursued further as it does not give any conceptually different receiver designs.

A summary of all the proposed receivers is given in Table 2.1, with the performance increasing from left to right and from down up.

---

[4]Interestingly enough, the "optimal" value for $A^{(l)}$ was found to be constant over iterations.

	PO	QDF	PD	CDF
M	eq. (2.6)	eq. (2.10)	eqs.(2.4c),(2.5)	eq. (2.23)
E	eq. (2.7)	eq. (2.7)[a]	eq. (2.7)[b]	eq. (2.7)[c]
EDR	eq. (2.8)	eq. (2.8)[a]	eq. (2.8)[b]	eq. (2.8)[c]

[a]with $\hat{\theta}$ taken as angle of w_i of eq. (2.11)

[b]with $\hat{\theta}$ obtained using $\mathbf{z}^i$

[c]with E_p replaced by $E_{p,\text{eff}}$ of eq. (2.25)

Table 2.1: Summary of Proposed Receivers

2.4 Performance Analysis using Density Evolution

In this section, performance analysis of the proposed receivers is done using density evolution techniques. Density evolution is a powerful tool for analysis of various message passing algorithms, and in particular the sum-product algorithm. It evaluates the performance of the receiver when averaged over the ensemble of codes with common parameters by keeping track of the densities of the messages transmitted in the graph during decoding. It was proven that for large blocklength, performance of almost all the codes in the ensemble approaches this average behavior, and therefore results obtained using density evolution should predict the true performance of the particular code from that ensemble.

In addition to the regular assumptions for the density evolution analysis, in the following we also assume that the code is big enough and sufficiently interleaved, so that any two variable nodes in the same block are connected only through the functional F-node, and thus messages arriving at any variable node could be considered independent. For coherent channel it is proven that the performance of the sum-product algorithm is independent of the transmitted codeword and therefore the all-one codeword can be assumed which enables analysis through density evolu-

tion [77]. It turns out that this property holds for the noncoherent channel for all the receivers considered herein. First we show that this is true for the case of optimal sum-product decoding (*i.e.*, when the message update rules of (2.5e) are used at the F-node of the factor graph in Fig. 2.1(b)).

Let $\mathbf{z}$ be the output of the channel when the transmitted codeword is $\mathbf{x} = (s_1, \ldots, s_{N-1})$. This channel can be represented by an equivalent multiplicative channel as $\mathbf{z} = U\mathbf{z}'$, where $\mathbf{z}'$ is the output of the noncoherent channel when the all-one codeword is sent, and U is a diagonal matrix with entries $(1, s_1, \ldots, s_{N-1})$ on the main diagonal. Let μ, μ' represent the messages that are interchanged in the graph when the codeword $\mathbf{x}$, and the all-one codeword is sent, respectively. Also assume that the incoming messages to node F are related as $\mu_{X_iF} = s_i\mu'_{X_iF}$. If we prove that at any iteration, the outgoing messages from node F satisfy $\mu_{FX_i} = s_i\mu'_{FX_i}$ then, exactly as in [77], we can deduce that all the messages interchanged in the graph will satisfy this as well, including the ones on which the hard decisions will be made, thus proving our claim. By changing the variable of summation in (2.5e) to $\mathbf{x}' = U'\mathbf{x}$, where U' is a diagonal unitary matrix with $(s_1, \ldots, s_{N-1})$ on the main diagonal, we get

$$L_i(a) = \sum_{\mathbf{x}:x_i=a} I_0\Big(\frac{|\mathbf{y}^T\mathbf{z}|}{\sigma^2}\Big) \prod_{\substack{j=1,j\neq i\\ j:x_j=1}}^{N-1} e^{\mu_{X_jF}} = \sum_{\mathbf{x}':x'_i=as_i} I_0\Big(\frac{|\mathbf{y}'^T\mathbf{z}'|}{\sigma^2}\Big) \prod_{\substack{j=1,j\neq i\\ j:x'_j=s_j}}^{N-1} e^{s_j\mu'_{X_jF}} \tag{2.12}$$

$$= \sum_{\mathbf{x}':x'_i=as_i} I_0\Big(\frac{|\mathbf{y}'^T\mathbf{z}'|}{\sigma^2}\Big) \prod_{\substack{j=1,j\neq i\\ j:x'_j=1}}^{N-1} e^{\mu'_{X_jF}} \left(\prod_{\substack{j=1,j\neq i\\ j:s_j=-1}}^{N-1} e^{-\mu'_{X_jF}}\right) = L'_i(as_i) \prod_{\substack{j=1,j\neq i\\ j:s_j=-1}}^{N-1} e^{-\mu'_{X_jF}}, \tag{2.13}$$

where $\mathbf{y}' = U\mathbf{y}$. The first equality is the definition (2.5e) and the second equality follows by observing that the set over which the summation is performed after the change of variables becomes $\{U'\mathbf{x}|x_j = a\} = \{\mathbf{x}'|x'_j = as_j\}$. Therefore, the message

from node F to node X_i becomes

$$\mu_{FX_i} = \log \frac{L_i(+1)}{L_i(-1)} = \log \frac{L'_i(s_i)}{L'_i(-s_i)} = s_i \log \frac{L'_i(+1)}{L'_i(-1)} = s_i \mu'_{FX_i}. \tag{2.14}$$

Since at the first iteration messages μ_{X_iF} are taken to be zero, this relationship is true for the initial messages as well, and therefore all the messages during the decoding will undergo related sign changes depending on the transmitted codeword and so will the hard decisions made at the end of the decoding process.

In order to apply density evolution, the densities of the initial messages for the proposed receivers need to be evaluated. Since the problem is considerably different for PO and QDF receivers two cases are considered separately.

2.4.1 PO receivers

For the case of PO receivers, the μ_{FX} messages are evaluated only in the beginning, and are repeatedly used in subsequent iterations. For all three PO receivers presented herein, the initial messages satisfy (2.14), and therefore it is adequate to assume the transmission of the all-one word.

For the M-PO receiver, observe that the variables z_i and z_0 have a known joint cumulative distribution function (cdf), from which it is straightforward to get the cdf of the message in (2.6). The derivation of the M-PO message cdf is presented in Section A.1 of Appendix A.

For the E-PO receiver, first the conditional density of the phase estimate $p(\hat{\theta}|\theta) = T(\hat{\theta} - \theta)$ for the case of ML phase estimate $\hat{\theta} = \angle z_0$ is evaluated as

$$T(x) = \frac{1}{2\pi} e^{-\frac{E_p}{2\sigma^2}} + \sqrt{\frac{E_p}{2\pi\sigma^2}} \cos(x) e^{-\frac{E_p}{2\sigma^2}\sin^2(x)} \times \frac{1}{2}\Big(1 + \operatorname{erf}\Big(\sqrt{\frac{E_p}{2\sigma^2}} \cos(x)\Big)\Big) \tag{2.15}$$

where $\operatorname{erf}(\cdot)$ is the error function (see [47] for a detailed discussion on $T(x)$). Then,

the expression in (2.7) is rewritten in terms of $r = |z_i|$ and $t = \angle z_i - \hat{\theta}$ as

$$\mu_{FX}(r,t) = \log \frac{\int_{-\pi}^{\pi} e^{\sqrt{E_s} r \cos(t-x)/\sigma^2} T(x) dx}{\int_{-\pi}^{\pi} e^{-\sqrt{E_s} r \cos(t-x)/\sigma^2} T(x) dx}. \tag{2.16}$$

By observing that $T(x)$ is an even function, one can verify that the expression in (2.16) is an even function of t, and is odd symmetric around point $t = \pi/2$ (*i.e.*, $\mu_{FX}(r,t) = -\mu_{FX}(r,\pi - t)$). Therefore, the Fourier expansion in t will only consist of cosine terms with odd frequencies. This function is very well approximated by the first term in this expansion, $h(r)\cos(t)$ where $h(r) = \mu_{FX}(r,t)|_{t=0}$, with a relative error less than 10^{-5}, for the E_p values considered herein. With this approximation, it is straightforward to derive the distribution of the E-PO message, which is given in Section A.2 of Appendix A.

Finally, the distribution of the EDR-PO receiver can be easily derived noting that conditioned on θ, the message is Gaussian, resulting in

$$P(\mu_{FX} \leq q) = \int_{-\pi}^{\pi} \frac{1}{2} \operatorname{erfc}\left(\frac{\sigma^2 q - 2\cos(x)}{2\sqrt{2}\sigma}\right) T(x) dx, \tag{2.17}$$

with $T(x)$ defined in (2.15).

2.4.2 QDF receivers

Because the messages μ_{FX_i} change with iterations, density evolution for this receiver scheme is considerably different from that for PO schemes. However, an important observation regarding the distribution of the message in (2.10) makes density evolution feasible for this receiver. The idea is that the cdf of the M-QDF message in (2.10) can be written as the average of some distribution functions which do not change with iterations. These distribution functions can be evaluated before hand, thus making density evolution tractable.

To start with, the cdf of the message in (2.10) can be found by first observing

that the following relationship holds (with $q_i = \mu_{FX_i}$)

$$P(q_i < q|\mathbf{x}) = \sum_{\mathbf{v}^i} P(q_i < q|\mathbf{x}, \mathbf{v}^i) p(\mathbf{v}^i) \tag{2.18}$$

$$= \sum_{n_+=0}^{N-2} \sum_{n_-=0}^{N-2-n_+} P(q_i < q|\mathbf{x}, n_+, n_-) p(n_+, n_-) \tag{2.19}$$

$$p(n_+, n_-) = \binom{N-2}{n_+, n_-} p_+^{n_+} p_-^{n_-} p_0^{n_0} \tag{2.20}$$

$$p_+ = P(\mu_{XF} > A^{(l)}) \tag{2.21}$$

$$p_- = P(\mu_{XF} < -A^{(l)}), \tag{2.22}$$

where $1+n_+$ and n_- denote number of positive and negative terms in $\mathbf{v}^i$, respectively, $n_0 = N-2-n_+-n_-$ and $p_0 = 1-p_+-p_-$. Since the weights are assumed constant (*i.e.*, $a^{(l)} = a = \sqrt{E_p}$), the only variables in this equation that change with iterations are p_+ and p_-. Therefore, the expressions $P(q_i < q|\mathbf{x}, n_+, n_-)$ can be evaluated off-line and stored for use in subsequent iterations[5].

Furthermore, due to the similarity of the M-QDF message in (2.10) with the messages for the M-PO receiver, the same approach can be followed for the evaluation of these conditional cdfs, which are given in Section A.3 of Appendix A. To prove that the probability of error is independent of the transmitted codeword, observe that the quantization mapping is an odd function $\mathcal{Q}(-x) = -\mathcal{Q}(x)$, and consequently w_i in (2.11) is independent of the transmitted codeword. Therefore, the relationship in (2.14) holds for the M-QDF receiver as well.

2.4.3 Correct Decision Feedback Bound

It is customary in decision-feedback systems to use the so-called "genie-aided" receiver to obtain upper bounds on the performance of the analyzed receiver (for

[5] $N(N-1)/2$ distribution functions need to be stored.

instance, see [22, 43] and references therein). In this section we apply this idea to the benchmark receiver by considering the following hypothetical system. In particular, it is assumed that a "genie" corrects all the messages μ_{XF} (*i.e.*, $\mu_{XF} = +\infty$, assuming the all-ones word is transmitted), so that these messages provide a perfectly correct decision to the F nodes. As a result, those messages do not change with iterations and neither do the messages μ_{FX} from the node F to the variable nodes. This implies that the hypothetical receiver under consideration is similar to the M-PO receiver, with the initial messages μ_{FX_i} evaluated as (compare with (2.10))

$$\mu_{FX_i} = \log \frac{p(\mathbf{z}|\mathbf{x}^i, x_i = +1)}{p(\mathbf{z}|\mathbf{x}^i, x_i = -1)} = \log \frac{I_0(|p_i\sqrt{E_{p,\text{eff}}} + z_i\sqrt{E_s}|/\sigma^2)}{I_0(|p_i\sqrt{E_{p,\text{eff}}} - z_i\sqrt{E_s}|/\sigma^2)}, \tag{2.23}$$

where

$$p_i = \frac{\mathbf{y}^{iT}\mathbf{z}^i}{\sqrt{E_{p,\text{ eff}}}}. \tag{2.24}$$

Intuitively, in this hypothetical receiver, all coded bits, other than the ith bit can be considered as pilot symbols, resulting in an effective pilot symbol energy

$$E_{p,\text{eff}} = \|\mathbf{y}^i\|^2 = E_p + (N-2)E_s. \tag{2.25}$$

It can be seen from (2.24) that p_i is a Gaussian random variable with mean $e^{j\theta}\sqrt{E_{p,\text{eff}}}$ and variance σ^2, conditioned on θ. Thus, the message in (2.23) will have exactly the same distribution as the M-PO message in (2.6) with $E_{p,\text{eff}}$ substituted for E_p. This receiver, which takes advantage of the *correct decision feedback* given by the hypothetical "genie", will be referred to as the M-CDF receiver.

2.5 Numerical Results

In this section the performance of the proposed receivers is presented for a concrete example of a rate 1/2 LDPC code with parameters $d_v = 3$ and $d_c = 6$. Discretized density evolution is considered, which corresponds to the exact performance

of a receiver operating with quantized messages in the sum-product algorithm [15]. A uniform quantizer was used with 127 levels[6] in the range from -40 to 40, so that the cell size was 80/127=0.63. In our implementation of density evolution reaching a BER of less than 10^{-30} in at most 1000 iterations, was considered as achieving error-free communication.

For each pair of E_p and σ (the normalization $E_s = 1$ was used), density evolution was run to check if the BER converged to zero. The highest σ for which BER convergence occurred was reported as the σ threshold for that value of E_p. For the case of PO receivers since the messages μ_{FX} do not depend on N, the same density evolution results are used for different values of N, where E_b/N_0 is calculated from (2.3).

In the case of M-QDF receiver, the messages μ_{FX} change with iterations and depend on N. As a result, separate density evolution runs should be performed for different values of N. Moreover, the number of initial distribution functions that need to be evaluated is quadratic in N. A considerable reduction in complexity was achieved by observing that the set of pairs of non-negative numbers $B_N = \{(n_+, n_-)|0 \leq n_+ + n_- \leq N-2\}$ is included in the set $B_{N'}$ for any $N' > N$. This fact implies that the set of distribution functions that need to be evaluated and stored for N' includes all the distribution functions already calculated for N. Therefore, for fixed σ and E_p, when increasing the channel coherence interval from N to $N+1$, only N additional distribution functions need to be evaluated, as opposed to $N(N-1)/2$. This observation was the basis for the numerical results for M-QDF receiver presented herein. In particular, the values of σ and E_p were fixed and density

[6]It is noted however, that discretized density evolution with this resolution is not sufficient to obtain accurate threshold values for a receiver using *unquantized* messages.

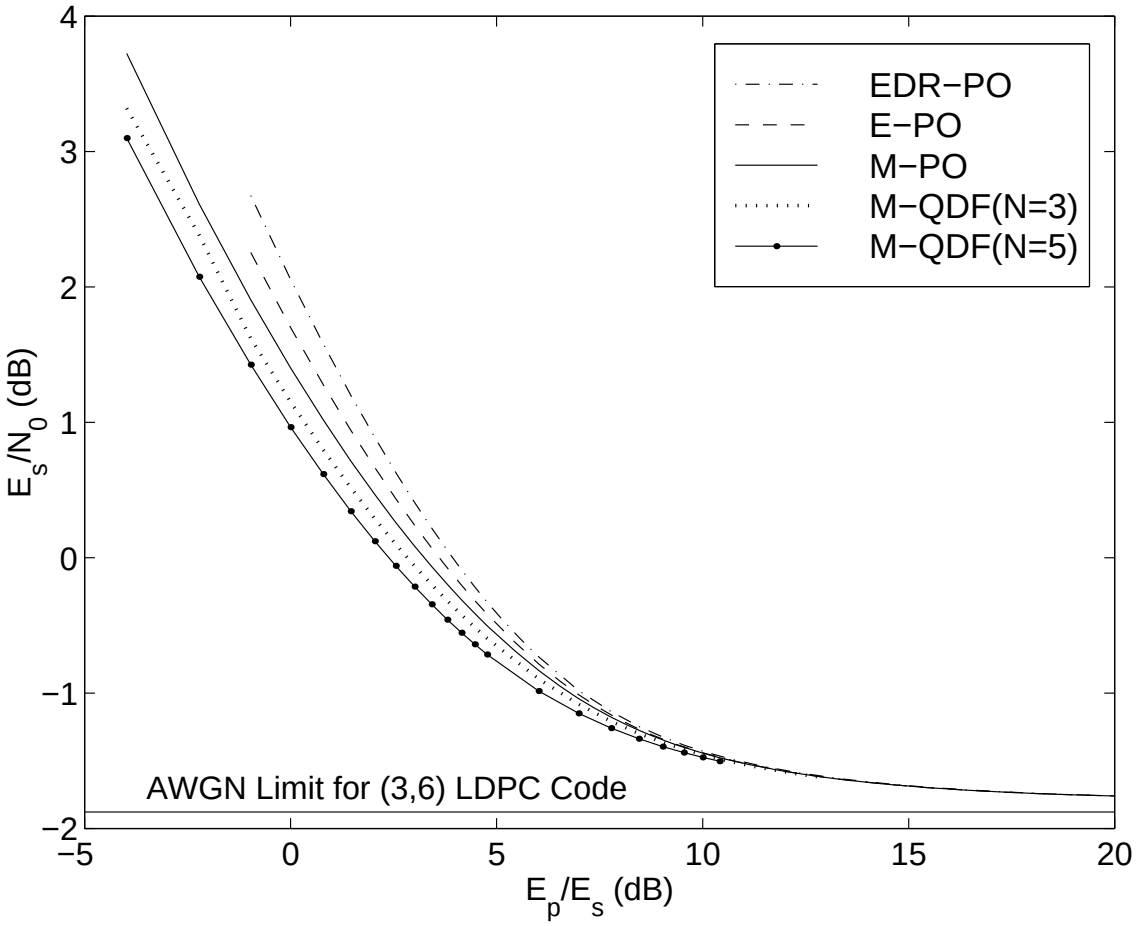


Figure 2.3: Minimum required E_s/N_0 versus E_p/E_s.

evolution was performed for increasing values of N until BER convergence occurred, say for $N = N^*$. The resulting triplet of values (E_p, σ, N^*) is achievable, thus the corresponding E_b/N_0 can only be viewed as an upper bound on the minimum achievable E_b/N_0 (because of the suboptimal power allocation, that was specified before hand). Repeating this procedure for several pairs E_p and σ, a set of achievable points is obtained. Moreover, since the performance increases with N, an achievable triplet of values (E_p, σ, N^*) implies that the triplet (E_p, σ, N') is also achievable for $N' > N^*$. Thus a set of E_b/N_0 values can be obtained through the use of equation (2.3) for all $N' > N^*$.

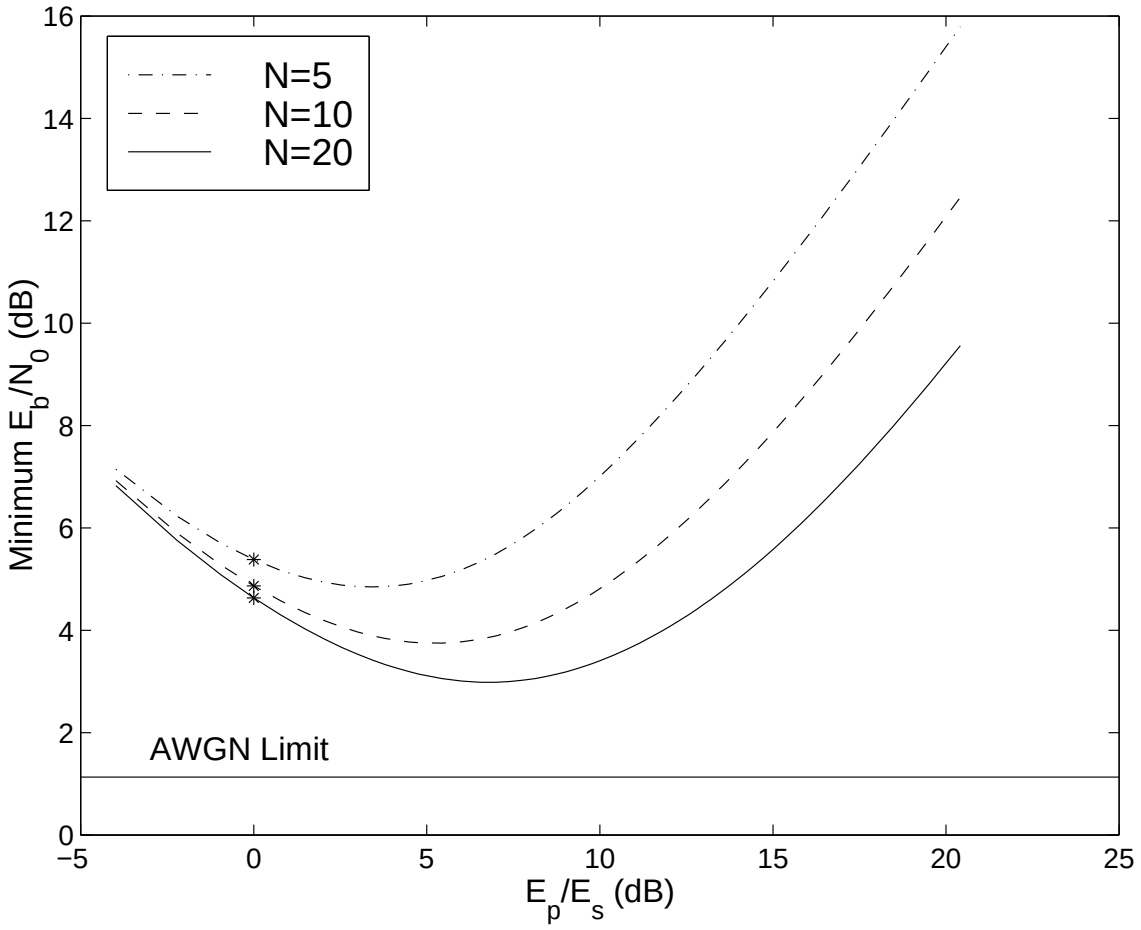


Figure 2.4: Minimum E_b/N_0 required versus E_p/E_s for M-PO receiver for different values of N.

Fig. 2.3 shows the minimum coded bit SNR, E_s/N_0, that is required versus the pilot to symbol power ratio, E_p/E_s, for M-QDF and all three PO receivers. For comparison, the horizontal line indicates the threshold for the coherent AWGN receiver (sum-product algorithm) [77]. The convergence of these curves to the performance of AWGN receiver with increasing E_p/E_s can be observed from the figure. Furthermore, the performance of the M-QDF receiver is increased with increasing N, and is consistently better than that of the PO receivers.

In Fig. 2.4, the tradeoff associated with the power allocation to the pilot symbol is exemplified for the M-PO receiver. This graph depicts the overall E_b/N_0 required

for error-free transmission versus E_p/E_s for several values of N. The points denoted by stars on the curves correspond to E_b/N_0 that would be required if non-optimal $E_p = E_s$ power allocation were used. The gain over these points achieved by using the optimal power allocation is considerable, thus proving the importance of power allocation (*e.g.*, a gain of 1.5 dB at $N = 20$ is observed). Also observe that optimal power allocation happens somewhere between 3-6 dB, which means that pilot energy E_p should be roughly from 2 to 4 times the coded symbol energy E_s.[7] We note that the curves for different N were produced using the single curves in Fig. 2.3, combined with (2.3).

The behaviour of the SNR gains as a result of pilot power optimization for different N and receivers is depicted in Fig. 2.5. Minimum required E_b/N_0 for these receivers is drawn versus the channel coherence interval, together with minimum required E_b/N_0 when $E_p = E_s$ is used. As can be seen from the figure, optimizing the pilot energy results in 0.8 dB gain at $N = 10$ (1.2 dB at $N = 20$) for the M-QDF receiver, while for the M-PO receiver the corresponding gains are 1 dB and 1.8 dB.

Fig. 2.6 compares the performance of M-PO and M-QDF receivers to the unconstrained and modulation-constrained capacities of this channel. Minimum required E_b/N_0 for these receivers is drawn versus the channel coherence interval. Also depicted in this figure is the E_b/N_0 curve for the hypothetical M-CDF receiver (Section 2.4.3) and the noncoherent BPSK capacity from [68], which together constitute a lower bound on the performance of all the above receivers including the benchmark receiver. It is noted that the performance of the sum-product receiver described in

---

[7] One can also distribute the pilot energy $E_p = mE_s$ into m symbols, each of energy E_s. This might be more advantageous if additional constraints are imposed by the amplifier dynamic range. However, such a scheme will result in a rate loss of $(N - m)/N$.

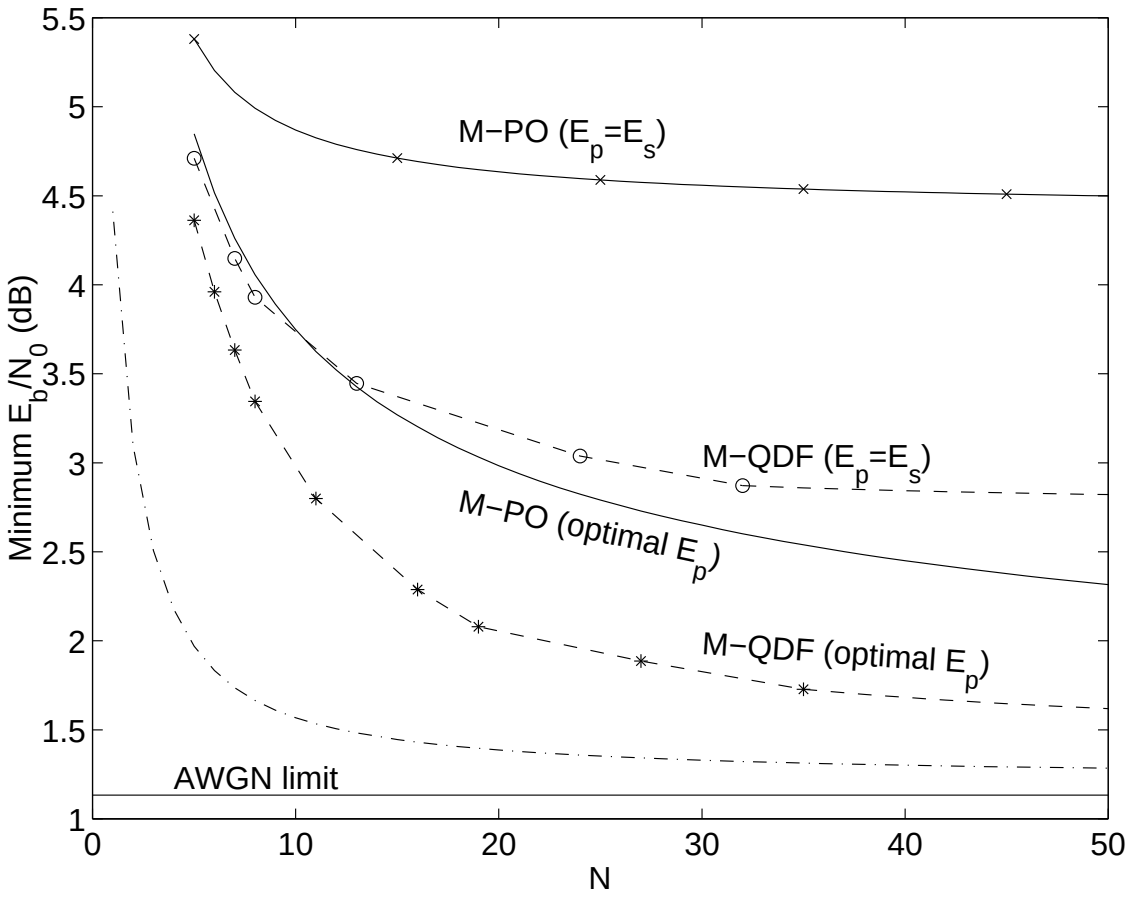


Figure 2.5: Density evolution results: Effect of pilot power optimization for different receivers.

(2.4c) falls between the curves of M-QDF and this combined lower bound (BPSK capacity curve for $N \leq 11$, and M-CDF curve for $N > 11$). The lowest curve on this figure is the capacity of the noncoherent channel without any modulation constraints taken from Chapter III. The difference between this curve and the BPSK capacity shows how much information bit SNR could be gained by utilizing more suitable modulation schemes (see Chapter III for a detailed discussion of capacity results). Density evolution results for irregular LDPC codes revealed the performance gain for irregular codes in the noncoherent channel (solid and dashed M-PO curves) is accurately predicted by the gain in the AWGN channel (solid and dashed

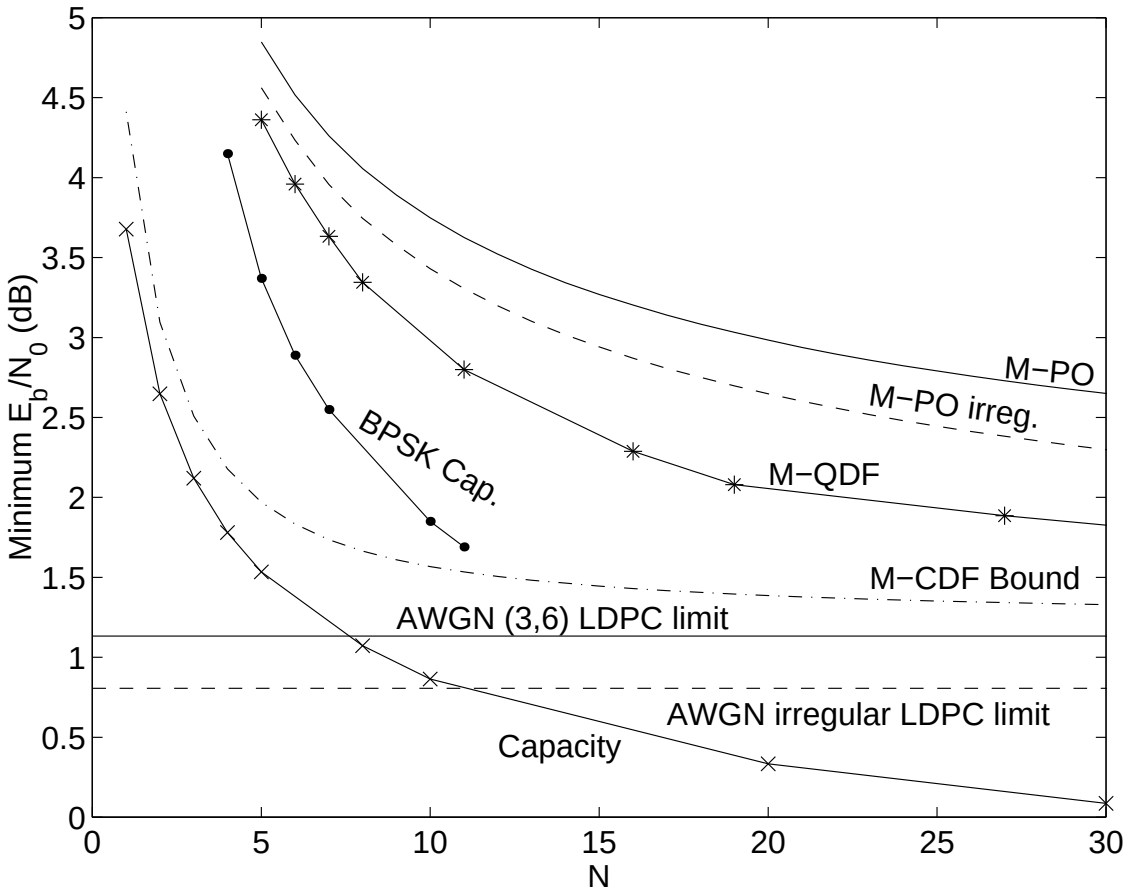


Figure 2.6: Density evolution results: Minimum required E_b/N_0 versus N curves for different receivers.

AWGN horizontal lines). No attempt was made to optimize the irregular codes for this channel; rather, the irregular code of [76] with maximum variable- and check-node degree of 4 and 6, respectively, was used. Finally, density evolution results, not shown here, confirmed that the performance of orthogonal modulation using symbol-by-symbol noncoherent detection is far worse than the performance of the proposed PSA schemes even when simple PO receivers are utilized (*e.g.*, the minimum required E_b/N_0 was found to be 7.41 dB).

Fig. 2.7 shows the optimal power allocation E_p/E_s required to achieve the minimum E_b/N_0 values, reported in Fig. 2.5, for the proposed receivers versus N. The

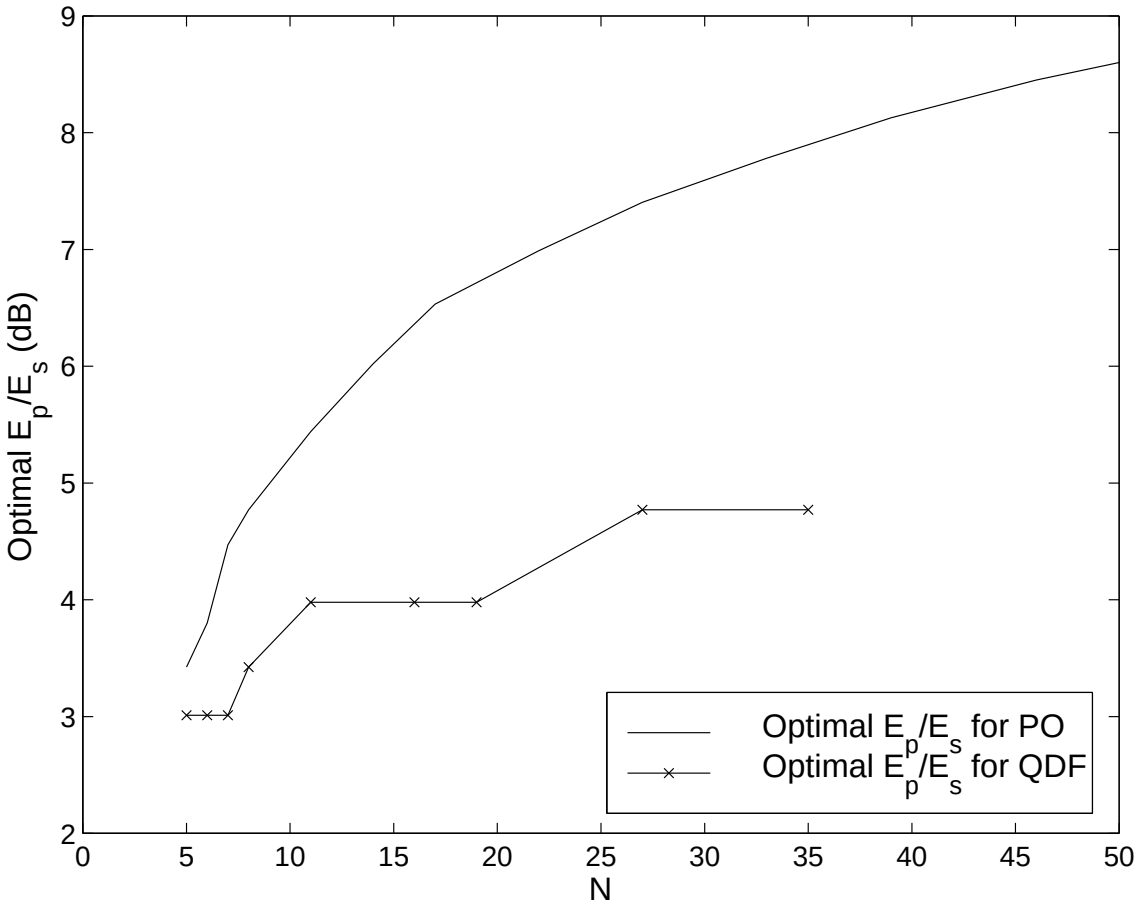


Figure 2.7: Density evolution results: Optimal power allocation E_p/E_s versus N for different receivers.

dependence of the optimal power allocation on N as well as on the particular algorithm used is illustrated. In particular, the increase of E_p/E_s with N is observed, as expected. Also, it could be seen from this figure that the better the estimation performed by the particular receiver, the less pilot power it requires for optimal operation. These facts should be taken into account when designing transmission systems; required pilot power is the property of not only the code but also the particular receiver utilized.

The figures shown previously are the results of density evolution which is an analytical tool and assumes infinite iterations and codelength. It is therefore important

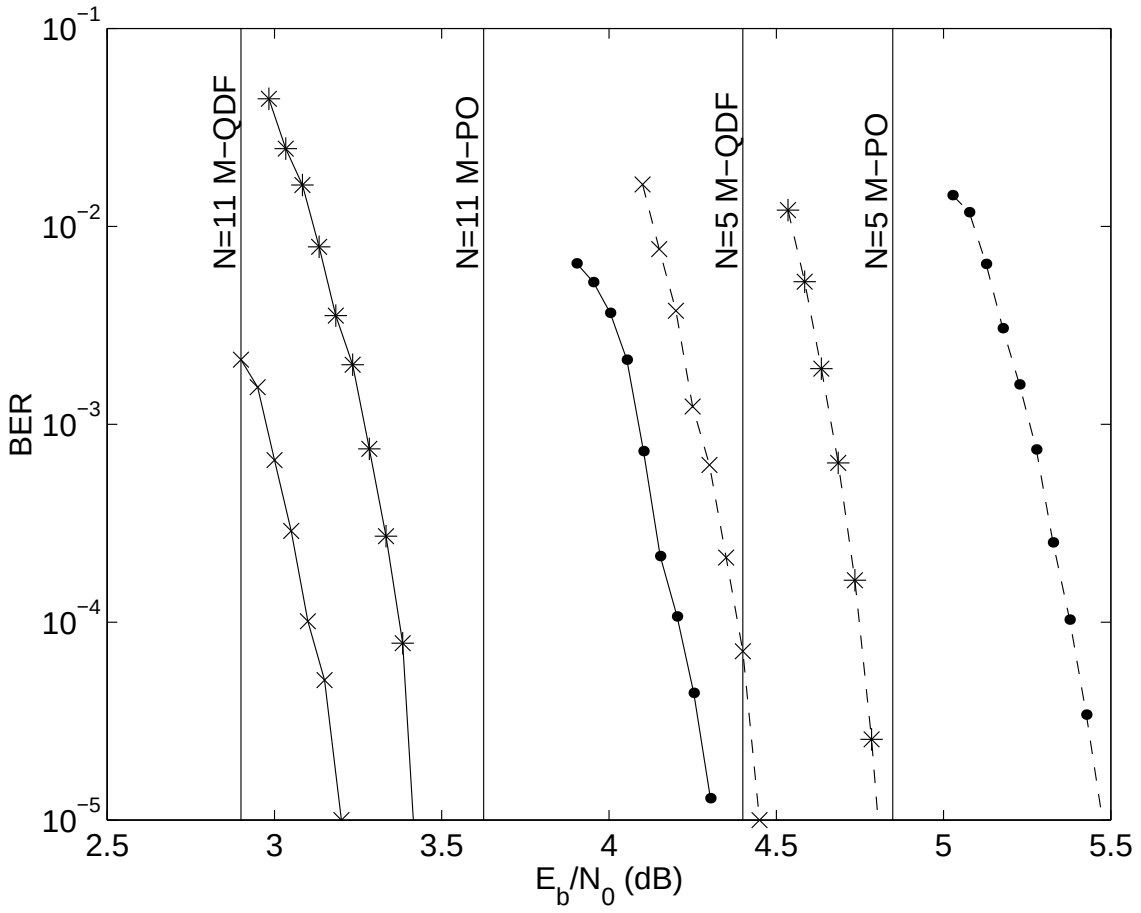


Figure 2.8: Simulation results for an $(n, k) = (4000, 2000)$ LDPC code. Curves show $N = 5$ (dashed) and $N = 11$ (solid) lines, with M-PO ($\bullet$), M-QDF ($*$) and QPH ($\times$) algorithms.

to quantify the effects of using small codes and finite iterations. Fig. 2.8 shows BER versus E_b/N_0 curves for M-PO and M-QDF receivers, obtained by simulating an $(n, k) = (4000, 2000)$ LDPC code. Vertical lines are the performance limits obtained by using density evolution. It is observed that the relative performance obtained using density evolution predicts the true performance of LDPC codes even for such small codes. Also plotted on this figure is the performance of another suboptimal receiver (denoted QPH) that works, by *quantizing* the *phase* variables θ, on the factor graph which explicitly models these phase variables (see Fig. 2.1(a)). Recalling

that $f_i(x,\theta) = \exp[\mathrm{Re}(\sqrt{E_s}xz_ie^{-j\theta})/\sigma^2]$, the message passed by this algorithm to the LDPC decoder can be written as

$$\mu_{fX_i} = \log \frac{L_i'(+1)}{L_i'(-1)} \tag{2.26}$$

$$L_i'(a) = \sum_{\theta\in\Theta}\Big[f_i(a,\theta)\exp(\mathrm{Re}(\sqrt{E_p}z_0e^{-j\theta})/\sigma^2)\times \times \prod_{\substack{j=1\\ j\neq i}}^{N-1}(f_j(+1,\theta)e^{\mu_{X_jf}} + f_j(-1,\theta))\Big], \tag{2.27}$$

where Θ denotes a finite set of quantized phase values. It was shown in [70, 36] that the performance of this suboptimal receiver is close to the optimal sum-product receiver even for a small number of quantization intervals (*e.g.*, 8 samples are adequate for almost identical performance in the case of BPSK). The pilot power allocation was optimized using simulations in this case, and the optimal E_p/E_s values were found to be 0.4 dB for $N = 5$ and 1.7 dB for $N = 11$. As expected, both values are below the optimal values predicted by density evolution for the M-QDF receiver, since the QPH receiver closely approximates the sum-product algorithm, and thus, utilizes the pilot energy more efficiently. Also observe that the performances of the M-QDF and QPH receivers are close (*i.e.*, within 0.5 dB), suggesting that the performance of the benchmark receiver is closer to the M-QDF curve than to the lower bound in Fig. 2.6.

To further justify the results of this chapter we applied these ideas to a more bandwidth efficient constellation than BPSK, namely quadrature PSK (QPSK), which results in twice the transmission rate. However, since density evolution of non-binary codes and the QPH receiver is complicated (if at all possible), the pilot power allocation was optimized using simulations. The use of binary modulation and a simple receiver for analysis/design and QPSK modulation for actual implementation of a

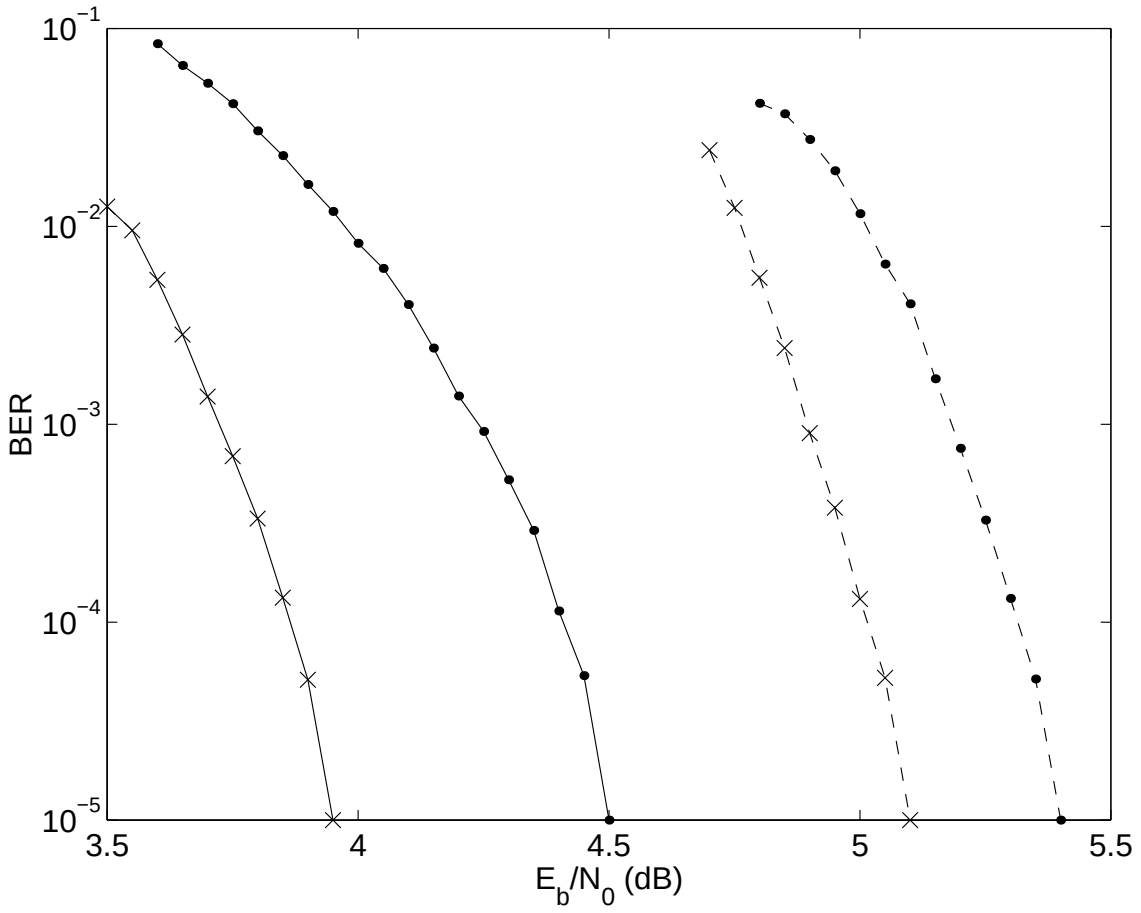


Figure 2.9: Simulation results for an $(n, k) = (4000, 2000)$ LDPC code with QPSK modulation. Curves show $N = 5$ (dashed) and $N = 11$ (solid) lines for the QPH algorithm with ($\times$) and without($\bullet$) pilot power optimization.

more complicated (*i.e.*, QPH) receiver is justified as can be seen in recent works [36]. The optimal E_p/E_s values were found to be 1.3 dB for $N = 5$ and 2.3 dB for $N = 11$, which is larger compared to the BPSK example, since phase estimation is more critical in QPSK. Simulation results are depicted in Fig. 2.9 (the same LDPC code as in Fig. 2.8 was used) with and without pilot power optimization. It can be observed that pilot power optimization still yields a gain of 0.6 dB (0.3 dB) at $N = 11$ ($N = 5$).

2.6 Conclusion

In this chapter, the design and analysis of PSA codes for the block-independent noncoherent AWGN channel was investigated. Several approximate receivers were proposed, which perform carrier-phase estimation either separately from decoding, or jointly as part of an iterative phase estimation/decoding process. The performance of these receivers was analyzed using density evolution. Based on these approximate receivers, a simple upper bound to the performance of any iterative joint phase estimation/decoding algorithm was derived. Utilizing density evolution as an analysis and optimization tool, the power allocation to the pilot symbol was quantified, and it was shown that a considerable performance gain can be obtained by designing codes with the optimal power allocation. Furthermore, this optimal allocation, depends highly on the channel coherence interval, and the particular algorithm used, and plays an increasingly critical role for fast channel dynamics. The development of these coding/decoding schemes is not specific to the particular regular LDPC codes used for demonstrating the concepts. For instance, irregular LDPC codes [76] can be used, providing a considerable additional performance gain.

The subject of the next two chapters is the design of more direct coding schemes that operate on the noncoherent channel when high dynamics are present. In this respect, the presented PSA codes and the corresponding receivers can serve as baseline – yet powerful – systems, against which other code designs can be compared.

CHAPTER III

CAPACITY OF THE NONCOHERENT CHANNEL

3.1 Introduction

In the previous chapter we analyzed a simple coding scheme for the block-independent noncoherent channel of (1.2), which used pilot symbols to aid phase estimation. It was shown that for high values of the channel coherence time N, such a system performs fairly well and permits the use of a vast amount of error correcting codes designed for the coherent AWGN channel. On the other hand, for faster phase dynamics, or equivalently, smaller values of N, such a segregated design is highly inefficient.

To combat this shortcoming we need to consider a more direct approach to coding by trying to utilize a close-to-optimal signaling scheme designed specifically for the noncoherent channel. Towards this goal we investigate the information capacity and the capacity achieving signaling scheme for this channel. The next chapter uses the results found herein to design specific coding schemes that come close to the theoretical limit and yet have practical complexity.

The capacity of the block-independent noncoherent channel has been investigated in the literature. In the early work of [12], the capacity of a frequency shift keying

(FSK) system was investigated numerically. In [68] the capacity of this channel was analyzed for M-ary phase shift keying (M-PSK) transmission. In particular, it was shown that the capacity-achieving inputs are independent and identically distributed M-PSK symbols. A partial characterization of the capacity-achieving distribution for non-constraint inputs was done in [18]. It was shown that the capacity-achieving input signal consists of N complex variables whose phases are independent identically distributed (i.i.d.) random variables, uniformly distributed over $[-\pi, \pi)$, and also independent of the amplitudes. Very recently [38, 39], the capacity-achieving distribution was found to have a discrete amplitude nature with infinite number of mass points for the special case of a memoryless phase process, *i.e.*, for $N = 1$. Finally, upper and lower bounds on the capacity for the case of $N = 1$ were derived in [44]. To date, a complete characterization of the capacity-achieving input distribution is not available for arbitrary phase dynamics, *i.e.*, for arbitrary N.

In Section 3.2, four facts about the structure of the capacity achieving input distribution for this channel are proved:

- The maximizing input density has circular symmetry, that is, all directions in the complex N-dimensional space should be used equally probable.
- There exists a unique amplitude distribution that maximizes the mutual information.
- The maximizing amplitude density is *discrete*.
- The maximizing amplitude density has infinite number of mass points.
- The maximizing density always has a mass point at zero.

We note that in a recent independent work [39], the validity of the third and

fourth facts mentioned above was established for a special case of a memoryless (*i.e.*, $N = 1$) noncoherent channel.

Based on the above characterization of the maximizing input distribution, asymptotic expressions are derived in Section 3.3 that relate the mass-point probabilities with the mass-point locations. These expressions suggest a double exponential decrease of the probabilities with respect to the mass-point locations. Motivated by these expressions, several numerical optimization results on capacity are reported in Section 3.4. In particular, the gain from utilizing more than one mass points in the amplitude density of the transmitted vector is investigated. These results indicate that for the range of signal-to-noise ratios (SNRs) and code rates considered herein, the amplitude density of the capacity-achieving input is well approximated by a two-point discrete density, with one point at zero. Furthermore, the capacity loss incurred by the use of practical modulation schemes is quantified.

3.2 Capacity Characterization

Recall the input/output relationship for the considered channel

$$\mathbf{y}_k = \mathbf{x}_k e^{j\theta_k} + \mathbf{n}_k, \tag{3.1}$$

where $\mathbf{x}_k$, $\mathbf{y}_k$ and $\mathbf{n}_k$ are complex sequences of length N, denoting the k-th block of the transmitted sequence, the observed sequence, and the noise sequence of independent identically distributed (i.i.d.), zero-mean, circular, complex Gaussian random variables with variance $\sigma^2 = N_0/2$ per real dimension, respectively. The unknown phase rotation in each block is represented by the variables θ_k, which are modeled as i.i.d. random variables uniformly distributed in $[-\pi, \pi)$. If blocks of length N are treated as symbols, this channel is memoryless, time invariant and is completely

specified by its transition probability

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^N} e^{-\frac{\|\mathbf{x}\|^2+\|\mathbf{y}\|^2}{2\sigma^2}} I_0\Big(\frac{|\mathbf{x}^H\mathbf{y}|}{\sigma^2}\Big). \tag{3.2}$$

Here and in the following, $I_n(\cdot)$ denotes the nth-order modified Bessel function of the first kind, and $\mathbf{x}^H$ and $\|\mathbf{x}\|$ denote the complex conjugate transpose, and the norm of $\mathbf{x}$, respectively. It is noted that the presence of the Bessel function in (3.2) significantly complicates analysis for this channel (as will be evident in the next sections) and is one of the reasons why this channel has not been as well studied as the corresponding block-independent fading channel.

We are interested in characterizing the capacity achieving distribution of $\mathbf{x}$ under an average power constraint $E(\|\mathbf{x}\|^2) \leq \mathsf{P}$. This problem can be precisely formulated as

$$\mathcal{C} = \sup_{\substack{p(\mathbf{x}) \\ E(\|\mathbf{x}\|^2)\leq \mathsf{P}}} I(p) \quad \text{with } I(p) \triangleq I(\mathbf{x};\mathbf{y}), \tag{3.3}$$

where $I(\mathbf{x};\mathbf{y})$ denotes the mutual information between $\mathbf{x}$ and $\mathbf{y}$

$$I(\mathbf{x};\mathbf{y}) = \iint\limits_{\mathbb{C}^{2N}} p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \log \frac{p(\mathbf{y}|\mathbf{x})}{\int p(\mathbf{y}|\mathbf{x}')p(\mathbf{x}')d\mathbf{x}'} d\mathbf{x}d\mathbf{y}. \tag{3.4}$$

A note regarding the convention used herein is in order. Although density functions instead of distribution functions are used throughout this chapter, this is only a notational convention. All derivations and results can be restated in a more rigorous manner by substituting densities with distributions and interpreting the integrals in the Lebesgue-Stieljes sense. Another approach is followed in proving certain facts in Section B.3 of Appendix B, where the density functions are interpreted as linear functionals on a space of bounded continuous functions. For smoothness of presentation we will refrain from introducing such complications early in the text and do so as they become necessary. Likewise, most of the details of the proofs are collected in Appendix B at the end of the thesis.

Generally, this problem involves an optimization over the densities of N dimensional random complex vectors, since there is no smaller-dimensional sufficient statistic for the output of the channel. However, the following lemma shows that we can restrict ourselves to a much smaller space of densities, namely the space of *circularly symmetric* densities. An N-dimensional complex random vector is called circularly symmetric if its density function depends only on the magnitude of the vector, *i.e.*,

$$p_{\mathbf{x}}(\mathbf{x}) = f(\|\mathbf{x}\|) \qquad \forall \mathbf{x} \in \mathbb{C}^N, \tag{3.5}$$

for some nonnegative function $f : [0, \infty) \to [0, \infty)$. Simply put, circularly symmetric random vector is one which uses all directions in the N-dimensional complex space equally likely, and is completely specified by the density of its amplitude. Therefore, the space of such functions can be parameterized in terms of one real function (that is, the density of $\|\mathbf{x}\|$). If $\mathfrak{S}$ denotes the space of circularly symmetric densities then we have

Lemma 3.1. *The optimization problem of* (3.3) *simplifies to*

$$\mathcal{C} = \sup_{\substack{p(\mathbf{x}) \\ E(\|\mathbf{x}\|^2) \le \mathsf{P}}} I(p) = \sup_{\substack{p(\mathbf{x}) \in \mathfrak{S} \\ E(\|\mathbf{x}\|^2) \le \mathsf{P}}} I(p). \tag{3.6}$$

Proof. We prove this lemma by showing that for any given density $p_0(\mathbf{x})$ there exists a density $p(\mathbf{x}) \in \mathfrak{S}$ that achieves mutual information not less than the one achieved by original density, that is $I(p) \ge I(p_0)$. This fact is a special case of a more general result obtained in [53] for a multiple-input/multiple-output (MIMO) communication system over a complex Gaussian block-independent fading channel. Although the channels under consideration are different, the only requirement for the proof is that the channel transition probability satisfies $p(\mathbf{y}|\mathbf{x}) = p(U\mathbf{y}|U\mathbf{x})$ for any deterministic unitary matrix U, which is true in this case as can be seen from (3.2). The proof

proceeds by first showing that $I(p(\mathbf{x})) = I(p(U\mathbf{x}))$, for any unitary U, and then forming a new random variable $\mathbf{x}' = \Phi\mathbf{x}$, where Φ is an *isotropically distributed* random matrix independent of $\mathbf{x}$, *i.e.*, it is an $N \times N$ random unitary matrix whose distribution is unchanged when it is multiplied by a deterministic unitary matrix (see [53] for a more detailed discussion on the isotropically distributed random matrices and vectors). The desired result follows by observing that $\|\mathbf{x}'\| = \|\mathbf{x}\|$ and using the fact that for a fixed transition probability, the mutual information is a convex cap functional of the input density. ■

The lemma indicates that all directions in the N-dimensional complex space should be used equally likely. In other words, the maximizing input vector — conditioned on its amplitude — is uniformly distributed on a sphere in the N-dimensional complex space. More precisely, the maximizing input vector is of the form $\mathbf{x}' = r\mathbf{h}$, where r is a real, nonnegative random variable, and $\mathbf{h}$ is an isotropically distributed random *vector*[1], independent of r, with density

$$p(\mathbf{h}) = \frac{(N-1)!}{\pi^N}\delta(\|\mathbf{h}\|^2 - 1). \tag{3.7}$$

The independence of r and $\mathbf{h}$ follows because $\mathbf{x}' = \Phi\mathbf{t}\|\mathbf{x}\|$, where $\mathbf{t} = \mathbf{x}/\|\mathbf{x}\|$. Since Φ is isotropically distributed, $\mathbf{h} = \Phi\mathbf{t}$ conditioned on $\mathbf{t}$ and r has the density given by (3.7), and hence is statistically independent of $\mathbf{t}$ and $r = \|\mathbf{x}'\|$.

Furthermore, this lemma implies the conclusion, previously reached in [18], that the phases of the optimally distributed input vector $\mathbf{x}$ are i.i.d. uniformly distributed and independent of their amplitude, namely if $\mathbf{x} = [r_1 e^{j\theta_1}, \ldots r_N e^{j\theta_N}]^T$ then

$$p(r_1, \ldots r_N, \theta_1 \ldots \theta_N) = p(r_1, \ldots, r_N)\left(\frac{1}{2\pi}\right)^N. \tag{3.8}$$

---

[1]If $\mathbf{z}$ is a vector of N i.i.d. zero-mean complex Gaussian random variables, then $\mathbf{h} = \mathbf{z}/\|\mathbf{z}\|$.

The main advantage of this lemma is that it effectively reduces the dimensionality of the optimization problem; instead of searching over densities of N complex variables, we only need to search over densities of one real variable. Using this fact, the mutual information in (3.3) that corresponds to a circularly symmetric input with normalized amplitude $a = \|\mathbf{x}\|/\sigma$ with density $p_a(a)$, can be written as (see Section B.1 of Appendix B for derivation details)

$$I(p_a) = \int_0^\infty F_{p_a}(a) p_a(a) da. \tag{3.9}$$

where

$$F_{p_a}(a) = \int_0^\infty p_1(r|a) \log \frac{p_1(r|a)}{r} dr - \int_0^\infty p_N(r|a) \log \frac{p_N(r) c_{N-1}}{r^{2N-1}} dr \tag{3.10}$$

In the above equation, $c_n = (2e)^n n!$, and the density of the normalized output variable $\|\mathbf{y}\|/\sigma$ is $p_N(r) = \int p_N(r|a) p_a(a) da$, with conditional density given as

$$p_n(r|a) = e^{-(r^2+a^2)/2} r \left(\frac{r}{a}\right)^{n-1} I_{n-1}(ar) \tag{3.11}$$

With this notation, the optimization problem in (3.3) becomes

$$\mathcal{C} = \sup_{\substack{p_a(a) \\ E(a^2) \le \gamma}} \int_0^\infty F_{p_a}(a) p_a(a) da, \tag{3.12}$$

where $\gamma = \mathsf{P}/\sigma^2$.

The necessary and sufficient conditions (referred to as "Kuhn-Tucker condition") for an input amplitude density to be the maximizing one are derived in Section B.2 of Appendix B as

$$F_{p_a}(a) \le \lambda + \mu a^2, \tag{3.13}$$

for some positive λ and μ, with equality if a is a mass point, that is $\int_{a-\delta}^{a+\delta} p_a(a') da' > 0$ for any $\delta > 0$. The set of all such points is called the support set. Observe that parameters λ and μ are related to the capacity as $\mathcal{C} = \lambda + \mu\gamma$.

Furthermore, in Section B.3 of Appendix B it is shown that the supremum in (3.12) is *achieved* by a *unique* density $p_a(a)$. This, in turn, implies that there exists exactly one capacity-achieving circularly symmetric input density for the input vector $\mathbf{x}$. This is true since, as mentioned earlier, a circularly symmetric input density is precisely identified by the corresponding amplitude density. However, it does *not* imply the uniqueness of the capacity-achieving input density, as there can be capacity achieving input densities that are not circularly symmetric. On the contrary, there indeed exist several such input densities. For instance, restricting any one symbol in the N dimensional input complex vector $\mathbf{x}$ to be real, does not reduce the achievable capacity, and yet such an input density is not circularly symmetric. It is, however, noted that all the capacity achieving input densities have exactly the same amplitude variation, as it easily follows from the application of the methods used in the proof of Lemma 3.1.

The next two theorems prove that the optimizing density of a is *discrete* with *infinite* number of mass points, but only finitely many in any bounded interval.

Theorem 3.2. *The capacity achieving density of $a = \|\mathbf{x}\|/\sigma$ is discrete, that is a can have at most finite number of mass points in any bounded interval.*

Proof. The proof is by contradiction. First observe that by using the fact

$$\int r^2 p_N(r|a)dr = 2N + a^2, \tag{3.14}$$

we can rewrite the Kuhn-Tucker condition in the more suitable form as

$$\int_0^\infty e^{-r^2/2} r \left(\frac{r}{a}\right)^{N-1} I_{N-1}(ar) \log \phi(r) dr = \int_0^\infty e^{-r^2/2} r I_0(ar) \log I_0(ar) dr \tag{3.15}$$

for all a such that $p(a) > 0$, where

$$\phi(r) = \frac{p_N(r) c_{N-1}}{r^{2N-1}} e^{\lambda+1+(\mu+1)(r^2-2N)} \tag{3.16}$$

Both sides of (3.15) can be analytically extended to the entire finite complex plane as

$$\Psi_L(z) = \int_0^\infty e^{-r^2/2} r \left(\frac{r}{z}\right)^{N-1} I_{N-1}(zr) \log \phi(r) dr \tag{3.17a}$$

$$\Psi_R(z) = \int_0^\infty e^{-r^2/2} r I_0(zr) \log I_0(zr) dr \tag{3.17b}$$

where for $\Psi_R(z)$ the integral should be understood in the Cauchy sense (see below for details).

Assume, on the contrary, that the statement of the theorem is not true, *i.e.*, in some (closed) finite interval there are infinitely many mass points. This in turn implies that the support set has an accumulation point. Equation (3.15) and the above assumption imply that the two analytic functions in (3.17) coincide on a set which has an accumulation point inside the region of analyticity. Therefore, it follows from the identity theorem [45, Sect. I-20] that these functions are identical in the entire region of analyticity, that is in the entire finite complex plane. We will arrive at a contradiction by showing that this statement is violated on the imaginary axis.

On the imaginary axis $z = jb$ we have

$$\Psi_L(jb) = \int_0^\infty e^{-r^2/2} r \left(\frac{r}{b}\right)^{N-1} J_{N-1}(br) \log \phi(r) dr \tag{3.18}$$

where $J_n(\cdot)$ is the n-th order Bessel function of the first kind. The expression in (3.17b) can be used directly (the integral understood in the Riemann sense) for any complex z with positive real part. However, the extension of this function to the imaginary axis and beyond is not straightforward. This is due to the presence of the $\log I_0(zr)$ component, which is pathological since $I_0(\cdot)$ has zeros on the imaginary axis. This obstacle is overcome by finding an alternative representation of $\Psi_R(z)$ that is valid in a region containing the imaginary axis. One such representation valid in

a region defined by $\arg(z) \in (\pi/4, 3\pi/4)$ is presented in Section B.4 of Appendix B, which for the case of $z = jb$ yields

$$\Psi_R(jb) = \int_0^\infty e^{-r^2/(2b^2)} \frac{r}{b^2} J_0(r) \log |J_0(r)| dr + j\pi \int_0^\infty e^{-r^2/(2b^2)} \frac{r}{b^2} J_0(r)\xi(r) dr. \quad (3.19)$$

where $\xi(r) = 0$ if $r \in (0, \alpha_1)$ and $\xi(r) = k$ if $r \in (\alpha_k, \alpha_{k+1})$, with $\alpha_k, k \geq 1$ being the positive zeros of $J_0(r)$ in increasing order. We will now show that the second integral in the above expression (*i.e.*, the imaginary part of $\Psi_R(jb)$) is not identically zero, thus arriving at a contradiction since $\Psi_L(jb)$ in (3.18) is real for all values of b. We accomplish this by bounding the imaginary part away from zero as (see Section B.5 of Appendix B for the details)

$$\frac{|\mathrm{Im}(\Psi_R(jb))|}{\pi} \geq (|J_0(\beta)| - \frac{b^2}{\beta}) e^{-\beta^2/(2b^2)} - (1 + \frac{2b^2}{\pi\alpha_2}) e^{-\alpha_2^2/(2b^2)} \quad (3.20)$$

where β is *the* positive zero of $J_1(r)$ less than α_2. This expression is positive for small enough values of b (*e.g.*, at $b = 1$ it is 2.9×10^{-5}).

Hence, our original assumptions are not valid which means that the capacity achieving density is indeed discrete. ■

At this point it is worth noting that the general methodology followed in proving the above theorem parallels the approach of [82], [25], and [38] for the amplitude-constrained AWGN channel, the fast (*i.e.*, $N = 1$) Rayleigh fading channel, and the fast (*i.e.*, $N = 1$) noncoherent AWGN channel, respectively. The technical difficulties in the channel under consideration originate from the fact that (i) we are considering the general case of $N \geq 1$ and (ii) the presence of Bessel functions in (3.15) (instead of the exponential functions that appear in [25]) prevents the use of the Laplace transform for solving explicitly (3.15) for $\phi(r)$.

As a side note, it is observed that the expression in (3.18) can be recognized as the Hankel transform of index $N - 1$ of the function $e^{-r^2/2} r^{N-1} \log \phi(r)$ divided by

b^{N-1}, and hence, one can invert[2] the resulting expression to get

$$\log \phi(r) = \frac{e^{r^2/2}}{r^{N-1}} \mathcal{H}_{N-1}^{-1}(b^{N-1}\Psi_R(jb)) = \frac{e^{r^2/2}}{r^{N-1}} \int_0^\infty \Psi_R(jb) b^{N-1} b J_{N-1}(br) db, \quad (3.21)$$

where $\mathcal{H}_{N-1}^{-1}$ denotes the inverse Hankel transform of index $N-1$. This expression would be useful for obtaining an explicit solution if $\Psi_R(jb)$ was a real function. However, as it is shown in the above proof, $\Psi_R(jb)$ has an imaginary component, and thus no positive solution for $\phi(r)$ exists.

The theorem implies that the only way to have infinitely many mass points is if they form a sequence extending to infinity in which case the support set has an accumulation point at infinity. As it turns out, unlike the case of Rayleigh fading [25] or amplitude-constrained [82] channels, the maximizing density for the noncoherent AWGN channel does have infinitely many mass points, as the following theorem shows. An intuitive explanation of this result is that the noncoherent AWGN channel does not suffer from the detrimental effects of a random amplitude, and thus, the risk involved in receiving a very small signal even though a very large signal is sent, is smaller compared to the case of Rayleigh channels.

Theorem 3.3. *The capacity achieving density of $a = \|\mathbf{x}\|/\sigma$ has an infinite number of mass points.*

Before proving this theorem we need the following lemma, the proof of which is in Section B.6 of Appendix B.

Lemma 3.4. *The Lagrange multiplier μ in* (3.15) *is less than 1/2.*

We are now ready to prove the theorem.

---

[2]Assuming $b^{N-1}\Psi_R(jb)$ satisfies some sufficiency conditions (*e.g.*, see [19, p. 71]).

Proof of Theorem 3.3. The proof is by contradiction. First, let a_k and p_k be the positions and probabilities of the mass points and rewrite the Kuhn-Tucker Condition (3.13) as

$$\lambda + \mu a^2 - F_{p_a}(a) \geq 0 \tag{3.22}$$

with equality at mass points $a = a_k$ for $k \geq 1$. Let

$$S_n(x) = c_n \frac{I_n(x)}{x^n}$$

and observe that $S_n(x)$ is an increasing function with $S_n(0) = e^n$ for any integer $n \geq 0$.

Assume on the contrary that the capacity achieving density has finitely many mass points. Then there exists M such that $a_M > a_k$ for any $k \neq M$. Therefore using (3.11) we have

$$\begin{aligned} \frac{c_{N-1} p_N(r)}{r^{2N-1}} = \sum_k p_k \frac{c_{N-1} p_N(r|a_k)}{r^{2N-1}} &= \sum_k p_k e^{-r^2/2} e^{-a_k^2/2} S_{N-1}(a_k r) \\ &\leq e^{-r^2/2} S_{N-1}(a_M r) \end{aligned} \tag{3.23}$$

Combining this bound with the fact that $S_n(x) \leq \exp(x+n)$, we can estimate the second term in (3.10) as

$$\begin{aligned} \int_0^\infty p_N(r|a) \log \frac{c_{N-1} p_N(r)}{r^{2N-1}} dr &\leq \int_0^\infty p_N(r|a) \log(e^{-r^2/2} S_{N-1}(a_M r)) dr \\ &= -\frac{a^2}{2} - N + \int_0^\infty p_N(r|a) \log S_{N-1}(a_M r) = -\frac{a^2}{2} + O(a) \end{aligned} \tag{3.24}$$

where the $O(a)$ term applies for a going to infinity. Similarly the first term in (3.10) is also of order[3] at most $O(a)$. Finally, combining these evaluations with (3.22) we get that

$$\lambda + \mu a^2 - [\frac{1}{2} a^2 + O(a)] \geq \lambda + \mu a^2 - F_{p_a}(a) \geq 0 \tag{3.25}$$

---

[3] In fact, it behaves as $O(\log a)$.

which implies that $\mu \geq 1/2$. This is a contradiction to the result of Lemma 3.4; therefore, the original assumption is incorrect, and thus there are infinitely many mass points which form a sequence extending to infinity. ■

The results so far show that the input vectors $\mathbf{x}$ that achieve capacity are isotropically distributed on infinitely many concentric spheres in the N-dimensional complex space. The next theorem takes the characterization of the maximizing density one step further and proves that there always exist a mass point at the origin. Observe that for the case of $N = 1$ this result is quite intuitive and straightforward to show: if there is no mass point at zero, then reducing the smallest amplitude to zero increases the "mass-point separation" at no extra cost (to say nothing of the reduced average power, which can be used to increase the amplitude of the other mass points). The proof in that case is easily carried out by evaluating the derivative of the mutual information with respect to the smallest amplitude and observing that the latter is negative as long as the smallest amplitude is nonzero (this is essentially the technique used in [25]). In the case of general N, there is some information stored in the phase (direction vector), and hence reducing a mass point to the origin has two contradicting effects: (i) the separation between spheres increases (for the same power) thus increasing the mutual information, and (ii) the information transmitted through the phase of the smallest sphere is lost which reduces the mutual information. In general, reducing the smallest mass point to zero does *not* necessarily increase the mutual information. The next theorem, follows a different approach to prove the existence of a mass point at zero.

Theorem 3.5. *The optimizing probability mass function of $a = \|\mathbf{x}\|/\sigma$ has a mass point at the origin for all values of SNR and channel coherence time N.*

We note that the proof of this theorem was by far the most complicated compared to the rest of the results in this thesis. As the details are lengthy and tedious, we give the general idea of the proof here, with the complete proof presented in Section B.7 of Appendix B.

Sketch of the Proof. We are actually proving a somewhat stronger fact from which this theorem follows as a consequence. In particular, we are proving that for any given input density that does not have a mass point at zero, we can construct a corresponding input density with a mass point at zero that results in higher mutual information. This way we are establishing the "introduction of a mass point at zero" as a simple and useful tool for modulation design for this channel.

Let $\mathbf{a} = \{0, a_1, a_2 \ldots\}$ and $\mathbf{p} = \{p_0, p_1, p_2 \ldots\}$ be the locations and probabilities of the mass points, respectively, so that the pair $(\mathbf{a}, \mathbf{p})$ denotes a valid input amplitude density. For notational simplicity also assume that the sequence $\mathbf{a}$ is (strictly) increasing and let $I(\mathbf{a}, \mathbf{p})$ be the mutual information corresponding to this input density.

Assume an input density with no mass point at the origin, *i.e.*, $p_0 = 0$ and consider the input density pair

$$\mathbf{p}^*(x) = \{x, p_1 - x, p_2, p_3 \ldots\} \tag{3.26}$$

$$\mathbf{a}^*(x) = \{0, a_1\sqrt{\frac{p_1}{p_1 - x}}, a_2, a_3, \ldots\}, \tag{3.27}$$

which corresponds to introducing a mass point at zero of probability x at the expense of the minimum amplitude mass point, and accordingly increasing the latter to keep the average power the same. Observe that the case of $x = 0$ corresponds to the original input density pair. We show that

$$\frac{d}{dx} I(\mathbf{a}^*(x), \mathbf{p}^*(x))\Big|_{x=0} > 0, \tag{3.28}$$

which implies that there is a $x_0 > 0$ so that $I(\mathbf{a}^*(x_0), \mathbf{p}^*(x_0)) > I(\mathbf{a}, \mathbf{p})$, that is, the original provided input density can not be the optimal one. Since the initial density was chosen arbitrarily (with no mass point at zero), this result implies that the optimal density must have a mass point at the origin, proving our claim.

Observe that

$$\frac{d}{dx} I(\mathbf{a}^*(x), \mathbf{p}^*(x))\Big|_{x=0} = \frac{\partial I}{\partial p_0} - \frac{\partial I}{\partial p_1} + \frac{1}{2}\frac{a_1}{p_1}\frac{\partial I}{\partial a_1}, \tag{3.29}$$

where the partial derivative with respect to a_k is the left-hand side of (B.28) in Section B.6 of Appendix B multiplied by $p_k a_k$, and the partial derivative with respect to p_k is given by $\frac{\partial I}{\partial p_k} = F_{p_a}(a_k) - 1$. Therefore, the derivative in (3.28) after some manipulation becomes

$$\begin{aligned} \frac{a_1^2}{2} + \int_0^\infty \left(\frac{a_1^2}{2} p_2(r|a_1) - (1 + \frac{a_1^2}{2}) p_1(r|a_1) \right) \log I_0(a_1 r) dr \\ - \int_0^\infty \left(p_N(r|0) + \frac{a_1^2}{2} p_{N+1}(r|a_1) - (1 + \frac{a_1^2}{2}) p_N(r|a_1) \right) \log \frac{p_N(r)}{r^{2N-1}} dr. \end{aligned} \tag{3.30}$$

The sum of the first two terms in the above expression is always positive, while the last term can be positive or negative. The proof is established by showing that the sum of the first two terms is greater than the last term. ■

3.3 Asymptotic Results

Aided by the results of the previous section on the properties of the capacity-achieving input distribution, one can proceed with a numerical evaluation of the capacity. This task involves numerical optimization over the positions and the probabilities of the mass points of the amplitude density. Before we do that, we present an asymptotic result on the relationship between the positions of the mass points and their corresponding probabilities. The value of such an expression is twofold: (i)

it can potentially aid the aforementioned numerical optimization by utilizing an additional constraint, thus limiting the search over a smaller set of unknowns, and (ii) it provides additional insight on the maximizing input distribution in some limiting cases, as will be discussed in the following.

Let a_k and p_k be the locations and probabilities of the mass points, respectively, with $a_0 = 0$. From the Kuhn-Tucker condition we know that for all k

$$F_{p_a}(a_k) = \lambda + \mu a_k^2. \tag{3.31}$$

Using the bound $p_N(r) = \sum_i p_i p_N(r|a_i) \geq p_k p_N(r|a_k)$ we get that

$$\lambda + \mu a_k^2 = F_{p_a}(a_k) \leq -\log p_k + g_N(a_k), \tag{3.32}$$

with

$$g_N(a_k) = \int_0^\infty p_1(r|a_k) \log \frac{p_1(r|a_k)}{r} dr - \int_0^\infty p_N(r|a_k) \log \frac{p_N(r|a_k) c_{N-1}}{r^{2N-1}} dr \tag{3.33}$$

which implies the bound

$$p_k \leq \exp(g_N(a_k) - \lambda - \mu a_k^2). \tag{3.34}$$

Although one can numerically evaluate the function $g_N(\cdot)$, an asymptotic expression, derived from well-known asymptotic expressions for the involved Bessel functions, is shown below

$$p_k \simeq p_0 e^{-\mu a_k^2} \frac{(a_k^2 + 2N - 1)^{N-1}}{c_{N-1}}. \tag{3.35}$$

The right hand side of (3.34) and the approximation in (3.35) are plotted in Fig. 3.1 for a special case of $N = 7$, $\mu = 0.2$ and $p_0 = 0.25$. It is evident from this figure that the asymptotic approximation of (3.35) is indeed accurate for high values of a.

Several observations regarding these expressions are in place here.

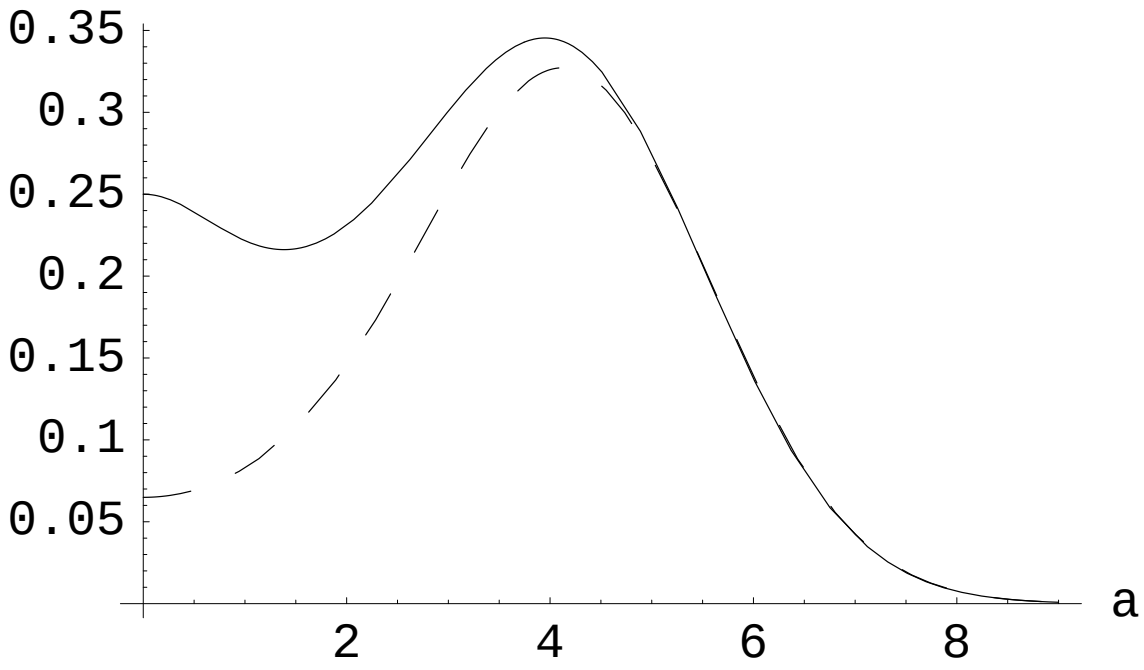


Figure 3.1: Bound on the mass point probabilities versus mass point locations for $N = 7$, $\mu = 0.2$ and $p_0 = 0.25$. The dashed curve is the asymptotic approximation of (3.35).

- It can be shown that the smaller the Lagrange multiplier μ (*i.e.*, the higher the SNR), the more accurate the expression in (3.35) is.

- For large values of the amplitude a_k, the corresponding probabilities decrease as $\exp(-\mu a_k^2)$. The square exponential decrease of the mass point probabilities is the main reason why there is only marginal error when high amplitude terms are truncated (*e.g.*, when numerical evaluation is performed).

- This approximate expression hints at densities that can be used to derive lower bounds on the capacity. For instance, a lower bound was derived in [38] (for the case of $N = 1$) using a half-Gaussian continuous amplitude density. This choice is justified as a good lower bound from the asymptotic result presented above. This asymptotic result is also consistent with the results of [44], where

it was shown that the half-Gaussian density achieves capacity at high SNR values.

- By defining a normalized amplitude $\beta = a/\sqrt{2N}$, one can observe a sphere hardening effect for large values of N (details provided below); this is consistent with the fact that for large N, the block-independent noncoherent channel resembles a coherent AWGN channel.

Sphere Hardening Effect: Now we show that the optimum input density exhibits the sphere hardening effect, that is with increasing N, the amplitude variation collapses to a single point. For that, fix μ, and assuming $N > 1$ rewrite (3.35) as

$$p(a) \simeq Ke^{-\mu a^2}(a^2 + 2N - 1)^{N-1}, \tag{3.36}$$

where K is some constant chosen to make the maximum of the right hand side of (3.36) equal to 1, that is

$$K = \left(\frac{\mu}{N-1}e^{1-2\mu}\right)^{N-1} e^{-\mu}. \tag{3.37}$$

Using this choice[4] of K and substituting $a = \beta\sqrt{2N}$, the right hand side of (3.36) becomes

$$f_N(\beta) = \left[e^{1-X}(X + \frac{X-\mu}{N-1})\right]^{N-1} e^{\mu - X}, \tag{3.38}$$

where $X = 2\mu(\beta^2 + 1)$. Taking advantage of the following well-known limit

$$\lim_{n\to\infty}\left(x + \frac{t}{n}\right)^n = \begin{cases} e^t & x = 1 \\ 0 & 0 \le x < 1 \\ \infty & x > 1 \end{cases} \tag{3.39}$$

[4]Because we are interested in the relative behaviour of the different amplitudes, a, the choice of K does not affect the sphere hardening effect.

and noting that $Xe^{1-X} \leq 1$ with equality if and only if $X = 1$, we obtain the desired result

$$\lim_{N\to\infty} f_N(\beta) = \begin{cases} 1, & X = 1 \\ 0, & \text{otherwise,} \end{cases} \tag{3.40}$$

that is, the limit of $f_N(\beta)$ is nonzero if and only if $\beta = \sqrt{\frac{1}{2\mu} - 1}$, which establishes the desired result. □

In view of the above discussion, discrete optimization methods can be utilized to evaluate (or approximate) the maximizing input distribution. We have seen little performance loss (for the range of values of SNR and N that we are interested in) when the number of mass points was restricted to 2, one at zero and the other at a non-zero location. Therefore, in the following by "capacity" of this channel we refer to this two-mass-point capacity. It should be noted that, although we have proven the existence of a mass point at zero, this is not necessarily true for the case of a *fixed* number of mass points. However, for the range of values of interest, the two-point optimizing density was found to have a mass point at zero.

3.4 Numerical Results

Throughout this section we quantify two important measures of performance. The first one, referred to as "discretization loss", is the result of using modulation schemes that approximate the circularly symmetric signaling suggested by Lemma 3.1. In other words, this loss is due to the fact that not all but only a finite number of points on the surface of a sphere are used in a specific modulation scheme. The second performance measure, referred to as "shaping gain[5]", is the gain

[5]The term "shaping" is used in general to describe the process of adapting the transmitted signal to match the optimal input distribution. This process has been studied extensively for the AWGN

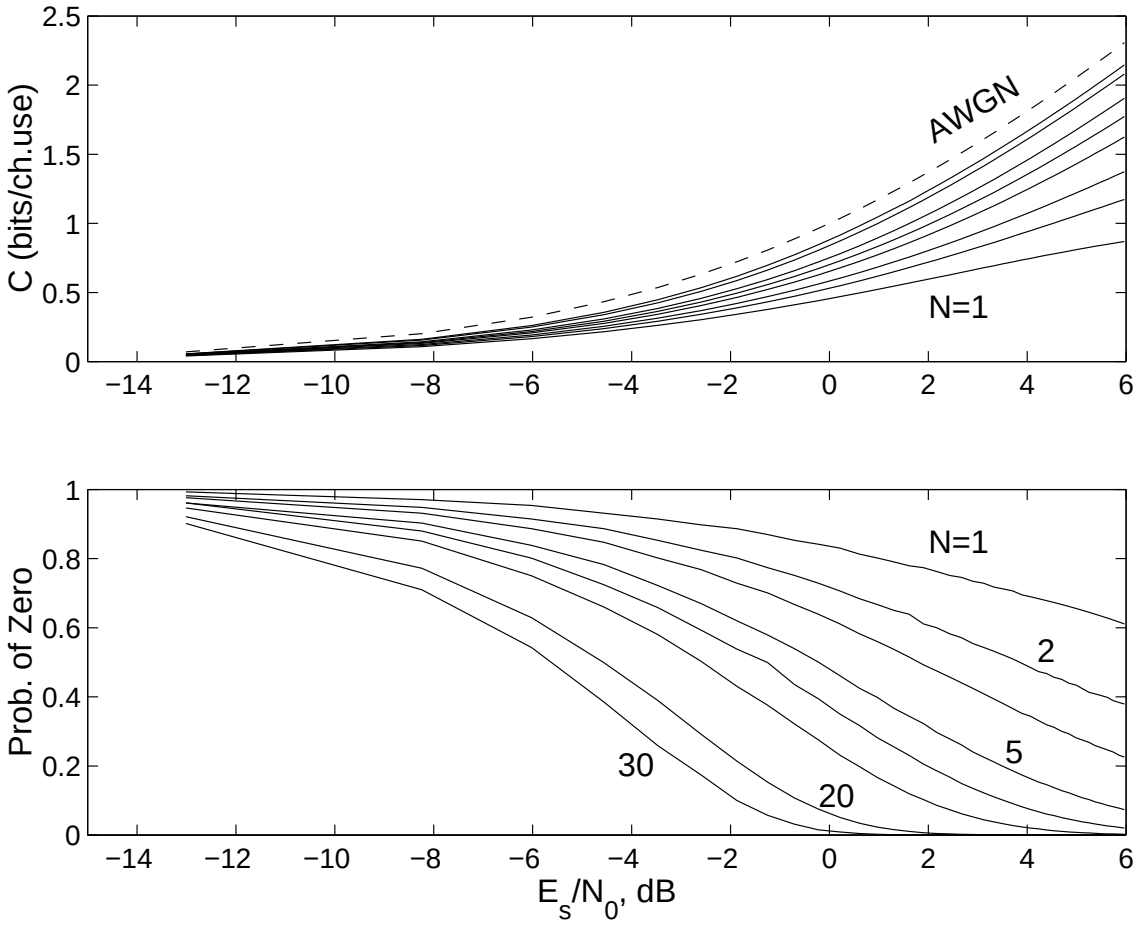


Figure 3.2: Capacity and optimal probability of zero versus E_s/N_0 for $N = 1, 2, 3, 5, 7, 10, 20, 30$. The dashed curve represents the capacity of the coherent AWGN channel.

obtained by using the above described two-mass-point density (*i.e.*, the mass point at zero in addition to a nonzero mass point), versus an input density consisting of a single sphere. The logarithm of base two and the following normalization for the capacity is used throughout

$$C = \frac{I(\mathbf{x}, \mathbf{y})}{N} \qquad \text{(bits/complex dimension)}. \tag{3.41}$$

Fig. 3.2 presents capacity versus coded bit SNR $E_s/N_0 = \mathsf{P}/(NN_0)$ curves for

channel [27, 23]. In this work, however, we use it in the more limited sense described above.

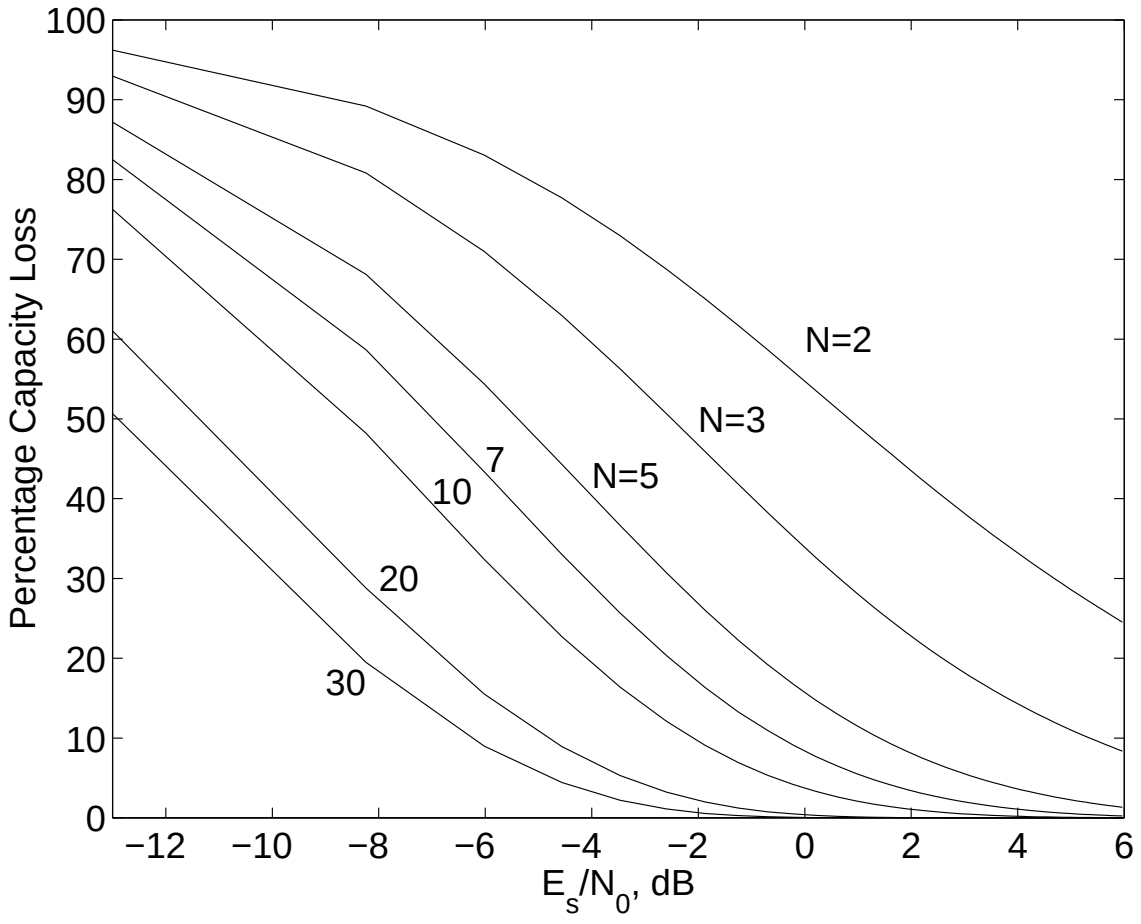


Figure 3.3: Percentage loss in capacity as a result of not using shaping.

several values of N, together with the maximizing probability of zero. Also shown in this figure is the capacity of the coherent AWGN channel for comparison. As can be seen, for low SNR values higher probability of zero is required and with increasing SNR and N this probability decreases. In Fig. 3.3, the relative loss in capacity resulting from not utilizing the zero mass point is quantified. It is observed that a significant loss of as much as 50% occurs when information is transmitted using points on the surface of a single sphere (*e.g.*, for E_s/N_0 in the range of 0-6 dB). This observation will be our starting point for developing capacity-achieving codes in the next chapter.

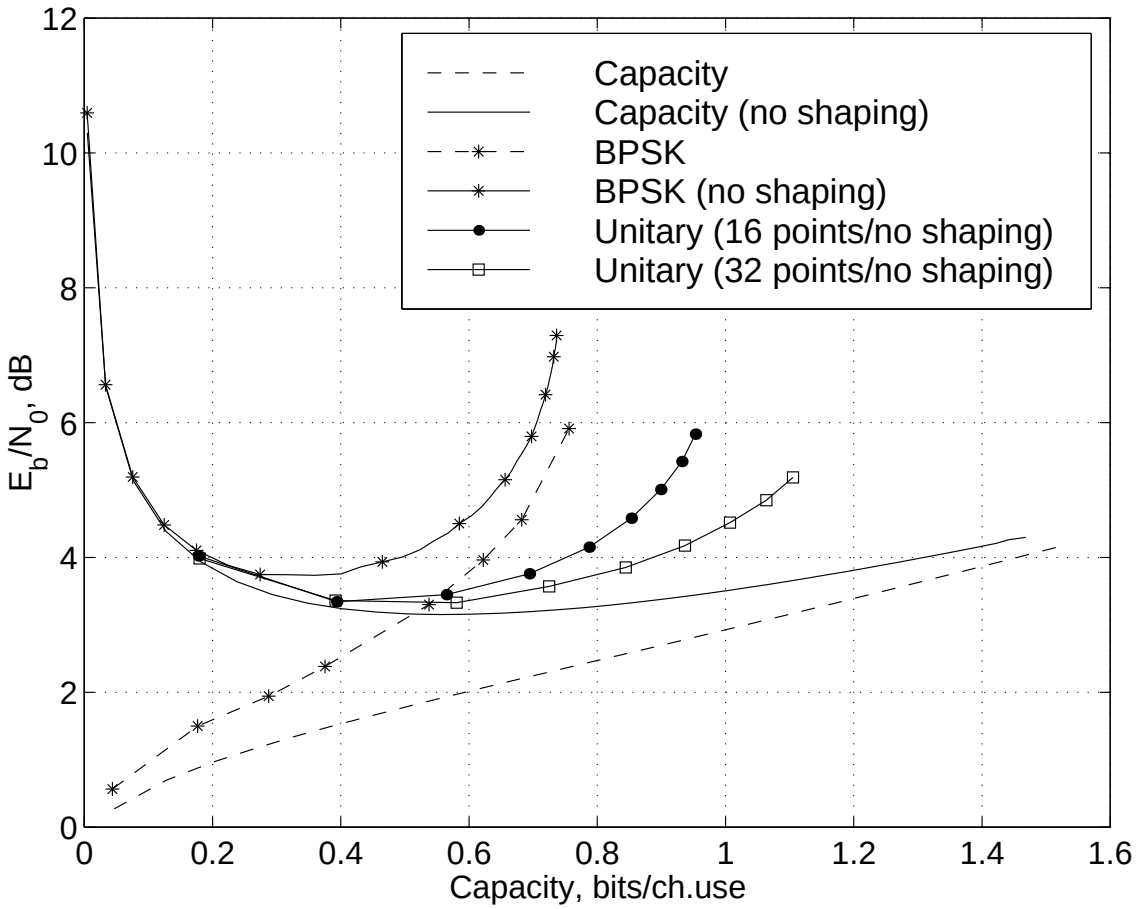


Figure 3.4: E_b/N_0 required to operate at rate equal to capacity for $N = 4$ and several modulation schemes.

In practical situations one can utilize fixed modulations, that approximate the circularly symmetric structure of the capacity achieving distribution suggested by Lemma 3.1. Fig. 3.4 depicts the minimum required information bit SNR $E_b/N_0 = (E_s/N_0)/C$ versus capacity for several modulation-constrained transmission schemes for a fixed channel coherence time $N = 4$.

In particular, the achievable capacity for BPSK signaling was evaluated with and without shaping. We also evaluated the performance of the more advanced "unitary modulation" schemes of [34]. The capacity curves, with and without shaping, with isotropic signaling (no discretization) are also plotted for comparison. It can be ob-

served that shaping is important at low rates, *e.g.*, at $C = 1/2$, shaping results in a performance gain of 1.4 dB and 1 dB for the isotropic and BPSK modulations, respectively. Unitary constellations seem promising (*i.e.*, they result in low discretization loss) as they perform close to the capacity curve (without shaping) for low to medium rates. In the next chapter, we try to capture these gains through specific code designs that imitate the maximizing input distributions described above.

3.5 Conclusion

Capacity and the structure of the capacity-achieving density for the noncoherent AWGN channel was investigated. It was proven that the capacity achieving input density is circularly symmetric, discrete in amplitude with infinitely many mass points, one of which is at the origin for any channel coherence time N. Asymptotic results were derived that further justify the well-known sphere hardening effect, and several numerical results were presented. In addition, two performance measures were identified: discretization loss that originates from using discrete modulation alphabets and shaping gain that occurs when using multiple amplitude levels. Numerical optimization was performed to quantify these performance measures for different transmission schemes. Apart from their theoretical value, these results can significantly aid the code/modulation design as will be demonstrated in the next chapter.

CHAPTER IV

CAPACITY-INSPIRED CODING FOR THE NONCOHERENT CHANNEL

4.1 Introduction

In this chapter we present coding and modulation design techniques that are inspired by the capacity studies presented in Chapter III. It is shown in Chapter II that for high values of channel coherence time N, pilot-symbol-assisted modulation paired with AWGN codes performs fairly well on this channel, while for small or moderate values of N this approach results in considerable performance loss. Therefore, our focus here is on coding for fast noncoherent channels, *i.e.*, for small and moderate values of N. Unfortunately, unlike the case of coherent AWGN channel, no satisfactory metric for code design has been found for the noncoherent channel. In particular, for a code spanning over L blocks of length N, it can be shown that the pairwise error probability between two codewords $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_L]$ and $\mathbf{c}' = [\mathbf{c}'_1, \mathbf{c}'_2, \ldots, \mathbf{c}'_L]$ depends on the block-wise Hamming distance (*i.e.*, the number of different blocks), and the absolute value of the cross correlation between these blocks

$$\rho_k = \frac{|\mathbf{c}_k^H \mathbf{c}'_k|}{\|\mathbf{c}_k\| \|\mathbf{c}'_k\|}. \tag{4.1}$$

However, it is unknown exactly how these two parameters are combined to form the pairwise error probability[1]. For example, the pairwise error probability between two codewords that differ in only one block (*i.e.*, having block-wise Hamming distance equal to one) with cross correlation of $\rho = 0$ in that block (*i.e.*, orthogonal) is essentially the same (over a large range of SNR values) as that of two codewords that differ in two blocks with correlation of $\rho = 0.65$ in each of the blocks, or with $\rho = 0.75$ in three blocks. Furthermore, even if such an expression for the pairwise error probability were known, it is not clear if this would be a good code-design metric for close-to-capacity performance.

The proposed codes are concatenated designs consisting of outer binary codes of rate k/l, and simple inner *modulation* codes which map l bits onto $Q = 2^l$ complex sequences of length N. The purpose of the outer code is to introduce memory between the blocks of length N, while the purpose of the inner code is to guarantee small cross correlation between the complex sequences used in each block. This approach is motivated by the conjecture reached in [55] that LDPC – or in general turbo-like – codes can achieve capacity for a wide range of channels if suitable modulation schemes are designed to interface with the channel. The overall throughput of these codes is $R = k/N$ (bits/complex dimension). Observe that for a given throughput R, the parameter l is a design choice. A large l results in a more powerful outer code, while a small l results in smaller modulation alphabets, and thus, smaller cross correlation between blocks. The choice of l in the proposed codes is suggested by the numerical results of Section 3.3 (*e.g.*, see Fig. 3.4).

Regarding the modulation code, our approach is to design alphabets that imitate

---

[1]Such asymptotic expressions for the pairwise error probability were the starting point for the design of good codes for the MIMO block-independent fading channel in [34].

the input distribution that maximizes the mutual information, as characterized in Section 3.2. Towards this goal we concentrate on modulation schemes that utilize the zero mass point, as well as points on the surface of a single N-dimensional sphere. More complex modulation schemes (*e.g.*, ones that use points on multiple spheres) are not considered herein; this approach is justified by the asymptotic and numerical results of Section 3.3. Furthermore, we consider the case of small N separately from the case of moderate N. This is done because complexity considerations at the demodulator suggest different designs in these two cases. In particular, for small N, we concentrate on highly optimized modulation codes (*e.g.*, *unitary modulation*-based schemes of [34]), while for larger values of N we concentrate on modulation schemes that are "separable", that is, modulation schemes that are in the form of a Cartesian product of simpler modulations (*e.g.*, M-PSK modulations).

4.2 Small channel coherence time

In this operating scenario the goal is to utilize highly optimized modulation alphabets in conjunction with powerful serially concatenated convolutional codes (SCCC).

The problem of designing good modulation alphabets, can be stated as finding Q sequences in N complex dimensions with small maximum cross correlation, as defined in (4.1). This is a long-standing problem, and so far no satisfying answer has been provided regarding the best achievable solution (see [93, 37] for a theoretical discussion of this problem). However, several techniques have been suggested in the literature that result in good designs. One such method, named unitary modulation, was proposed in [34], where certain columns of the discrete Fourier transform matrix

are used as the codewords, namely

$$\mathbf{x}_m = \begin{bmatrix} 1 \\ e^{j\frac{2\pi}{Q}u_1 m} \\ e^{j\frac{2\pi}{Q}u_2 m} \\ \vdots \\ e^{j\frac{2\pi}{Q}u_{N-1} m} \end{bmatrix} \qquad m = 0, 1 \ldots Q-1, \tag{4.2}$$

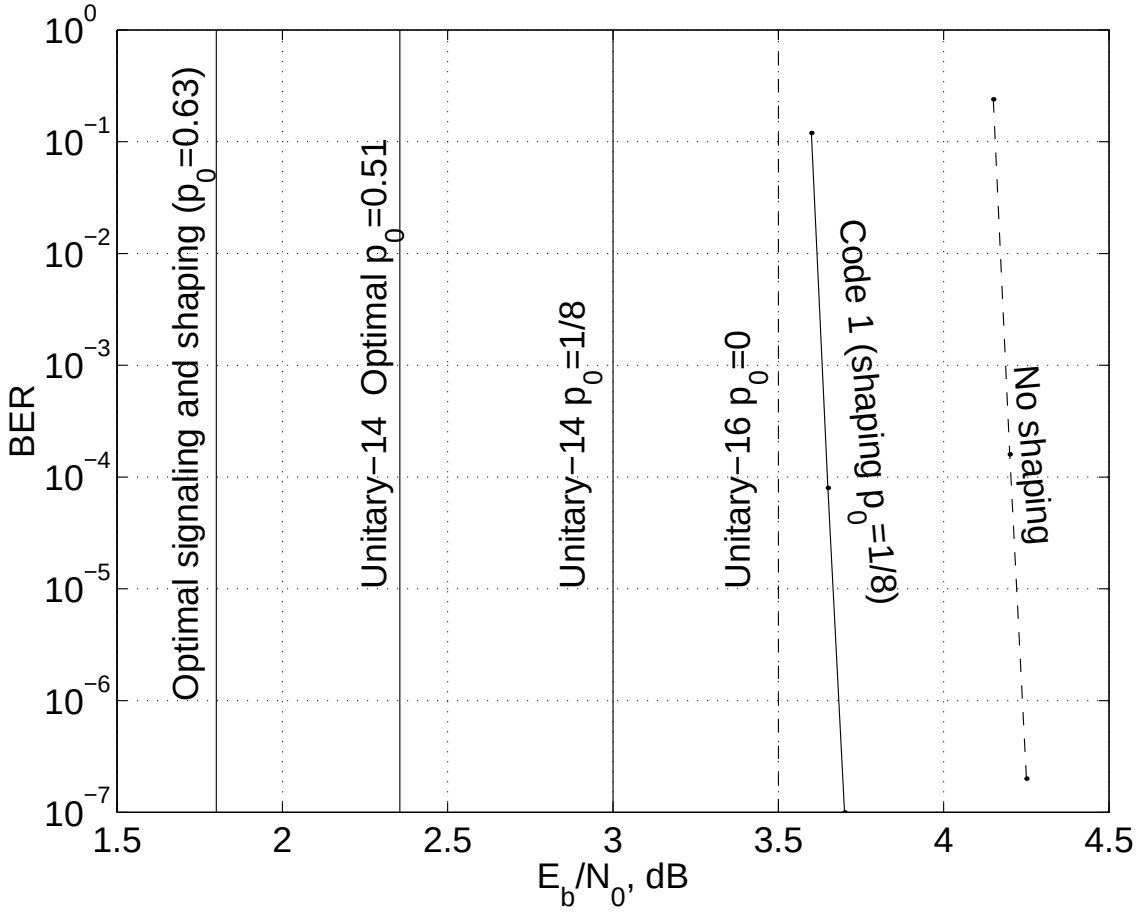
where $0 < u_1 < u_2 \ldots u_{N-1} \leq Q-1$ is a sequence of integers selected such that the maximum cross correlation is minimized. The unitary modulations approximate the isotropic distribution suggested by the capacity results. In this work, shaping gain is achieved by augmenting the unitary modulations with the zero sequence.

Regarding the design of the binary code, standard rules for SCCC design [5, 20] can be adapted to the noncoherent AWGN scenario. In this case, the modulation code can be incorporated into the inner code structure without any increase in complexity. In addition, the introduction of the zero sequence is handled by the inner code. A standard soft-input/soft-output (SISO) decoder [6] can be utilized for iterative decoding. The memory requirements for the inner SISO are proportional to the size Q of the output alphabet used, while the complexity depends on the number of transitions in the inner trellis.

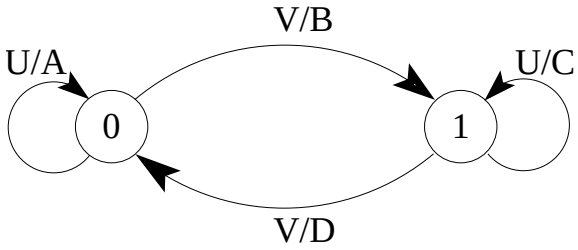
The same set partitioning and code design techniques of Ungerboeck [86] can be employed for the design of the inner code. Namely, the Q modulation codewords can be partitioned into sets with decreasing maximum cross correlation within the partitions. It is noted that the introduction of the zero sequence in the modulation alphabet imposes a complexity constraint on the inner code. In particular, in order to achieve a high probability of zero (as suggested by the numerical results in the previous section) the zero sequence should appear frequently in the trellis. However,

this requires a corresponding increase in the trellis complexity of the inner code, because otherwise the code will exhibit a catastrophic behavior.

Example 4.1. As an example of such a scheme we consider a noncoherent code with throughput 1/2 (bits/complex dimension). It is a serially concatenated turbo code with a rate-2/3 16-state outer code (this code was designed in [20] and has a generator matrix $G(D)$ shown in Fig. 4.1) and a rate-3/4 2-state inner code followed by a 14-point unitary modulation in addition to the mass point at zero in $N = 4$ complex dimensions. The optimized 14-point unitary modulation is designed according to (4.2) (with $[u_1, u_2, u_3] = [8, 10, 13]$ found using numerical optimization) and has the property that the cross correlation between the signals of the same parity is $\rho_{\text{even}} = 0.35$, and for opposite parity $\rho_{\text{odd}} = 0.5$. This way a two-way partition of this signal set into even and odd signals can be formed where each subpartition has a better cross correlation than the original modulation. In addition, these partitions can be further divided into sets of 4 signals with the addition of the zero mass point, that is, if $\emptyset$ denotes the zero mass point, then the four-way partition is $A = \{0, 2, 4, 6\}$, $B = \{8, 10, 12, \emptyset\}$, $C = \{1, 3, 5, 7\}$, $D = \{9, 11, 13, \emptyset\}$. These sets are assigned to the four transitions in the 2 state inner code, with input bit mapping designed to make the inner code recursive (the state diagram of the inner code is shown in Fig. 4.1). In order to avoid catastrophic behavior special care needs to be taken to assign the zero mass point to the same input binary word. This way the probability of zero achieved is $p_0 = 1/8$, assuming that all the input binary words at the input of the inner code are equally likely. It is worth noting that for this channel the optimal p_0 (for a two mass point density) was found to be as high as 0.63. Thus the designed modulation code can be improved by increasing the frequency of the zero mass point transmission. However, to achieve such high probability of zero without making the



$$G(D) = \begin{bmatrix} 1 & 1+D & 1+D+D^2 \\ 1+D+D^2 & 1+D+D^2 & 1 \end{bmatrix}$$



$$U = \{000, 011, 101, 110\} \quad A = \{0, 2, 4, 6\} \quad C = \{1, 3, 5, 7\}$$
$$V = \{001, 010, 100, 111\} \quad B = \{8, 10, 12, \emptyset\} \quad D = \{9, 11, 13, \emptyset\}$$

Figure 4.1: BER versus E_b/N_0 for the SCCC design (a length 6,000 bit interleaver was used). Vertical lines correspond to E_b/N_0 required suggested by the capacity result for different modulation and shaping. Also shown are the generator matrix and state diagram of the outer and inner codes, respectively.

code catastrophic requires the increase in the number of states in the inner code. Since we are interested in low-complexity implementations, we have not explored this option further.

The bit error rate (BER) versus information bit SNR (E_b/N_0) curve is plotted in Fig. 4.1. In the same figure, the minimum E_b/N_0 required for error-free transmission for the isotropic modulation as well as for the specific modulation utilized by this code is plotted. For comparison, a similar code was designed that utilizes only a single amplitude, by means of a 16-point unitary modulation (the modulation is specified by (4.2) with $[u_1, u_2, u_3] = [2, 11, 15]$). It can be observed from the simulation results that the E_b/N_0 gain by using a code with zero mass point transmission is essentially the same (0.5 dB) as suggested by the capacity results. In addition, the performance of this code (at a BER of 10^{-6}) is 0.7 dB away from the constraint capacity for the same p_0, and 1.5 dB away from the constraint capacity for optimized p_0. □

4.3 Moderate channel coherence time

As discussed in the previous section, the complexity of the demodulator is linear with the alphabet size, and thus exponential with the channel coherence time N (for a fixed transmission rate). As a result, the schemes described in the previous section might not be desirable for moderate values of N. One solution is to use modulation schemes with additional structure and take advantage of this structure at the demodulator. Examples of these codes include the codes introduced in [32] which are based on transforming linear codes using the Cayley transform. Here we take a different approach. The modulation scheme used incorporates the mass point at zero into the modulation code and yet allows demodulation with complexity linear in the channel coherence time N. This scheme is described in more detail in

the following.

Modulation is performed using m blocks of length N at a time as shown in Fig. 4.2(a). Of these m blocks, $m - \lfloor mp_0 \rfloor$ blocks consist of M-PSK symbols, each having energy E_s, and in the remaining $\lfloor mp_0 \rfloor$ blocks no transmission is taking place (zero mass point). These blocks will be referred to as the "zero blocks". The first symbol in each block—other than the zero blocks—is a pilot symbol of energy E_p (which can be different from the coded M-PSK symbol energy E_s). The position of the zero blocks is not fixed but can be any of the $\binom{m}{\lfloor mp_0 \rfloor}$ available positions. This is accomplished by utilizing $b = \lfloor \log_2(\binom{m}{\lfloor mp_0 \rfloor}) \rfloor$ bits, referred to as the "mode bits" that determine which $\lfloor mp_0 \rfloor$ out of the m blocks will be the zero blocks. The remaining $(m - \lfloor mp_0 \rfloor)(N-1)$ complex dimensions available for information transmission are used for the transmission of M-PSK symbols, thus carrying a total of $(m - \lfloor mp_0 \rfloor)$(N-1)$\log_2(M)$ coded bits. This transmission scheme is depicted in Fig. 4.2, and the overall throughput of this modulation code is

$$\frac{b + (m - \lfloor mp_0 \rfloor)(N-1)\log_2(M)}{mN} \quad \text{(coded bits/complex dimension)}. \tag{4.3}$$

The overall code is completed by using outer powerful irregular LDPC codes, optimized for the particular choice of the channel and modulation parameters.

The main advantages of this simple scheme are: (i) an arbitrary probability of zero, p_0, is obtained without the problem of catastrophic behavior mentioned in the previous section, and (ii) it permits the use of an iterative (message-passing) decoder that provides soft decisions to the outer code, and has linear complexity in N. In the following, the details of the demodulator that operates with linear complexity are described for the special case of $mp_0 = 1$, *i.e.*, for one zero block in m blocks. This special case is chosen for notational simplicity; however it conveys the basic idea of

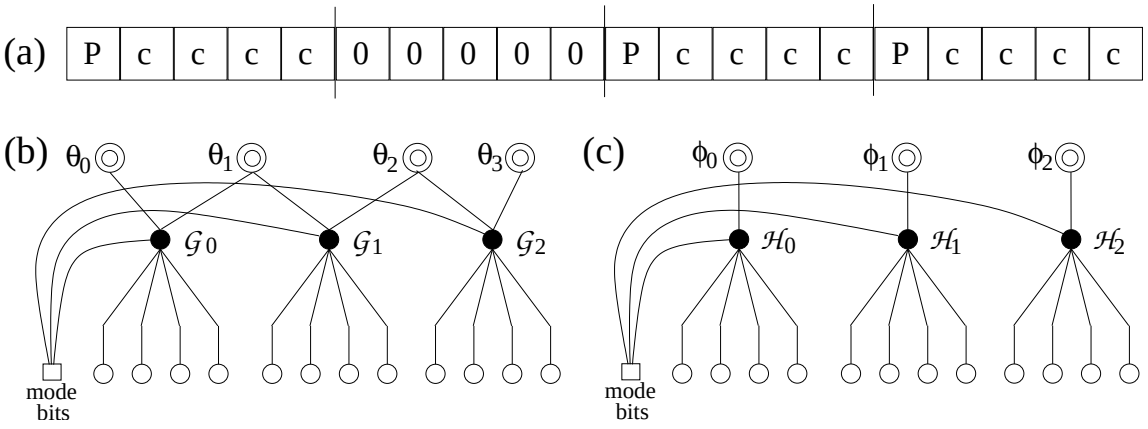


Figure 4.2: (a) Pilot-symbol-assisted transmission scheme for $b = 2$ ($m = 4$ blocks). Pilot symbols are denoted by "P", coded symbols by "c", and no transmission by "0". In this figure, the zero block is in the second position. (b) Factor graph representation of proposed modulation scheme for $b = 2$ implied by (4.4). (c) Modified factor graph representation of proposed modulation scheme for $b = 2$ implied by (4.7).

this approach.

The generic factor graph representation of the demodulator for this modulation code for the case of $b = 2$ ($m = 4$) is presented in Fig. 4.2(b). Circles represent the M-PSK variables $\mathbf{x}_k = [x_{k,1} \dots x_{k,N-1}]$ for $k = 0, \dots, m-2$, double circles represent the phase variables, and squares denote the variable w representing the b mode bits (taking values in $0, \dots, m-1$). The likelihood function to be marginalized can be factored as

$$\mathcal{G}(\mathbf{x}_0, \dots \mathbf{x}_{m-2}, w, \theta_0, \dots \theta_{m-1}) = \prod_{k=0}^{m-2} \mathcal{G}_k(\mathbf{x}_k, w, \theta_k, \theta_{k+1}), \tag{4.4}$$

where

$$\mathcal{G}_k(\mathbf{x}_k, w, \theta_k, \theta_{k+1}) = \exp(\sqrt{E_p} z_{q,0} e^{-j\theta_q}) \prod_{i=1}^{N-1} \exp(\sqrt{E_s} z_{q,i} x_{k,i}^* e^{-j\theta_q}), \tag{4.5}$$

$z_{q,i}$ is the i-th received symbol in the q-th block, with $z_{q,0}$ being the pilot symbol in that block (the transmitted pilot symbol is assumed to be 1), and θ_q (for $q =$

$0, \ldots, m-1$) are the unknown phase variables in each of the m blocks. The function $q(k, w)$ quantifies the random position of the zero block and is defined as

$$q = q(k, w) = \begin{cases} k, & \text{if } k < w; \\ k+1 & \text{if } k \geq w. \end{cases} \tag{4.6}$$

The existence of the function $q(k, w)$ in $\mathcal{G}(\cdot)$ results in a factor graph with cycles, as shown in Fig. 4.2(b). However, cycles in the factor graph are undesirable when message-passing algorithms are performed. Fortunately, cycles can be eliminated without increasing the complexity of a message-passing algorithm by taking advantage of the fact that only $m-1$ of the blocks are *active* at any time. This is true since for fixed w, variables in the w-th block do not appear in (4.4); hence, one of the phase variables can be eliminated. This way the above factor graph reduces to the factor graph shown in Fig. 4.2(c), where the equivalent likelihood function to be marginalized can be written as a product of $m-1$ simpler functions

$$\mathcal{H}(\mathbf{x}_0, \ldots, \mathbf{x}_{m-2}, w, \phi_0, \ldots, \phi_{m-2}) = \prod_{k=0}^{m-2} \mathcal{H}_k(\mathbf{x}_k, w, \phi_k) \tag{4.7}$$

with

$$\mathcal{H}_k(\mathbf{x}_k, w, \phi_k) = \exp(\sqrt{E_p} z_{q,0} e^{-j\phi_k}) \prod_{i=1}^{N-1} \exp(\sqrt{E_s} z_{q,i} x_{k,i}^* e^{-j\phi_k}), \tag{4.8}$$

where the new $m-1$ phase variables ϕ_k are used to emphasize that these are the only phases *active* at each particular sequence of m blocks. Furthermore, since each $\mathcal{H}_k(\cdot)$ is a separable function in $x_{k,i}$'s, the marginalization over these variables can be done with linear complexity, which can be observed through the following chain

of equations

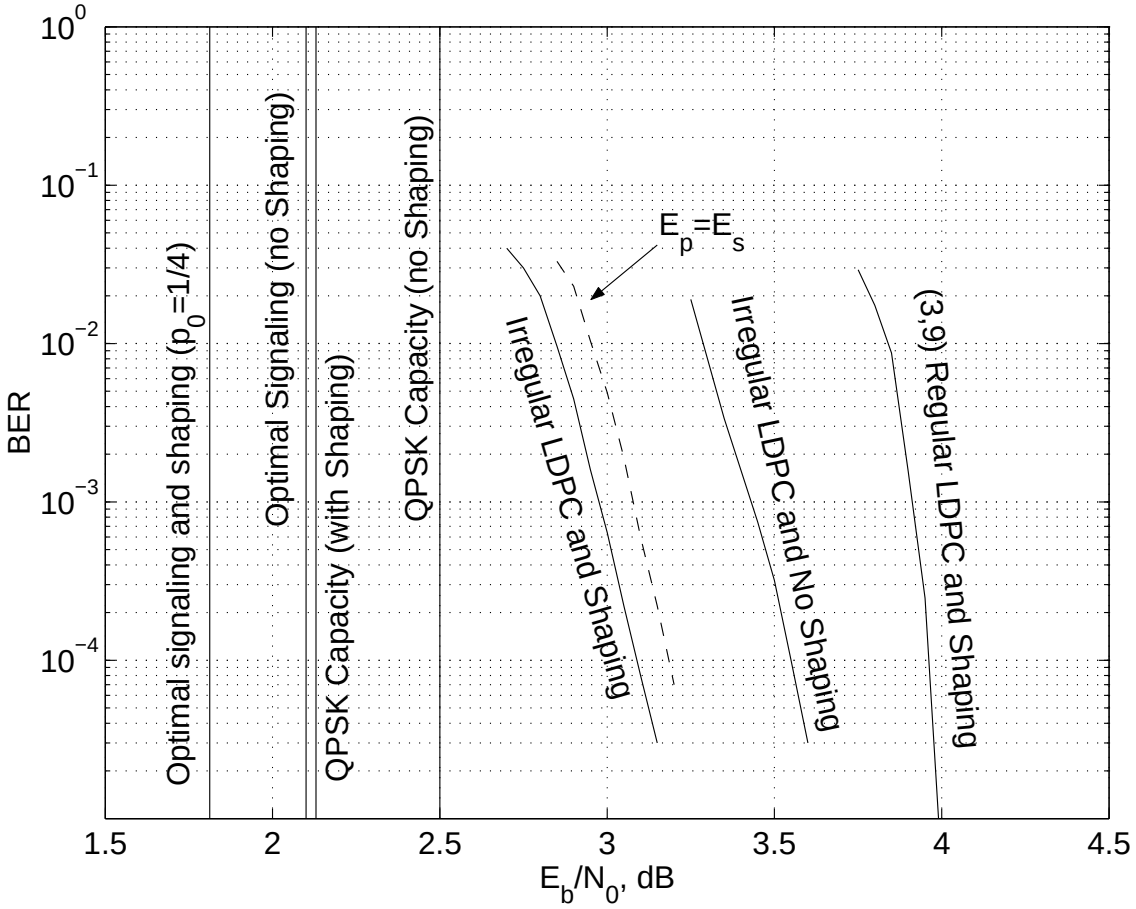
$$\sum_{\phi_k,\mathbf{x}_k} \mathcal{H}_k(\mathbf{x}_k, w, \phi_k) = \sum_{\phi_k} \exp(\sqrt{E_p} z_0^q e^{-j\phi_k}) \sum_{\mathbf{x}_k} \prod_{i=1}^{N-1} \exp(\sqrt{E_s} z_{q,i} x_{k,i}^* e^{-j\phi_k})$$
$$= \sum_{\phi_k} \exp(\sqrt{E_p} z_{q,0} e^{-j\phi_k}) \prod_{i=1}^{N-1} \Big(\sum_{x_{k,i}} \exp(\sqrt{E_s} z_{q,i} x_{k,i}^* e^{-j\phi_k}) \Big). \quad (4.9)$$

Here, the summation of M^{N-1} terms (over all $x_{q,i}$ variables) is substituted by the product of $N-1$ summations with M terms each, thus reducing the complexity from exponential to linear in N. In particular, it can be shown that the precise complexity per channel symbol of a message-passing algorithm operating on the factor graph of Fig. 4.2(c) is increased by $2(m-1)/m$, when compared to a modulation scheme using the same M-PSK alphabet but no zero block. Also note that summation over the phase variables, can be efficiently implemented by appropriate discretization of $[0, 2\pi)$ as noted in Section 2.5 of Chapter II.

Example 4.2. As a specific example, we consider the above described modulation scheme with QPSK (*i.e.*, $M = 4$) and $b = 2$, thus achieving the probability of zero $p_0 = 1/4$. For the case of $N = 7$, this probability is the optimal probability of zero as suggested by the numerical results of Section 3.3. The outer code considered is an irregular LDPC code with rate 24/38, while the modulation code has throughput of 38/28 (bits/complex dimension), as obtained from (4.3). The resulting throughput of the overall code is 6/7 (bits/complex dimension)[2].

The irregular LDPC code was optimized by an algorithm very similar to the differential evolution idea of [84, 81, 35], resulting in the variable and check degree distribution indicated in Fig. 4.3. The only difference is that instead of density

[2]This throughput is selected so that the results can be compared in a fair way to the best known codes in the literature [36]. These codes have a nominal throughput of 1 (bit/complex dimension), but assume that there is an overlap of 1 complex dimension between adjacent blocks.



v	N_v
1	2202
2	2886
3	1966
4	1478
10	177
15	411

c	N_c
5	1154
9	111
10	2095

Figure 4.3: BER versus E_b/N_0 for the LDPC code (a codeword length of 9120 bits was considered). Vertical lines correspond to E_b/N_0 required suggested by the capacity result for different modulation and shaping. (N_v, N_c) are the number of variable and check nodes with degree (v, c) respectively.

evolution for the testing part we use the simulated performance of the particular code for fixed codelength and number of packets. The pilot energy is optimized as well in order to minimize the BER (see Chapter II for details). Performance curves and LDPC code details are depicted in Fig. 4.3.

The performance of the proposed code is compared to a code that does not use zero block transmission and pilot power optimization, but uses an (optimized) irregular LDPC code of rate 1/2 (with the same complexity) and direct QPSK modulation yielding a throughput of 6/7 (bits/complex dimension). Regarding the performance of these codes, we observe that zero block transmission and pilot power allocation yield an E_b/N_0 gain of around 0.4 dB. A code similar to the simulated no-shaping code using differentially encoded QPSK was proposed in [36]. The code proposed herein outperforms such a system by at least[3] 0.3 dB.

In order to isolate the effect of pilot symbol optimization from that of shaping, the performance of the proposed code with $E_p = E_s$ is also shown. It is observed that pilot optimization is responsible for only a fraction of the gain, *i.e.*, 0.1 dB.

It is also noted that on the noncoherent AWGN channel, irregular LDPC code optimization has a larger impact on the performance compared to the case of coherent AWGN channels. To illustrate this point, we have simulated a regular LDPC code (with variable and check node degrees 3 and 9, respectively) with shaping and pilot optimization and compare it with the best proposed code. As can be seen form the figure, there is a loss of more than 1 dB when compared to the irregular case.

---

[3]It is noted that the codes in [36, 70] assume a more favorable channel model than the block-independent channel considered here. For a given code, the performance gain due to this more favorable channel can be as much as 0.7 dB, as observed via simulations of the code in [36] on this more favorable model and the block-independent model. The gains reported here are in addition to this gain. Therefore, for a fair comparison, 0.7 dB should be added to the gains reported herein.

In coherent AWGN channel however, the difference between regular and optimized irregular codes of roughly the same complexity is smaller (see for instance the codes in [76]). □

4.4 Conclusion

Coding for the block-independent noncoherent channel was investigated in this chapter. More precisely, theoretical and numerical results of Chapter III were applied to code and modulation design for this channel. Due to the complexity constraints, the case of small channel coherence time N, was treated separately from the case of moderate values of N. In particular, two coding/modulation schemes with distinct advantages were designed that imitate the signaling structure suggested by the capacity results of Chapter III. Two examples of these coding schemes were presented and their performance was evaluated through simulations. Binary code optimization was also performed to match the channel characteristics, and it was shown that the latter plays a significant role in the overall performance. These combined code/modulation designs yield practical transmission schemes that outperform the best known codes so far.

CHAPTER V

ROTATIONAL INVARIANCE AND ROBUSTNESS

5.1 Introduction

In the previous chapters we dealt with information-theoretic and practical aspects of communication over the block-independent noncoherent channel model. Two different approaches to coding were investigated: in Chapter II a simple modulation scheme was analyzed that allowed the utilization of a large collection of off-the-shelf AWGN binary codes, while in Chapter IV codes were designed directly for the specific channel model under consideration.

In this chapter noncoherent communication is considered from a slightly different perspective. In that, unlike the previous chapters, we do not assume the rigid block-independent channel model, but rather consider more relaxed modeling of the phase process. We investigate certain properties of the coding scheme that, under some circumstances, make the coding scheme less susceptible to the harmful effects of random phase. We start with identifying the potential problem when communicating over the noncoherent channel, which forms one of the main reasons to consider the aforementioned properties.

Regardless of the particular structure of the coding scheme, a potential problem

can occur when communicating over the noncoherent channel. This problem can be explained by observing that if two sequences of length N differ by a mere phase rotation, they will be indistinguishable at the receiver side. Therefore, such pairs should be avoided during the transmission as they will lead to a catastrophic behavior. On the other hand, if all such sequences are included in the codebook and are assigned the same input sequence, then the catastrophic behavior is avoided. This is true since no input errors will occur as a result of a random phase rotation. Coding schemes that resolve the inherent phase uncertainty in the transmitted sequence in this particular way are called rotationally invariant (RI), and are essentially generalizations of the differential phase encoding to higher signaling alphabets.

Because rotational invariance is a property of interest in both the coherent and the noncoherent setting, in this chapter we consider the behavior of RI codes for several channel models. We start by considering rotational invariance in the traditional setting of a channel with slow phase dynamics. In this setting we investigate RI code design for serially concatenated turbo codes. These codes are then investigated in the presence of faster phase dynamics. It is observed that a stronger property than rotational invariance, namely *rotational robustness* (RR), is required to avoid noncatastrophic behavior in these channels.

In the remainder of this section we briefly describe rotational invariance, rotational robustness and the channel models that are relevant in the investigation of each of these properties.

5.1.1 Rotational Invariance

Rotational invariance was originally introduced in the context of *coherent* detection and/or decoding, as a method for resolving the phase ambiguities, resulting

from non-pilot-aided carrier-phase estimation, and rotationally symmetric constellations. More specifically, if the particular modulation used in a coded scheme is unchanged by phase rotations through some base angle ϕ (*e.g.*, M-PSK signal set is unchanged by phase rotations by an integer multiple of $2\pi/M$), then any phase estimation scheme, which does not utilize pilot symbols, can only estimate the phase modulo ϕ. The overall model for this channel is one in which, apart from corruption by additive white Gaussian noise, the signal undergoes a *discrete* (*i.e.*, integer multiple of $2\pi/M$) phase rotation which remains *constant* over the entire codeword. In the following, this channel model will be referred to as the *discrete constant rotation* (DCR) channel.[1] It is noted that the maximum likelihood sequence detection (MLSD) receiver corresponding to the DCR channel is a complicated algorithm, due to the need for averaging the sequence likelihoods over all possible discrete rotations. This last observation was the main motivation for introducing rotationally invariant (RI) codes, which, due to their symmetrical structure, enable near-optimal decoding in the DCR channel even when conventional coherent decoders are used (*e.g.*, the Viterbi algorithm (VA) for the case of trellis-based RI codes).

An extensive literature on RI trellis-coded modulation (RI-TCM) schemes exists. The pioneering work in this area was done by Wei [89, 90, 91], whose ideas were later applied by many authors to TCM code construction. In [11, Ch. 8], differential encoding schemes that achieve rotational invariance were investigated. The theoretical framework underlying RI codes was presented in [85], together with necessary and sufficient conditions for a code and an encoder to posses such a property, while [7] investigated rotational invariance of group codes. In [49], sufficient conditions for

[1] A summary of all channel models considered herein is given in Table 5.1.

designing RI serially concatenated TCM (SCTCM) codes were derived; however, the suggested codes can only be viewed as examples, rather than powerful RI-SCTCM designs.

In the first part of this chapter several powerful RI-SCTCM codes are designed that utilize 8-PSK and 16-QAM constellations, thus resulting in high bandwidth efficiency. The performance of the proposed designs is evaluated through simulations.

5.1.2 Rotational Robustness

Although the previous work on RI codes implies an underlying DCR channel, all practical channels are more complicated than the simplified DCR model. Of particular interest is the case when, due to a noise burst, the external phase estimator looses and reestablishes phase lock after some brief acquisition time. This phenomenon, known as a cycle slip, can be modeled by a channel that, corrupts the initial part[2] of the codeword only with AWGN, and, in addition, introduces a *discrete* phase rotation to the rest of the transmitted codeword. To distinguish from the DCR model, this more elaborate channel model is referred to as the *discrete partial rotations* (DPR) channel. Since MLSD decoding matched to the DPR channel is extremely complicated (the sequence likelihoods need to be averaged over all possible cycle-slip positions and all possible discrete rotations), it is desirable to design codes that are robust to these cycle slips when conventional, coherent decoders are utilized. Codes that are robust to cycle slips, *i.e.*, result in a small number of errors when transmitted through a DPR channel, and decoded by coherent detectors, will be referred to as rotationally robust (RR) codes. The significance of RR code design

[2]In this model, the position of the cycle slip is assumed uniformly distributed within the codeword.

can be demonstrated for a system that is designed such that the bit error rate (BER) is dominated by the probability of loosing lock, P_{ll} (*i.e.*, the BER in the tracking mode is small compared to P_{ll}). In such a system, when a cycle slip occurs during the transmission of a length L codeword of a non-RR code, approximately half of the bits will be in error with high probability, resulting in an approximate BER of $P_{ll}/2$. For an RR code on the other hand, only a small number of bits will be affected by the cycle slip, resulting in an approximate BER on the order of P_{ll}/L.

In this work, it is shown that although rotational invariance is, in general, a weaker property than rotational robustness, RI codes with small state space possess this stronger property and thus they are also RR. In particular it is shown that for an RI code, the number of input bit errors, due to a cycle slip, is upper bounded by a relatively small number that increases monotonically with the trellis size of the code, but does not depend on the codeword length. Although this statement directly implies that RI-TCM codes are also RR, it does not hint as to whether RI codes with large overall state space, *e.g.*, RI-SCTCM codes, possess rotational robustness or not. An incomplete answer to this question is given by simulating the specific RI-SCTCM codes designed herein on the DPR channel, and showing that the addition of a simple stopping criterion to the coherent iterative decoding algorithm is sufficient for rotational robustness in the DPR channel.

An even more interesting channel is the noncoherent AWGN channel which introduces an arbitrary (as opposed to discrete) phase rotation on the transmitted sequence. This model is well suited for receivers that perform phase estimation and decoding jointly. Clearly the DCR and DPR models cannot adequately describe this channel, thus the *arbitrary constant rotations* (ACR) and the *arbitrary partial rotations* (APR) models are introduced for the case of constant and partial arbi-

trary phase rotations throughout the transmitted codeword, respectively. It is a well known fact [75] that for transmission over the ACR channel the code should either be *anti-RI*, *i.e.*, no two codewords are rotated versions of each other, or it should be RI, namely any rotation of a codeword is a valid codeword and corresponds to the same input sequence. Any other possibility that does not fit into one of these categories will lead to catastrophic behavior, since rotated codewords cannot be distinguished on the receiver side in the ACR channel. However, the question of which of these two classes performs better is still unanswered.

In the last part of this chapter the performance of these two classes is investigated. In particular, it is proven that MLSD decoding matched to the ACR model is identical for RI and anti-RI codes, thus resulting in the same performance. At this point one might ask whether any of these classes of codes is preferable. A partial answer to this question is provided here by showing that non-concatenated RI codes are provably robust to partial arbitrary rotations, *i.e.*, are provably robust in the APR channel.

The rest of the chapter is structured as follows. In Section 5.2, we briefly present the theoretical background for studying RI codes, along the lines of [85]. The design of powerful RI-SCTCM is presented in Section 5.3. Rotational robustness in connection with the DPR channel is discussed in Section 5.4, while extensions to the noncoherent ACR/APR channels are investigated in Section 5.5. The conclusions are summarized in Section 5.6.

5.2 Theoretical Background

A rotation of a symbol $x \in \mathbb{C}$, where $\mathbb{C}$ denotes the set of complex numbers, by an angle θ is defined as $\rho_\theta(x) = xe^{j\theta}$. Rotation of a sequence of symbols in $\mathbb{C}^N$ is

defined as component-wise rotation, *i.e.*,

$$\rho_\theta((x_1, \ldots, x_N)) = (\rho_\theta(x_1), \ldots, \rho_\theta(x_N)), \tag{5.1}$$

and rotation of a set is defined as

$$\rho_\theta(S) = \{\rho_\theta(x) \mid x \in S\}. \tag{5.2}$$

A signal constellation S has an n-fold symmetry if it has $2\pi/n$ rotational invariance, *i.e.*, $\rho_{2\pi/n}(S) = S$ (for the highest such n, $2\pi/n$ is called the base angle). For example, the M-PSK signal set has an M-fold symmetry. In the following, ρ will denote the rotation by the base angle, with the latter understood from the content. Let S be a signal set with an n-fold rotational invariance, and $\Lambda = \{\lambda_1, \ldots, \lambda_m\}$ be a partition of S, *i.e.*, $\lambda_i \subset S$, $\lambda_i \cap \lambda_j = \emptyset$ if $i \neq j$ and $\bigcup_i \lambda_i = S$. We say that Λ is rotationally invariant if for every $\lambda_i \in \Lambda$, we have $\rho(\lambda_i) \in \Lambda$.

RI codes are codes that are invariant to rotations by the base angle. Specifically, rotation by any multiple of the base angle applied to all codewords leaves the code unchanged. In addition, an RI encoder is defined as the one-to-many mapping from an input sequence to all rotated versions of a codeword in an RI code, as depicted in Fig. 5.1. At this point, we emphasize the distinction between a code and an encoder. The code is the set of all possible codewords (*i.e.*, the codebook), whereas the encoder is the mapping of input sequences to codewords. Different encoders might produce the same code, and one of the issues dealt with in [85] is a systematic way of finding an RI encoder for a given RI code. The reason behind the one-to-many property of the RI encoder is that for non-catastrophic behavior of RI codes in the DCR channel, the mapping corresponding to the inverse of the encoder should associate all rotated versions of a codeword with the same input sequence. In the case of trellis codes, since the output sequence depends on the input sequence and the initial state, this

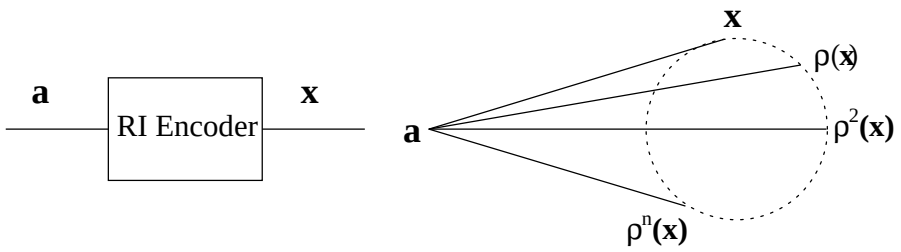


Figure 5.1: Structure of Rotationally Invariant Encoder.

one-to-many behavior is achieved by using the same input sequence with different initial states.

The design of RI coding schemes first starts with finding an RI code and then finding a suitable encoder. For a TCM code defined using a partition[3] Λ, a necessary condition for rotational invariance is that Λ is RI [85].

Necessary and sufficient conditions for a code and an encoder to be RI were given[4] in [85] in terms of the Fischer cover of the code and encoder, respectively. The Fischer cover of a code is defined as the unique state transition graph of the code that satisfies two additional properties: 1) all edges leaving a node have different labels and 2) the set of codewords initiated at any two states differ.[5]

In this chapter we will only use the sufficiency part of the results of [85]. These sufficient conditions apply to any state transition graph, and not only the Fischer cover. Denoting by t a transition from state s_1 to state s_2 associated with input a and output c, the sufficient conditions for a code and its encoder to be RI are summarized as follows

---

[3]Original design rules for constructing TCM codes in [86] use signal set partitioning and these rules are followed in almost all works dealing with TCM code design, including [4, 20].

[4]The sufficient condition was first introduced in [89, 90, 91].

[5]Since the encoder can be thought of as a code whose graph has output labels the input-output pairs (a, c), the above definition of the Fischer cover can be applied to the encoder as well.

1. The graph obtained by applying a rotation by the base angle to the output labels of the transitions of the state transition graph of the encoder, is isomorphic to the original graph, *i.e.*, there exists a bijective mapping (also denoted[6] by ρ) between states and transitions of the two graphs, which maps the transition t from state s_1 to state s_2 with output label c to a transition $\rho(t)$ from state $\rho(s_1)$ to state $\rho(s_2)$ with output label $\rho(c)$.

2. Every pair of transitions $(t, \rho(t))$ is associated with the same input label a.

If the phrase "state transition graph" is replaced by "Fischer cover" in the above statements, then these become necessary conditions for a code and an encoder to be RI, as well.

If the state transition graph of a code possesses only the first property above, then the corresponding code is still RI, but the encoder is not.[7] All the states of the state transition graph of an RI encoder are partitioned into *orbits* of size n, where the orbit of a state s is defined as the set $\{\rho^i(s)|i = 0, 1 \ldots n-1\}$. Therefore, the number of states of an RI encoder should be a multiple of n. Starting from the two conditions stated above, one can derive a systematic procedure to find an RI encoder for a given RI code, as explained in [85].

Finally, we point out that an important property of RI encoders is that they cannot be feedforward [85]. This is a favorable property for RI-SCTCM code design, since for good SCTCM codes the inner code *must* be recursive.

---

[6]Following the notation used in [85], we use the same symbol ρ to denote this bijective mapping of states and edges. The same symbol is also used to denote rotations of signals, and sets, as in (5.1), (5.2), respectively.

[7]In [75], a code is called rotationally transparent if both the code and the encoder are RI, and rotationally invariant if the code is RI only.

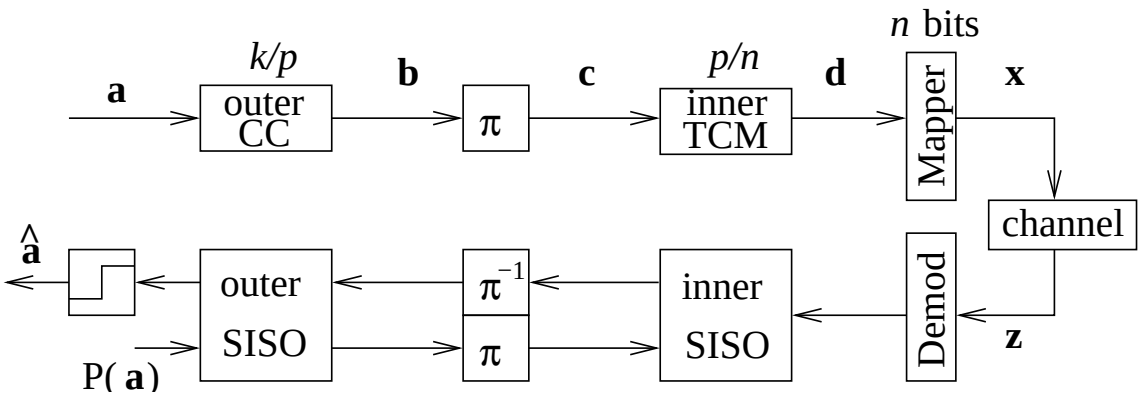


Figure 5.2: SCTCM encoder and iterative decoder structure

5.3 Design of RI-SCTCM codes

5.3.1 Design Guidelines

Serially concatenated trellis-coded modulation schemes were introduced in [4] as alternatives to the well known "turbo" coding schemes [9], which are parallel concatenated codes. A typical SCTCM, shown in Fig. 5.2, consists of a rate k/p outer convolutional code (CC), the output of which is fed into an inner TCM code of rate p/n, after symbol-wise or bit-wise interleaving (denoted by π). These n bits are then mapped onto a transmitted symbol (complex number), resulting in an overall throughput of k bits/symbol. For a length L interleaver the overall code can be thought of as an equivalent block code which takes Lk bits and outputs L symbols. The number of states in the equivalent code is large and thus the complexity of the maximum likelihood sequence detection (MLSD) decoder is very high. However, a practical sub-optimal decoder for such a code is the iterative decoder [4], which consists of two Soft-Input Soft-Output (SISO) modules [6] and an interleaver/deinterleaver pair, as shown in Fig. 5.2. Each SISO module has complexity proportional to the number of states of the corresponding constituent code (for a detailed discussion of SCTCM codes see Section C.1 of Appendix C.)

One might aim to find necessary and sufficient conditions for the SCTCM coding scheme to be RI. However, we will not follow this approach. Instead, we are interested in sufficient conditions that depend only on the constituent codes and not the interleaver. In [49] it was shown that a simple sufficient condition for the SCTCM scheme to be RI is that its inner TCM *encoder* is RI. Indeed, if an input sequence $\mathbf{a}$ encodes into $\mathbf{x}$ via the outer-code codeword $\mathbf{b}$, then $\mathbf{a}$ also encodes to $\rho(\mathbf{x})$ via the same outer codeword, the difference being the starting state of the inner encoder. However, observe that, if the inner *code* is RI, and the utilized *encoder* is not, then the overall code is not necessarily RI. In this case, although the rotated version of any codeword (of the inner code) is still a codeword, their corresponding input sequences differ, and since these can only be (interleaved) codewords of the outer code, the rotated codeword need not be a codeword of the concatenated code.

It should be emphasized that RI-SCTCM code design does not reduce to the design of good inner RI-TCM codes, *i.e.*, inner RI-TCM codes with good distance properties. The reason is that, the assignment of input sequences to output codewords (which is irrelevant for standard RI-TCM as long as this assignment satisfies property 2 of Section 5.2) is crucial for the case of RI-SCTCM, since the overall *code* depends on the particular inner *encoder*. As a result, RI-SCTCM code design should incorporate the rotational invariance property to the design rules of [4].

The design procedure for RI-SCTCM codes is similar to non-RI SCTCM codes. The outer code is selected to maximize the outer code free distance. Regarding the inner code, because of the constraints placed on the encoder by the RI property, it is only required to specify the labels for transitions leaving only one state in each orbit. Furthermore, in performing the assignment of input bits, the SCTCM design parameter $d^2_{f,\text{eff}}$ [4] is maximized. This parameter is defined as the minimum squared

Euclidean distance between inner-code input sequences that are Hamming distance 2 apart, *i.e.*,

$$d^2_{f,\text{eff}} = \min_{d^H(\mathbf{c},\mathbf{c}')=2} d^2(\mathbf{x}(\mathbf{c}), \mathbf{x}(\mathbf{c}')). \tag{5.3}$$

This procedure is followed in the next section, where specific examples of high-rate RI-SCTCM codes are given. We point out that these designs are the first powerful RI-SCTCM codes that appear in the literature.

Finally, it is mentioned that the iterative decoder for an RI-SCTCM code can be the standard coherent iterative decoder, except that for the inner SISO all the states in the orbit of the actual initial state are initialized to be equally likely. If the initial state is unknown to the receiver, then the starting state is initialized to be equally probable between all the available states [49].

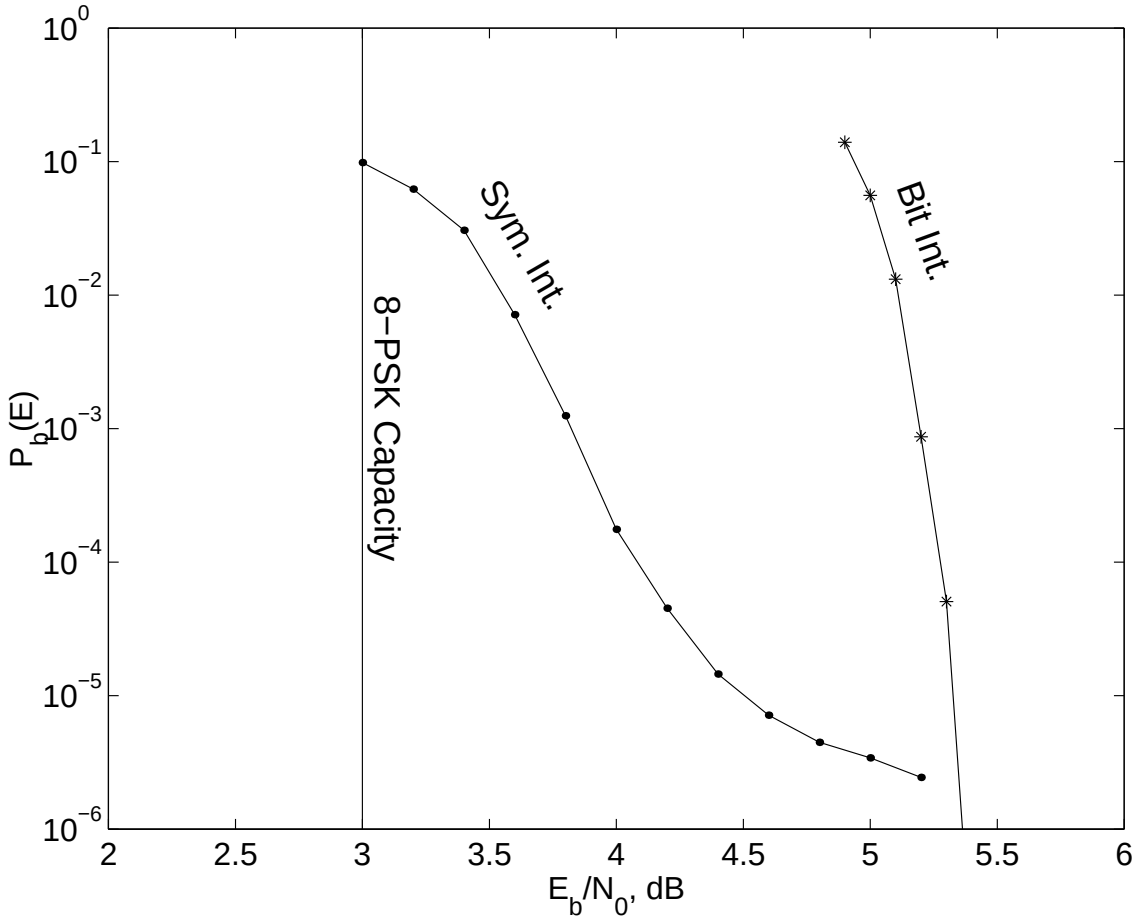
5.3.2 Design Examples

The first code (named Code 1) is an 8-PSK code of throughput 2 bits/symbol, constructed by a rate 2/3 4-state outer code and a rate 3/3 inner code. The outer code is obtained by using twice a rate 1/2 systematic code with coded bit polynomial $\frac{1+D}{1+D+D^2}$, and puncturing the fourth bit [20]. Since the codebook of the inner code is the set of all 8-PSK sequences, it is straightforward from the previous discussion that exactly 8 states are required for the RI inner encoder. Also, since $n = 8$, we have one orbit with 8-states which are numbered 0 to 7 such that $\rho(i) = (i+1) \bmod 8$. All the states are connected to each other, and all the transitions going to state i have the output label i, which corresponds to the 8-PSK symbol $\exp(j2\pi i/8)$. What remains to do is to assign input labels. It suffices to assign these to all the transitions leaving state 0, which will then induce the labels on the rest of transitions via the rotational invariance property. The table that describes the input output relationship of the

inner code at state 0 is given in Fig. 5.3 together with the state transition graph induced by the RI property. In this table, integers in the "Input" column are the decimal representation of the three-bit output of the inner code (this convention is followed throughout this section).

The performance of this code is depicted in Fig. 5.3. In the same figure, the capacity limit of the 8-PSK-constrained channel is plotted as a vertical line for comparison. Note that for this code symbol-wise interleaving results in approximately 1 dB loss from the modulation constrained capacity for BER of 10^{-4}, and symbol-wise interleaving outperforms bit-wise interleaving. As it will be demonstrated later, this result is not always true.

The last two codes are over the 16-QAM constellation. The first code, named Code 2, uses a rate 2/3 outer and a rate 3/4 inner code, resulting in a throughput of 2 bits/symbol. This code is the RI version of SCTCM code published in [20], obtained by making the inner code RI. Consequently, the outer code, whose generator matrix is provided in Fig. 5.4, is unchanged. The rotational symmetry group of 16-QAM consists of 4 rotations, and the inner code obtained has 4 states. The transitions leaving zero-state are given in Fig. 5.4 for the inner code. Output labels in this transition table correspond to the following 16-QAM labeling: the number $i = 4q + r$ (with q and r being integers between 0 and 3 inclusive) is mapped onto the 16-QAM point $((-3 + 2r) + j(3 - 2q))/\sqrt{10}$. The performance of Code 2 is depicted in Fig. 5.4, where it can be observed that symbol-wise interleaving outperforms bit-wise interleaving for this code, as well. The dashed curves correspond to the performance of the non-RI SCTCM code of [20] (named Code 2.1) with the same outer code and a 2-state inner code. It can be seen that despite the lower complexity, the non-RI code outperforms the RI code (at least for bit interleaving).



Input	Output	Next State
0	0	0
1	7	7
2	5	5
3	2	2
4	3	3
5	4	4
6	6	6
7	1	1

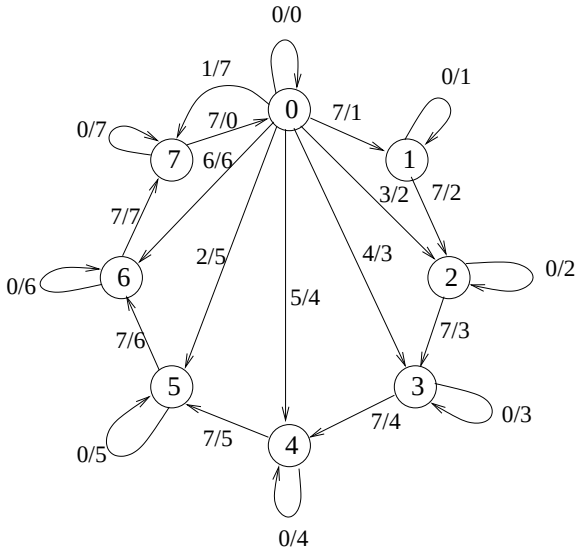
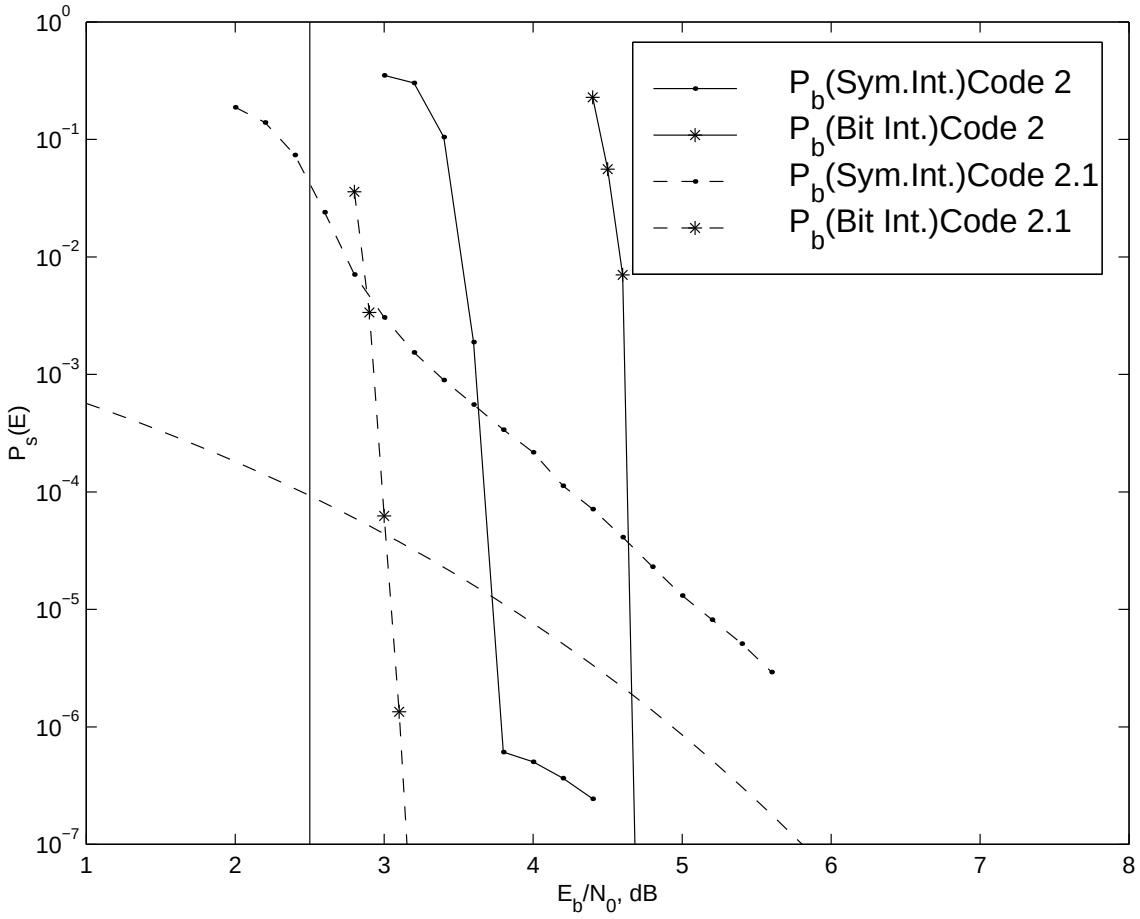


Figure 5.3: Code 1. Throughput 2 bits/symbol 8-PSK code with rate 2/3 outer and rate 3/3 inner code. Input block of $Lk = 16384$ bits is used. Inner code Description: transition structure at state 0, and partial state transition diagram.

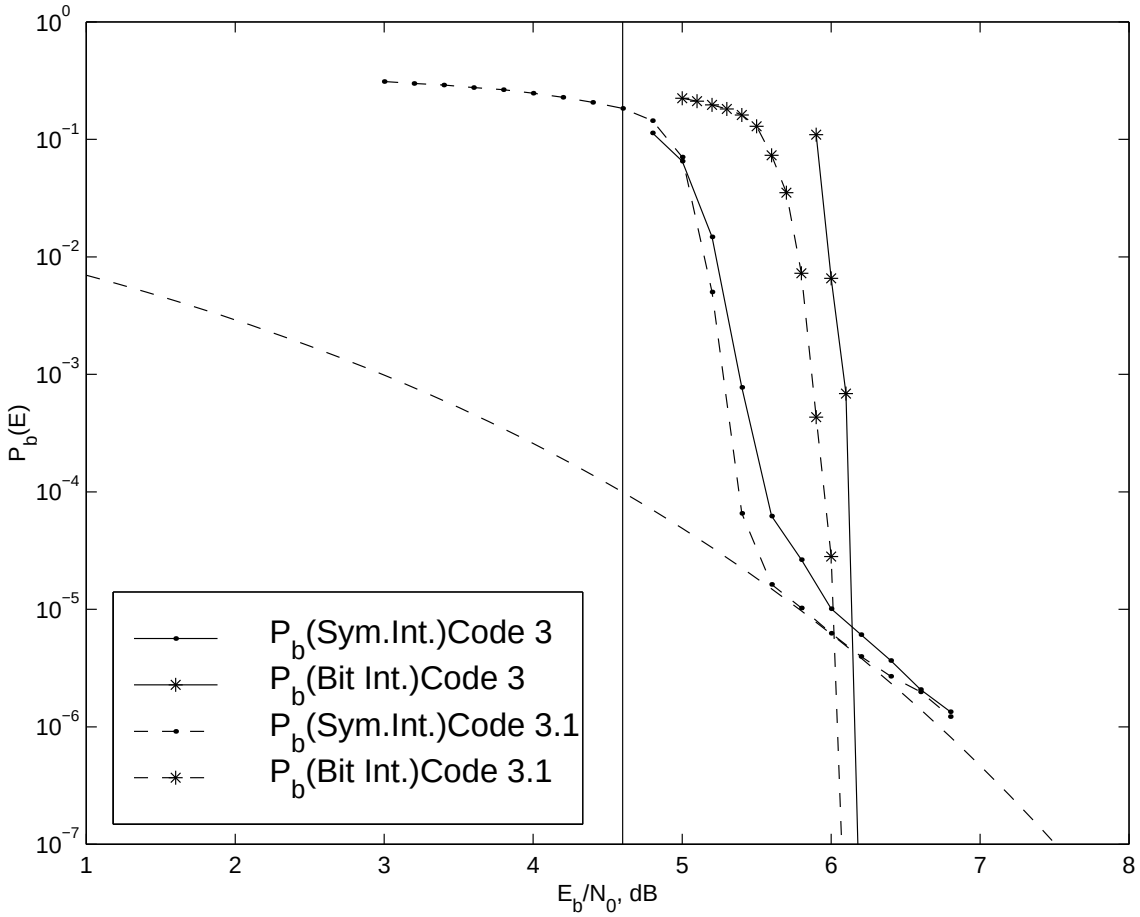


$$G_2(D) = \begin{bmatrix} 1 & 1+D & f(D) \\ f(D) & f(D) & 1 \end{bmatrix}$$

$$f(D) = 1 + D + D^2$$

Inp	0	1	2	3	4	5	6	7
Out	0	2	8	10	15	7	13	5
N.S.	0	2	2	0	3	1	1	3

Figure 5.4: Codes 2 and 2.1. Throughput 2 bits/symbol 16-QAM codes with rate 2/3 outer and 3/4 inner code. Input block of $Lk = 16384$ input bits is used. The lower dashed curve is a lower bound, derived in Appendix D, on the symbol-wise interleaved BER for Code 2.1.



$$G_3(D) = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & D & D & 1+D \\ 0 & 0 & 1 & 1+D \end{bmatrix}$$

Inp	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Out	0	3	1	2	15	12	14	13	7	4	6	5	8	11	9	10
N.S.	0	3	3	0	2	1	1	2	2	1	1	2	0	3	3	0

Figure 5.5: Codes 3 and 3.1. Rate 3 16-QAM code with rate 3/4 outer and rate 1 inner. Input block of $Lk = 16383$ input bits is used. The lower dashed curve is a lower bound, derived in Appendix D, on the symbol-wise interleaved BER for both codes.

The next 16-QAM code has an overall throughput of 3 bits/symbol. This code (named Code 3) consists of a rate 3/4 outer and a rate 1 inner code. The outer code is the best rate 3/4 convolutional code [20] with $d_{\text{free}} = 3$, the generator matrix of which is provided in Fig. 5.5. The inner code transition table at state 0 is also given in the same figure. The performance of this code is shown in Fig. 5.5 together with a non-RI code of [20] (named Code 3.1) of exactly the same complexity. It can be observed that, contrary to the previous case of Code 2 and 2.1, the penalty paid for achieving rotational invariance is not significant for Code 3.

As a general observation we can say that while for some codes rotational invariance is achieved at significant performance loss (Code 2 vs. Code 2.1), others show no considerable performance degradation (Code 3 vs. Code 3.1). This behavior is observed for non-concatenated RI-TCM codes as well [7], and a sufficient explanation is still not available.

5.4 Rotationally robust codes for the DPR channel

We now investigate codes for the DPR channel, that allows for different discrete rotations to affect different parts of the codeword, as opposed to the DCR channel, for which the same rotation affects the whole codeword. As mentioned in the Section 5.1, the DPR channel is sufficient for modeling cycle-slips occurring within the transmitted codeword. In particular, it is assumed that the discrete random rotation could change value in a random position k within a codeword, to some other random multiple of the base angle (see Table 5.1).

Although a partially rotated codeword is not necessarily a codeword of an RI code, it is desirable that the coherent MLSD receiver decodes a partially rotated codeword into an input sequence which differs from the original one in finitely many symbols,

	Constant	Partial
Discrete	DCR $z_i = x_i e^{j\phi m} + n_i$	DPR $z_i = \begin{cases} x_i e^{j\phi m} + n_i & i < k \\ x_i e^{j\phi m'} + n_i & i \geq k \end{cases}$
Arbitrary	ACR $z_i = x_i e^{j\theta} + n_i$	APR $z_i = \begin{cases} x_i e^{j\theta} + n_i & i < k \\ x_i e^{j\theta'} + n_i & i \geq k \end{cases}$

x_i, z_i represent the ith transmitted and observed symbol, respectively
n_i's are zero-mean complex Gaussian with variance σ^2
m and m' are independent random variables uniform in $\{0, 1, \ldots n-1\}$
θ and θ' are independent random variables uniform in $[-\pi, \pi)$
k is a random variable uniform in $\{0, 1, \ldots N-1\}$
$\phi = 2\pi/n$ is the base angle of the signal constellation used.

Table 5.1: Summary of Considered Channels

i.e., the errors are concentrated around the time index k and do not increase with the blocklength. This property, which is obviously stronger than rotational invariance, is called rotational robustness, and the corresponding codes are referred to as RR codes. We point out that, in general, rotational robustness is a property of both the code and the particular receiver used. It is true that standard (*i.e.*, non-concatenated) RI-TCM schemes with coherent MLSD decoding are also RR. This statement is made precise with the following theorem, which considers only a single partial rotation. The extension to arbitrary number of rotation changes is straightforward.

Before stating the theorem we introduce two parameters associated with a given encoder. First, the parameter K is the smallest number of steps required to reach any state from any other state on a trellis. It can be shown that for irreducible graphs (*i.e.*, graphs for which any two states are connected through a series of transitions), K is at most $S-1$, where S is the number of states in the graph. In the theorem below we also assume that we can reach any state in *exactly* K steps, which is true

for all RI encoders that we considered. For codes whose graphs are not irreducible we take K to be infinite.

The second parameter, L, is the maximum number of *consecutive* output symbols that agree for two paths in a trellis that do not share a common state. Also, let $\mathbf{c} = f(s, \mathbf{a})$ represent the codeword for a given input sequence $\mathbf{a}$ and starting state s. If $f(\cdot, \mathbf{a})$ is one-to-one[8] (viewed as a function of s, for a given input sequence $\mathbf{a}$), then the parameter L is infinite if and only if the code is catastrophic. Furthermore, for non-catastrophic codes, the parameter L can be bounded as $L \leq S(S-1)/2 - 1$. For binary linear codes, a tighter upper bound can be found as $L \leq \log_2 S$.

Theorem 5.1. *Let an input sequence* $\mathbf{a}$ *be encoded by an RI encoder into the codeword* $\mathbf{x}$*, and let* $\mathbf{x}'$ *be a partially rotated version of* $\mathbf{x}$ *so that all output symbols are rotated starting from some index* k*,* i.e.*,*

$$x'_i = \begin{cases} x_i & , i < k \\ \rho(x_i) & , i \geq k \end{cases} \tag{5.4}$$

If the minimal encoder for the RI code has finite parameters K *and* L*, then the input sequence* $\hat{\mathbf{a}}$ *corresponding to the closest (in Euclidean distance) codeword* $\hat{\mathbf{x}}$ *to* $\mathbf{x}'$ *satisfies*

$$d^H(\mathbf{a}, \hat{\mathbf{a}}) \leq (L+1) \times (K \times \frac{d_M^2}{d_m^2} + 2) + 1, \tag{5.5}$$

where d^H *denotes the symbol-wise Hamming distance, and* d_M^2 *and* d_m^2 *are the maximum and minimum Euclidean intra square distances within the output alphabet (signal constellation).*

Proof. Consider a trellis of the RI *code* and the path corresponding to the partially rotated sequence $\mathbf{x}'$, depicted in Fig. 5.6. Since the code is not catastrophic, we

[8]This, rather natural, condition is assumed throughout the chapter.

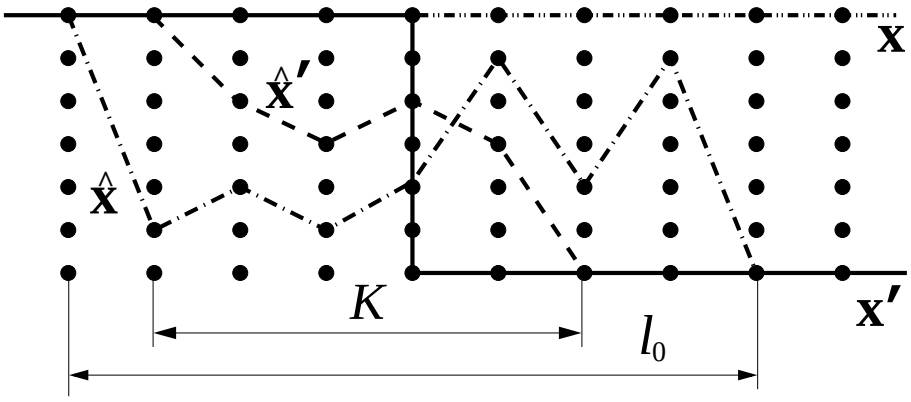


Figure 5.6: Trellis of RI code; $\hat{\mathbf{x}}$ denotes the codeword chosen by MLSD.

know that the path corresponding to $\hat{\mathbf{x}}$ should agree with the path of $\mathbf{x}'$ in all but l_0 transitions around time k.[9] We will first bound l_0. Since there exists a valid codeword $\hat{\mathbf{x}}'$ that agrees with $\mathbf{x}'$ in all but at most K symbols (trellis steps), the Euclidean squared distance is bounded by

$$d^2(\mathbf{x}', \hat{\mathbf{x}}') \leq K \times d_M^2. \tag{5.6}$$

On the other hand, since $\hat{\mathbf{x}}$ is the closest codeword to $\mathbf{x}'$ we have

$$d^2(\mathbf{x}', \hat{\mathbf{x}}') \geq d^2(\mathbf{x}', \hat{\mathbf{x}}) \geq l_1 \times d_m^2, \tag{5.7}$$

where l_1 is the number of output symbol disagreements between $\mathbf{x}'$ and $\hat{\mathbf{x}}$. Since the encoder has finite L parameter, given l_1 output symbol disagreements, l_0 is bounded by

$$l_0 \leq (L+1) \times (l_1 + 2) \tag{5.8}$$

(corresponding to the worst possible case of having each differing output symbol be

[9]The path followed by $\mathbf{x}'$ is not a valid path through the trellis since it involves a jump at time k

followed by L agreeing symbols[10]). Combining the above three inequalities we get

$$l_0 \leq (L+1) \times (K \times \frac{d_M^2}{d_m^2} + 2). \tag{5.9}$$

We note that in the above equations we use the parameters K and L of the minimal encoder rather than those of the RI encoder. This is valid since for an RI encoder the number of differing trellis steps for two given codewords is at most 1 more than the number l_0 corresponding to the minimal encoder used to generate the RI code following the procedure of [85]. This is true since, if a state is split into several states, all new states have the same outgoing transitions as the original states (see [85] for details on how the RI encoder is obtained given the RI code). The proof is concluded by using the fact that the number of input symbol disagreements $d^H(\mathbf{a}, \hat{\mathbf{a}})$ can be at most $l_0 + 1$. ■

Although the above theorem does not take into account the channel noise, there is a straightforward way to extend this result for the case of AWGN channel. We need to find a new upper bound for l_1 that takes into account the effects of the noise. The rest of the proof is identical to the above. For that purpose, assume operating signal-to-noise ratio of $\gamma = E_s/N_0$. The key idea in the AWGN case is that the decision metrics are Gaussian random variables. In this scenario, the condition that the metric corresponding to $\hat{\mathbf{x}}$ is smaller than the one corresponding to the codeword $\hat{\mathbf{x}}'$ can be written as

$$m(\hat{\mathbf{x}}) - m(\hat{\mathbf{x}}') = \sqrt{\gamma}(l_1 d_m^2 - K d_M^2) - n\sqrt{l_1 d_m^2 + K d_M^2} < 0, \tag{5.10}$$

where n is a zero-mean real Gaussian random variable with variance $1/2$. Two things are assumed in this equation: (i) the positions of l_1 disagreeing symbols between $\hat{\mathbf{x}}$

[10]Since $\mathbf{x}'$ is a concatenation of two valid partial paths, around time k there can be $2L$ consecutive agreeing symbols

and $\mathbf{x}'$ do not overlap with the K disagreeing positions between $\hat{\mathbf{x}}$ and $\mathbf{x}'$; (ii) distances within the mentioned l_1 and K disagreements are $E_s d_m^2$ and $E_s d_M^2$, respectively.[11] It is easy to see that these conditions indeed comprise the worst-case scenario (for the purposes of maximizing l_1). For any positive A, if $n < \sqrt{A}$, then by solving for l_1 from (5.10) we get that

$$l_1 \leq K \times \frac{d_M^2}{d_m^2} + \frac{A + \sqrt{8Kd_M^2 A\gamma + A^2}}{2d_m^2\gamma}. \tag{5.11}$$

Since this event has probability

$$P(n < \sqrt{A}) = 1 - \frac{1}{2}\operatorname{erfc}(A) \approx 1 - \frac{1}{2}\exp(-A), \tag{5.12}$$

we finally get that

$$d^H(\mathbf{a}, \hat{\mathbf{a}}) \leq (L+1) \times (K \times \frac{d_M^2}{d_m^2} + \frac{A + \sqrt{8Kd_M^2 A\gamma + A^2}}{2d_m^2\gamma} + 2) + 1. \tag{5.13}$$

with probability $1 - 0.5\exp(-A)$, for any positive A. Observe that the difference between the bounds for the noiseless and the noisy case is an additional term which is small for the parameter values of interest.

We emphasize that the above theorem is valid for all RI-TCM codes, *i.e.*, for standard non-concatenated RI-TCM schemes, as well as RI-SCTCM schemes. In particular, for good non-concatenated RI-TCM codes the parameter L is small.[12] Also, for codes with large input alphabet (compared to the number of states), and no parallel transitions, the parameter K is small. For instance all inner RI-TCM codes that we used have parameters $K = 1$ and $L \leq 1$. The ratio of squared

---

[11] Observe that d_m^2 and d_M^2 denote the normalized (*i.e.*, taking $E_s = 1$) minimum and maximum square distances within the constellation

[12] The parameter L was calculated for all good codes presented in [72, pp. 493-496] and found to be on the order of $(\log_2 S)/(n-k)$, for an (n, k) convolutional code.

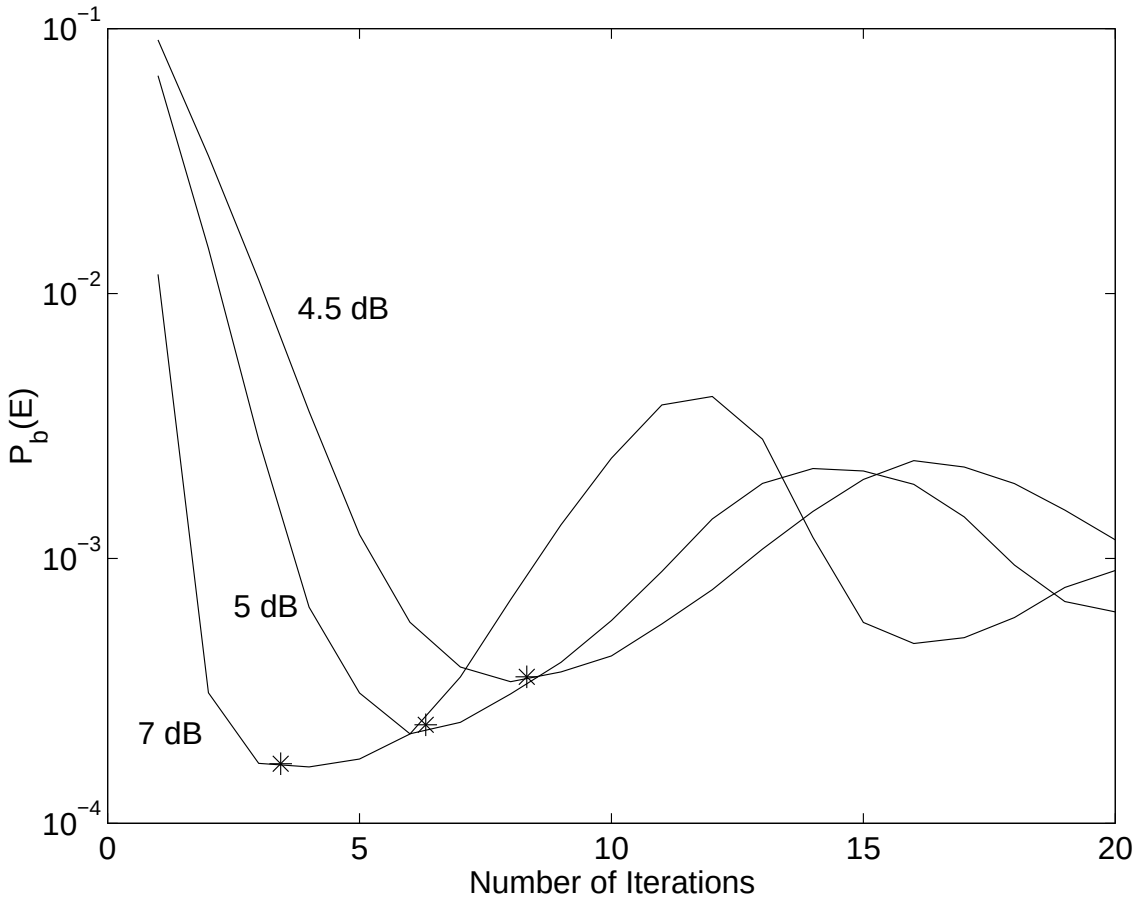


Figure 5.7: BER vs. number of iterations for Code 1 with partial rotations ($E_b/N_0 =$ 4.5, 5, and 7 dB). Stars on the curves correspond to the BER reached by the entropy minimizing stopping criterion plotted at the average stopping iteration number

distances that appears in the theorem is the only factor that relates to the output distance, since the rotational invariance is a property of the state transition graph. This ratio is 4/0.58=6.82 for 8-PSK, and 12.8/0.4=32 for 16-QAM constellation.

Although the theorem is valid for all RI codes, it does not provide any information about the robustness of RI-SCTCM codes. This is because, due to the presence of the interleaver, an SCTCM code is an equivalent TCM code with a huge number of states (and possibly with very large parameters K and/or L). On the other hand, observe that a better bound on the parameters K and L, that is connected to the

code performance rather than the number of states utilized, would make the above theorem useful for the SCTCM case as well.

In fact, simulation results have revealed interesting facts. In Fig. 5.7 the performance of the standard iterative receiver for Code 1 (with block length of 16384 bits) is shown as a function of the iteration number. In this simulation experiment, a single partial rotation of discrete but random magnitude is introduced in a random position inside each transmitted block. It can be observed that the iterative receiver does not converge to a final decision at least for the first 20 iterations. Furthermore, the lowest BER is achieved at a different iteration number for different E_b/N_0 values. This lowest achievable BER for $E_b/N_0 = 7$dB is on the order of 10^{-4} which suggests that a BER floor is reached.

The above simulated performance does not necessarily imply that the MLSD receiver is not robust to partial rotations for an RI-SCTCM scheme. In fact, the BER floor corresponds to an average of 2 bit errors per block, which strongly suggests that a useful upper bound exists for RI-SCTCM codes as well. On the other hand, the oscillatory behavior of the iterative receiver demonstrates its weakness in handling partial rotations. A possible solution is to modify the iterative receiver by adding some stopping criterion in the iterative loop, in order to detect the minimum BER point.

One such stopping criterion is investigated herein. Namely, at each iteration, the average entropy of the information bits is calculated, and its minimum is tracked. The iteration number at which this average entropy started to increase is chosen as the stopping criterion for iterations. Simulation results show that this simple filtering operation is sufficient to track the minimum BER point of the receiver quite successfully. These results are shown on the graph of Fig. 5.7 as "stars", denoting

the minimum BER reached by using this stopping criterion. A similar behavior was observed for several other codes, including the ones presented in Section 5.3.2. An intuitive explanation for its effectiveness is that such a stopping criterion is good for any iterative scheme that approximates aposteriori probabilities, since a small entropy implies a large probability. In this case, the additional feature is that the approximate aposteriori probability was observed to oscillate, so it is reasonable to stop at its first minimum.

Finally, it is worth mentioning that while Theorem 5.1 provides a generic bound for codes with given parameters, a better bound can be found for specific codes. For example, for the inner code of Code 1, the number of input symbol errors as a result of partial rotation is 1, whereas the upper bound suggested by the theorem is 9 (using $K = 1$ and $L = 0$).

5.5 Rotationally robust codes for the ACR/APR channels

In this section, we investigate noncoherent channels that introduce arbitrary rotations to the transmitted codewords. These phase changes either remain constant over the entire codeword, or affect only part of the codeword. The discussion is initiated with the ACR channel, *i.e.*, the channel which assumes rotations are arbitrary in $(-\pi, \pi]$, but otherwise constant over the whole codeword (see Table 5.1). As was noted in the Introduction, the only possible choices for noncatastrophic behavior over this channel are RI codes and anti-RI codes [75]. The obvious question is which of these two choices is preferable in the ACR channel.

It is now shown that as far as their performance on the ACR channel is concerned RI and anti-RI codes are equivalent. First, observe that there is a one-to-one mapping between RI and anti-RI codes. In particular, let $\mathcal{C}$ denote an anti-RI code over the

signal alphabet that has n-fold rotational symmetry, then by adding all the rotated versions of the codewords in $\mathcal{C}$ we get an RI code $\mathcal{C}_{\text{RI}} = \{\rho^m(\mathbf{x}) \mid \mathbf{x} \in \mathcal{C}, \quad m = 0, 1, \ldots n-1\}$. Similarly, starting from any RI code, by excluding all but one rotated versions of a codeword, an anti-RI code is generated. Moreover, if $\mathbf{x}(\mathbf{a})$ denotes the encoder mapping of the anti-RI code, the corresponding RI encoder transmits one of the rotated versions of $\mathbf{x}(\mathbf{a})$, $\rho^m(\mathbf{x}(\mathbf{a}))$ with equal probability,[13] when the input sequence is $\mathbf{a}$. The MLSD decoding rule for transmission over the ACR channel is exactly the same for both cases, since

$$\begin{aligned} \hat{\mathbf{a}}_{\text{RI}} &= \arg\max_{\mathbf{a}} p_{\mathbf{z}|\mathbf{a}}(\mathbf{z}|\mathbf{a}) = \arg\max_{\mathbf{a}} \frac{1}{n} \sum_{m=0}^{n-1} p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\rho^m(\mathbf{x}(\mathbf{a}))) \\ &= \arg\max_{\mathbf{a}} p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}(\mathbf{a})) = \hat{\mathbf{a}}_{\text{anti-RI}}, \end{aligned} \tag{5.14}$$

where we used the fact that for the noncoherent ACR channel $p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}) = p_{\mathbf{z}|\mathbf{x}}(\mathbf{z}|\mathbf{x}e^{j\theta})$ for any angle θ. Therefore, RI and anti-RI codes are equivalent for this channel model, which raises the question whether any of these classes of codes is preferable. A partial answer to this question is provided here by showing that non-concatenated RI codes are provably robust to partial arbitrary rotations, *i.e.*, are provably robust in the APR channel. The APR model is similar to the DPR, except now the rotations can be arbitrary (see Table 5.1). The reason for the interest in the APR channel, is that it sufficiently models noncoherent channels with fast dynamics where phase changes are often. It is emphasized that rotational robustness in this case is discussed in connection with the optimal MLSD decoder for the noncoherent ACR channel, since the MLSD decoder matched to the AWGN channel is completely useless in this scenario.

Let $\mathcal{C}$ be an RI code, $\phi = 2\pi/n$ the base angle of the RI signal set, θ an arbitrary

[13] This is achieved through the choice of the initial state of the RI encoder.

angle in $(-\pi, \pi]$, $\mathbf{x}$ the transmitted codeword corresponding to the input sequence $\mathbf{a}$, and $\mathbf{z}$ the partially rotated codeword at a certain time index k, *i.e.*,

$$z_i = \begin{cases} x_i & , i < k \\ x_i e^{j\theta} & , i \geq k. \end{cases} \tag{5.15}$$

In addition, let the utilized decoder be the standard MLSD receiver for the noncoherent channel (assuming all codewords $\mathbf{x}$ have equal energy), given by

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathcal{C}} \Lambda(\mathbf{x}, \mathbf{z}) = \arg\max_{\mathbf{x}\in\mathcal{C}} |\mathbf{x}^H \mathbf{z}| = \arg\max_{\mathbf{x}\in\mathcal{C}} \Big| \sum_i z_i x_i^* \Big|, \tag{5.16}$$

with $\mathbf{z}$ being the observed sequence. Associated with the decoded sequence $\hat{\mathbf{x}}$, through the inverse mapping of the encoder, is the corresponding input sequence $\hat{\mathbf{a}}$. In the following, for simplicity we assume M-PSK signals ($n = M$ in this case). The extension to the case of signal sets that utilize amplitude variations is straightforward. The following lemma proves that for RI codes, no decoding error occurs in the APR channel in the special case when the partial phase rotation is smaller than half the base angle. This result is then generalized by dropping this assumption.

Lemma 5.2. *If $|\theta| < \phi/2$ then the codeword chosen by the noncoherent MLSD decoder in* (5.16) *is $\hat{\mathbf{x}} = \mathbf{x}$.*

Proof. Let $\mathbf{y}$ be any sequence whose elements are from the constellation used for transmission, and let $A_l = k$ and $A_r = N - k$ be the number of symbols on the left and the right of the rotation boundary k. We have

$$\Lambda(\mathbf{y}, \mathbf{z}) = |\mathbf{y}^H \mathbf{z}| = \Big| \sum_{i=0}^{k-1} y_i^* x_i + e^{j\theta} \sum_{i=k}^{N-1} y_i^* x_i \Big| = \Big| \sum_{i=0}^{k-1} e^{jm_i\phi} + e^{j\theta} \sum_{i=k}^{N-1} e^{jm_i\phi} \Big|, \tag{5.17}$$

where we have used the fact that for M-PSK signals x_i, y_i, there is an $m_i \in \{0, \ldots, n-1\}$ such that $y_i^* x_i = e^{jm_i\phi}$. Collecting all terms with the same $m_i = m$ we get

$$\Lambda(\mathbf{y}, \mathbf{z}) \;\; = \;\; \Big| \sum_{m=0}^{n-1} (a_m + b_m e^{j\theta}) e^{jm\phi} \Big| = \Big| \sum_{m=0}^{n-1} d_m e^{jm\phi} \Big|, \tag{5.18}$$

where a_m is the number of terms in the left sum of (5.17) with $m_i = m$, and similarly for b_m (clearly, $\sum_{m=0}^{n-1} a_m = A_l$ and $\sum_{m=0}^{n-1} b_m = A_r$). Observe that $|\arg(d_m)| \leq |\theta| < \phi/2$, and also $|\arg(d_m d_{m'}^*)| \leq |\theta| < \phi/2$. Hence we have

$$\begin{aligned}\Lambda(\mathbf{y},\mathbf{z})^2 &= \Big|\sum_{m=0}^{n-1} d_m e^{j\phi m}\Big|^2 = \sum_{m=0}^{n-1} |d_m|^2 + 2\sum_{\substack{m,m'=0\\ m>m'}}^{n-1} |d_m d_{m'}| \cos((m-m')\phi + \alpha_{m,m'}) \\ &< \sum_{m=0}^{n-1} |d_m|^2 + 2\sum_{\substack{m,m'=0\\ m>m'}}^{n-1} |d_m d_{m'}| \cos(\alpha_{m,m'}) \\ &= \Big|\sum_{m=0}^{n-1} d_m\Big|^2 = |A_l + A_r e^{j\theta}|^2, \end{aligned} \tag{5.19}$$

where $\alpha_{m,m'} = \arg(d_m d_{m'}^*)$, and inequality follows from the fact $\cos(t\phi + \alpha_{m,m'}) < \cos(\alpha_{m,m'})$ for any non-zero integer t since $|\alpha_{m,m'}| < \phi/2$. On the other hand, for the transmitted sequence $\mathbf{x}$ we have $\Lambda(\mathbf{x},\mathbf{z}) = |A_l + A_r e^{j\theta}|$. The proof is completed by combining this result with (5.19) to get

$$\Lambda(\mathbf{x},\mathbf{z}) > \Lambda(\mathbf{y},\mathbf{z}) \Rightarrow \hat{\mathbf{x}} = \mathbf{x}. \qquad \blacksquare$$

The more general case when the condition of Lemma 5.2 does not hold, is addressed by the following theorem.

Theorem 5.3. *Under the assumptions of finite K and L parameters, the input sequence $\hat{\mathbf{a}}$, corresponding to the codeword $\hat{\mathbf{x}}$ chosen by the noncoherent MLSD decoder in* (5.16) *satisfies:*

$$d^H(\mathbf{a},\hat{\mathbf{a}}) \leq (L+1)\Big(K\frac{2}{1-\cos\phi}\cdot\frac{1}{1-|\tan(\theta'/2)|\sin\phi} + 2\Big) + 1, \tag{5.20}$$

where θ' is such that $\theta = s\phi + \theta'$, $|\theta'| \leq \phi/2$ and $s \in \{0, 1 \ldots n-1\}$, for any $\theta \in (-\pi, \pi]$,

Proof. To start with, we know that there exists a valid path through the trellis of length at most K steps which departs from the path of $\mathbf{x}$ at time index $k - k_l$ and

joins the path of $\rho^s(\mathbf{x})$ at some index $k + k_r$ ($k_l + k_r = K$). Call the codeword that starts as $\mathbf{x}$ and joins the path of $\rho^s(\mathbf{x})$, passing through this transition path, $\hat{\mathbf{x}}'$. Also, let $\mathbf{x}'$ denote the sequence which coincides with the path of $\mathbf{x}$ to the left, and with the path of $\rho^s(\mathbf{x})$ to the right of k (this sequence is not a valid codeword since it involves a state jump at time index k). Then, using a similar argument as the one used in the proof of Lemma 5.2, we have

$$\Lambda(\hat{\mathbf{x}}', \mathbf{z}) = |(A_l - k_l) + A_m + (A_r - k_r)e^{j\theta'}| \geq |(A_l - 2k_l) + (A_r - 2k_r)e^{j\theta'}|. \quad (5.21)$$

Similarly to Theorem 5.1, we only need to bound $l_1 = d^H(\hat{\mathbf{x}}, \mathbf{x}')$, the number of symbol disagreements between $\hat{\mathbf{x}}$ and $\mathbf{x}'$. Let m_l and m_r be the number of symbol disagreements between these two sequences on the left and on the right of k, respectively ($m_l + m_r = l_1$). Then the problem reduces to maximizing $m_l + m_r$ subject to the condition that

$$\Lambda(\hat{\mathbf{x}}, \mathbf{z}) \geq \Lambda(\hat{\mathbf{x}}', \mathbf{z}), \quad (5.22)$$

with both m_l and m_r nonnegative. This problem can easily be solved by either direct manipulation or by using the method of Lagrange multipliers. We have

$$\Lambda(\hat{\mathbf{x}}, \mathbf{z}) = |(A_l - m_l) + (B_l + B_r e^{j\theta'}) + (A_r - m_r)e^{j\theta'}|, \quad (5.23)$$

where $B_l = \sum_{i=1}^{m_l} \exp(jg_i\phi)$ and $B_r = \sum_{i=1}^{m_r} \exp(jh_i\phi)$ with g_i and h_i being integers between 0 and $n - 1$. With $\theta' > 0$ the maximum occurs[14] when $g_i = 1$ and $h_i = n - 1\ \forall i$. So the constraint in (5.22) becomes

$$|A_l - m_l + m_l e^{j\phi} + (A_r - m_r + m_r e^{-j\phi})e^{j\theta'}| \geq |A_l - 2k_l + (A_r - 2k_r)e^{j\theta'}|, \quad (5.24)$$

---

[14]The implicit assumption made at this point is that, the resulting m_l satisfies $m_l(\cos\theta'(1 - \cos 2\phi) + \sin\theta'(1 + \sin\phi)) < 2A_l \sin\phi \sin\theta'$. When this assumption is not true (*e.g.*, $\theta' \approx 0$) we have to take $h_i = 1\ \forall i$ which yields an even tighter bound.

which after some algebra yields

$$l_1 = m_l + m_r \leq K \frac{2}{1-\cos\phi} \cdot \frac{1}{1 - |\tan(\theta'/2)| \sin\phi} \tag{5.25}$$

where $K = k_l + k_r$ (we also assumed that A_l and A_r are much bigger than k_l and k_r). Finally, using the same line of reasoning as in Theorem 5.1, the desired result is established from the bound in (5.25). ■

Observe that for the case of discrete rotations we have $\theta' = 0$ and hence

$$d^H(\mathbf{a}, \hat{\mathbf{a}}) \leq (L+1)\Big(K \frac{2}{1-\cos\phi}\Big) + 1. \tag{5.26}$$

Also note that, in this case, the term $2/(1-\cos\phi)$ is the only term relating to the output signal set used, which is the counterpart of the term d_M^2/d_m^2 in Theorem 5.1 for coherent decoding. Furthermore, for M-PSK constellation these two terms are *exactly* the same, which implies the identical bound in this case.

We would like to point out that the MLSD receiver for the ACR noncoherent channel shown in (5.16) is too complex for practical implementation. If suboptimal receivers operating on a small observation window are used (*e.g.*, the receivers proposed in [75, 21, 17]), the results of Theorem 5.3 are not necessarily valid. In fact, one can claim that any receiver that operates on a size-N_r observation window is robust in the sense that when a partial rotation occurs, a maximum of approximately $2N_r$ bits will be lost (in the neighborhood of the partial rotation). This means that the fact that RI codes are RR in the APR channel, proven in Theorem 5.3, is true only when the observation window is much larger than the upper bound obtained in the Theorem. This assumption, however, is not unrealistic, since there exist suboptimal receivers (*e.g.*, per-survivor processing (PSP) receivers with phase-locked loop (PLL) phase estimation [74]) that operate on an—effectively—large window size (due to the

autoregressive nature of the PLL), but only search a part of the decision tree. These receivers can be interpreted as approximations of the algorithm in (5.16).

5.6 Conclusion

In this chapter several issues regarding RI codes were investigated. High-rate, powerful RI-SCTCM codes were designed and their performance was gauged through simulations. The use of RI codes in more complicated channel models than the one generally assumed in the literature was also investigated. Several channel models were considered, and certain RI codes were shown to posses the stronger property of rotational robustness, desired for reliable communication over these channels. Two theorems were proven that guarantee that RI codes with small state space posses this property, when used both in coherent and noncoherent channels. It was further demonstrated through an example that RI codes with larger state space, *e.g.*, RI-SCTCM codes, can also be made robust by adding a simple stopping criterion in the iterative decoding algorithm.

CHAPTER VI

SUMMARY AND FUTURE WORK

In this chapter we summarize the main contributions of this thesis, and present avenues for future research.

6.1 Summary

Several issues regarding communication over the phase-noisy AWGN channel have been investigated. A block-independent model for the phase process is assumed, which leads to a block-wise time-invariant memoryless complex vector channel. The dynamics of the phase process are modeled by the number of symbols, N, over which the phase is assumed constant. This parameter, which can be thought of as the channel coherence time, is assumed known both to the transmitter and the receiver.

Two main problems are investigated: (i) design and analysis of practical coding schemes that come close to the information theoretic limit and (ii) evaluation of information capacity and characterization of the structure of the capacity-achieving density.

Due to the similarity of this channel to the coherent AWGN channel, the problem of code design and analysis was initiated in Chapter II with a natural suboptimal

scheme that pairs a phase estimation with subsequent coherent decoding. In particular, the behaviour of the block independent noncoherent AWGN channel was investigated when pilot-symbol-assisted codes are utilized. The role of the pilot-symbol is to facilitate phase estimation and effectively translate the noncoherent channel to a coherent AWGN channel. The BPSK modulation was used in conjunction with the pilot symbols, and this modulation scheme was paired with the LDPC binary codes. Several approximate receivers were proposed, which perform phase estimation either separately from decoding, or jointly as part of an iterative detection/estimation process.

The performance of these coding schemes was analyzed using density evolution, a recently discovered technique for analysis of codes on graphs. Based on these approximate receivers, a simple upper bound to the performance of any iterative joint detection/estimation algorithm was derived. The trade-off, arising as a result of using pilot symbols, between the quality of the phase estimate and the effective SNR was demonstrated for the proposed receivers. Utilizing density evolution as an analysis and optimization tool, the power allocation to the pilot symbol was quantified, and it was shown that a considerable performance gain can be obtained by designing codes with the optimal power allocation. Furthermore, it was found that this optimal allocation depends highly on the channel coherence interval, as well as the particular algorithm used, and plays an increasingly critical role for fast channel dynamics. To further demonstrate the importance of optimizing pilot power allocation, similar simulations were run for more elaborate receivers as well as more bandwidth efficient modulation alphabets, for which the density evolution is not feasible. From these it is deduced that, the better the utilized receiver, in terms of joint estimation/decoding, the less pilot power is required. The presented pilot-symbol-assisted codes and the

corresponding receivers, apart from answering an important question of achievable performance with pilot symbols utilized to aid the phase estimation, can serve as a baseline, yet powerful, system, against which new code designs can be compared.

The pilot-symbol-assisted coding performs well for high values of N, as the loss in effective SNR as a result of pilot symbol insertion becomes negligible. On the other hand, for small or even moderate values of N, this SNR loss is considerable, and hence more elaborate signaling schemes should be considered.

As a first step in this direction, in Chapter III we approached the problem of optimal signaling from the information-theoretic point of view. Namely, we investigated the information capacity and, more importantly, the structure of the capacity achieving density.

Several interesting results were found regarding the structure of the capacity achieving density. It was shown that the capacity is achieved by the circularly symmetric input density, that is the density function that only depends on the amplitude of the variable. This implies that optimal signaling scheme uses all directions in the N dimensional complex space equally probable. Therefore, the problem of determining the structure of the optimal input density reduces to finding that of the amplitude variation of the input signal. Taking this as a starting point, it is proven that there exists a unique optimal amplitude density that is discrete and has infinitely many mass points with tails going to infinity. Furthermore, it is shown that there is always a mass point at zero, which implies that to achieve capacity the transmission should be kept *silent* with some positive probability.

The noncoherent channel resembles the two similar channel models: on one hand it has the amplitude variation of the coherent AWGN channel (constant envelope) while on the other it has phase variation of the Rayleigh fading AWGN channel

(uniform phase). The results about the structure of the capacity achieving density, outlined above, reveal where the noncoherent channel stands in comparison to these two close channel models. The fact that the optimum input signal is discrete in nature is reminiscent of the similar behaviour of the Rayleigh fading channel, while the fact that the tails of the optimal density go to infinity is similar to the behaviour of the coherent AWGN channel.

Complementary to these basic theoretical results, an asymptotic expression is derived that relates the mass point probabilities to mass point locations. It is shown that although the tails of the optimal density go to infinity, the probabilities decrease fast with increasing amplitude (square exponentially). As a result, numerical capacity evaluations that assume only a finite number of mass points result in small capacity loss. Several such evaluations were performed, and in particular it was shown that the mere introduction of a mass point at zero yields a considerable capacity gain over the more conventional single amplitude signaling (this gain is termed "shaping gain"). In addition to this, the loss associated with utilizing discrete modulation alphabets as opposed to the optimal circularly symmetric signaling (termed "disctretization loss") is identified and quantified. These two simple performance measures prove to be very useful from the code design point of view, which is the focus of Chapter IV, where these results are explicitly used to design the practical low-complexity coding schemes that come close to the capacity of this channel and outperform the existing designs.

In particular, Chapter IV considers the concatenated scheme where a high-performance binary outer code is used to introduce large memory, while a simple modulation code is used to match the channel characteristics. Due to the complexity constraints, the case of slow dynamics (small values of N) is considered separately from the case of

faster dynamics (moderate values of N). Within this context two different coding scheme designs that are inspired by the capacity results are proposed in Chapter IV.

For small values of N, the complexity at the demodulator (exponential in N) permits the use of efficient modulation codes with small discretization loss. In particular, a recently discovered unitary modulation, is used, which can be easily optimized to yield modulation code with good immunity to the harmful effects of random phase. As a binary code a serially concatenated turbo code is utilized. In this scenario, to benefit from the shaping gain, the inner code of the turbo code is used to introduce the transmission of zero mass point. It is demonstrated through an example that such a scheme powered with a relatively simple turbo code yields a very good bit-error rate performance, further emphasizing the capacity results.

In the second case, when moderate values of N are considered, a somewhat different approach is pursued. In this case, the complexity constraints prevent the use of powerful modulation codes, and as a result simple M-PSK modulation is used. A novel modulation scheme is proposed that, while using simpler modulation alphabets, handles the introduction of the zero mass-point. This modulation scheme is paired with the powerful low-density parity-check codes, that can be represented by factor-graphs and decoded quite reliably by using a message passing algorithm. Moreover, the factor-graph representation of the proposed modulation scheme allows a decoding with a message passing algorithm, that has linear complexity in N. Furthermore, utilized LDPC codes can be optimized taking into the account the overall structure of the factor graph including the modulation scheme. These ideas are demonstrated through a practical example. The obtained code outperforms all the existing designs by more than 0.5 dB at essentially the same complexity, and is within 0.3 dB of the capacity. It is shown that the channel-matched optimization of

the binary code plays a crucial role in the overall performance gain.

Regardless of the structure of the codes designed for the noncoherent channel, a certain condition has to be satisfied by all the codewords. This condition is the result of the fact that within the block of constant phase the two sequences that are rotated versions of each other are indistinguishable at the receiver side. All the codes designed in Chapters II and IV avoided this catastrophic behaviour by making sure that no such pair of codewords appears in the codebook. However, an alternative approach to this issue is possible: instead of avoiding codewords that are rotated version of each other, all such combinations are included in a codebook in a controlled manner, so that the ambiguity at the receiver side is vanished by assigning all such codewords to the same input sequence. The codes that possess such a property are called rotationally invariant (RI).

In Chapter V, we discuss RI codes and their properties under somewhat more flexible channel models. For analysis and code design purposes, several similar channel models are considered starting with a simple slow-dynamics model (constant phase over the entire codeword) and through a series of more realistic channel models finally arriving at a model that is very similar to the block-noncoherent channel model discussed earlier.

In the first part of Chapter V, RI codes are extended to the serially concatenated turbo codes, and the design guidelines for RI-SCTCM codes are outlined. Following these, several powerful high-rate RI-SCTCM codes are designed and simulated. As it is illustrated through these examples, in some cases RI property does not imply any penalty in the performance, while for others it comes at the cost of diminished performance.

In the second part, we introduce a notion of rotationally robust (RR) codes by

easing the condition of constant phase over the entire codeword. Namely, now the unknown phase can jump at random position within a codeword to another random realization. The RR codes are required to suffer only a finite – and independent from blocklength – number of input bit errors as a result of such a jump. We consider this property both in coherent and noncoherent setting, and prove that under certain conditions RI codes actually satisfy this property. To extend this result to SCTCM codes we propose a simple modification of the decoding algorithm, that makes the RI-SCTCM codes robust to phase jumps as well.

Aside from proving similar theorems of robustness to partial rotations for RI-SCTCM codes, designing practical RI-SCTCM schemes that are robust to partial rotations when decoded with standard iterative decoder (possibly with slight modifications), is a challenging problem for future studies.

These results constitute major contributions of this thesis. Though considerable progress has been made in both of the identified problems, coding and capacity, there are still possible improvements and extensions of these results. In the next section we investigate these future possible research topics.

6.2 Future Work

We initiate this section with a simple extension of the pilot-symbol-assisted modulation. A simple generalization of this coding scheme is to transmit the pilot symbol not in every block of constant phase, as was assumed in the above analysis, but once per every J number of blocks. The power loss due to pilot will now be on the order of $E_p/(JN)$ as opposed to E_p/N. However, it is not straightforward to predict which of the following two cases will have superior performance: 1) transmitting pilot of power E_p in every block; or 2) transmitting pilot of power JE_p once in every J blocks.

This is because, while power loss due to the pilot symbol insertion in both cases is essentially the same, it is not clear whether having very high quality phase estimate in one block (out of J) and none in others (case 2), is better/worse than having less reliable phase estimates in all the blocks (case 1). No matter what the conclusions will be, this approach gives another variable, J, for trade-off between the quality of estimation and effective SNR, and comprises an interesting future research direction.

The coding schemes presented in Chapter IV imitate the capacity results obtained in Chapter III and achieve very good performance. However, the two proposed schemes do not completely cover topic of the capacity-inspired coding. For instance, the coding scheme proposed in Section 4.3 for the moderate values of N, requires more involved mode-bit mappings. In particular, when the desired probability of zero is p_0, and m of the blocks are used at a time by the modulation scheme then $\log_2 \binom{m}{\lfloor mp_0 \rfloor}$ mode bits are required. The increase in m implies the increase of the rate gain in terms of number of mode bits per block of constant phase. Using the bound

$$\frac{1}{m+1} 2^{mH(p_0)} \leq \binom{m}{\lfloor mp_0 \rfloor} \leq 2^{mH(p_0)} \tag{6.1}$$

where $H(p_0)$ represents the binary entropy function, we can see that, in the limit, the rate gain as a result of mode bits is

$$\lim_{m \to \infty} \frac{1}{m} \log_2 \binom{m}{\lfloor mp_0 \rfloor} = H(p_0). \tag{6.2}$$

For example, the code presented as an example for the case of $N = 7$ has $p_0 = 1/4$ and uses $m = 4$. As a result, the rate gain is 0.5 (bits per block), while the achievable limit is $H(1/4) = 0.81$ (bits per block). This suggests that there is a potential 32% increase in the rate gain, which can prove very crucial. However, for higher values of m, there is a considerable complexity increase in comparison to the modulation

scheme that does not introduce the zero point. Therefore, the decoding algorithm has to be simplified – perhaps at the cost of abandoning optimal demodulation – to accommodate the increase in m without sacrificing low complexity that these coding schemes were designed for. These ideas form a research topic for future studies.

On the other extreme, the problem of capacity of the noncoherent channel has not been fully solved. The asymptotic results of Section 3.3 can be further extended to yield even better results. For example, the sphere hardening effect can be more rigorously proven using the derived asymptotic results. On the other hand, better bound and/or expressions for the capacity itself can be derived. An interesting approximation that can be derived by combining the results of Chapter III with the intuitive reasoning is

$$C_N(\gamma) = \frac{I(\mathbf{x},\mathbf{y})}{N} \simeq \frac{N-1}{N}\log_2\big(\gamma + 1 - \frac{1}{2N}\big) + \frac{1}{N}C_1(\gamma), \tag{6.3}$$

where γ denotes the signal-to-noise ratio. This simple approximation is very interesting, in that it reveals how the capacity of the coherent AWGN channel is approached with increasing N and γ. In particular, it shows that for a fixed N, and with increasing γ, the capacity of the noncoherent channel is approximately $N/(N-1)$ times that of the coherent channel plus the extra term that is capacity for the case of $N = 1$. This is expected, because although the high SNR case can be argued to translate this channel into the coherent case (*e.g.*, via phase estimation), exactly one out of the $2N$ real dimensions will be lost (out of N pairs of phases and amplitudes, one of the phases can not be recovered), and thus only $N-1$ complex dimensions exhibit coherent behaviour and the remaining amplitude variation results in the extra term. This result is in agreement with [56], where it was partially shown that for block memoryless channels the limit of the infinite blocklength, N, the capacity of the co-

herent case is approached. In addition, this approximation precisely quantifies this convergence. A more rigorous proof of this approximation is another future research topic.

Connected to this is the derivation of the efficient algorithm for the numerical optimization of the mutual information. More precisely, we know that for any input density the following two sets of conditions need to be satisfied

$$F_{p_a}(a_k) = \lambda + \mu a_k^2 \tag{6.4}$$

$$F'_{p_a}(a_k) = 2\mu a_k \tag{6.5}$$

by the two sets of variables, mass point probabilities p_k and positions a_k. Using these conditions, one can devise simple procedures to satisfy each of the conditions by varying one set of variables provided the others are fixed. In particular, the methods used in proving the existence of the mass point at zero in the Section B.7 of Appendix B can be adapted for this purpose. By utilizing such methods it can be shown that carefully chosen procedures will necessarily increase the mutual information at each step. This way an iterative algorithm is obtained that – because the mutual information is bounded for finite power – is guaranteed to converge. The proof of the fact that such a procedure yields a global optimum (or a sequence of local extreme points that approach the global optimum with increasing number of mass points), as well as a more rigorous derivation of the aforementioned procedures is of practical interest.

APPENDICES

APPENDIX A

DERIVATION OF THE DISTRIBUTION FUNCTIONS FOR PSA

In this appendix we derive the expressions for initial message distributions for the proposed receiver models of Chapter II assuming the transmission of the all-one codeword. These distribution functions are used to run density evolution in order to analyze these receivers. The expressions for messages are repeated here for convenience, and throughout the appendix, the normalization $E_s = 1$ is assumed.

A.1 M-PO Receiver

For M-PO receiver, the initial message is given by

$$\mu_{FX_i} = \log \frac{p(z_i, z_0|x_i = +1)}{p(z_i, z_0|x_i = -1)} = \log \frac{I_0(|z_0\sqrt{E_p} + z_i\sqrt{E_s}|/\sigma^2)}{I_0(|z_0\sqrt{E_p} - z_i\sqrt{E_s}|/\sigma^2)}. \tag{A.1}$$

Observe that z_i and z_0 have joint probability density function

$$\begin{aligned} p(z_i, z_0|x_i = 1) &= \int_0^{2\pi} p(z_i, z_0|\theta, x_i = 1)d\theta \\ &= \frac{1}{(2\pi\sigma^2)^2} \exp(\frac{-|z_i|^2 - |z_0|^2 - E_p - E_s}{2\sigma^2}) I_0\left(\frac{|z_0\sqrt{E_p} + z_i\sqrt{E_s}|}{\sigma^2}\right). \end{aligned} \tag{A.2}$$

After some algebra the cumulative distribution function of the message in (A.1) can be found as

$$P(\mu_{FX} \leq q) = \int_{\alpha_l}^{\infty} \int_{-1}^{1} \exp(\frac{-(E_p + 1)(\alpha + 4E_p)}{8E_p\sigma^2}) I_0(\frac{\sqrt{\alpha}}{\sigma^2}) \frac{g(\alpha, q, t)}{\sqrt{1 - t^2}} dt d\alpha, \tag{A.3}$$

where $g(\alpha, q, t)$ is given by

$$g(\alpha,q,t)=\frac{1}{a}\Big[e^{-ax}\cosh(b\sqrt{x})+\frac{\sqrt{\pi}}{2}\frac{b}{2\sqrt{a}}e^{\frac{b^2}{4a}}\big(\operatorname{erfc}(\frac{-b}{2\sqrt{a}}+\sqrt{ax})-\operatorname{erfc}(\frac{b}{2\sqrt{a}}+\sqrt{ax})\big)\Big],$$

with

$$a=\frac{E_p+1}{8E_p\sigma^2} \tag{A.4}$$

$$b=\frac{E_p-1}{4E_p\sigma^2}\sqrt{\alpha(1-t)/2} \tag{A.5}$$

$$x=\sigma^4 f^2(I_0(\frac{\sqrt{\alpha}}{\sigma^2})e^{-q}), \tag{A.6}$$

where $f(\cdot)$ is the inverse of $I_0(\cdot)$ (with the convention $f(y)=0$ if $y\leq 1$). The lower limit of integration is $\alpha_l=\sigma^4 f^2(e^q)$.

A.2 E-PO Receiver

For the case of PO receiver with explicit phase estimation, the message has the form

$$\mu_{FX}(r,t)=\log\frac{\int_{-\pi}^{\pi}e^{\sqrt{E_s}r\cos(t-x)/\sigma^2}T(x)dx}{\int_{-\pi}^{\pi}e^{-\sqrt{E_s}r\cos(t-x)/\sigma^2}T(x)dx}. \tag{A.7}$$

where $r=|z_i|$, $t=\angle z_i-\hat{\theta}$ and $T(x)=p_{\hat{\theta}|\theta}(x|\theta=0)$ is given in (2.15).

Utilizing the approximation $\mu_{FX}(r,t)=h(r)\cos(t)$, mentioned in Section 2.4.1 of Chapter II, results in the following expression for the cdf of μ_{FX}

$$P(\mu_{FX}\leq q)=\int_{-\pi}^{\pi}\int_{\pi/2}^{\pi}T(x)s(x,t)dt\,dx, \tag{A.8}$$

with $s(x,t)$ defined as

$$s(x,t)=e^{\frac{-\sin^2(x-t)}{2\sigma^2}}\Big(\frac{1}{\pi}e^{-H^2}+\frac{1}{\sqrt{2\pi\sigma^2}}\cos(x-t)\operatorname{erfc}(H)\Big) \tag{A.9}$$

where erfc denotes the complimentary error function and

$$H=\frac{h^{-1}(|q/\cos(t)|)}{\sqrt{2\sigma^2}}. \tag{A.10}$$

A.3 M-QDF Receiver

For the M-QDF receiver, the initial message resembles the M-PO message as can be evidenced from its definition

$$\mu_{FX_i} = \log \frac{I_0(|w_i a^{(l)} + z_i\sqrt{E_s}|/\sigma^2)}{I_0(|w_i a^{(l)} - z_i\sqrt{E_s}|/\sigma^2)}, \tag{A.11}$$

with all the involved variables defined in Section 2.3.2. As it was mentioned earlier, we need to find the conditional cdfs $P(\mu_{FX_i} \leq q|\mathbf{x}, \mathbf{v}^i)$. To do that, observe that w_i in (2.11) is a complex Gaussian random variable with mean $b(\mathbf{v}^i)e^{j\theta}$, variance σ^2 per real dimension and is independent of z_i conditioned on $\mathbf{v}^i$, $\mathbf{x}$ and the channel parameter θ, with

$$b(\mathbf{v}^i) = \frac{\mathbf{y}^T\mathbf{v}^i}{\|\mathbf{v}^i\|}. \tag{A.12}$$

Furthermore, due to the similarity of the M-QDF message in (A.11) with the message for the M-PO receiver, the same approach can be followed for the evaluation of the corresponding cumulative distribution functions, resulting in the following expressions

$$P(\mu_{FXi} \leq q|n_+, n_-) = K\int_{\alpha_l}^{\infty}\int_{\beta_l}^{\infty}\int_{-1}^{1}\big[G(\alpha,\beta,t,+1) + G(\alpha,\beta,t,-1)\big]\frac{dt}{\sqrt{1-t^2}}d\beta d\alpha \tag{A.13}$$

where $G(\alpha, \beta, t, c)$ is given by

$$G(\alpha,\beta,t,c) = \exp(-\frac{g(1/a,c)}{2\sigma^2})I_0(\frac{\sqrt{g(b/a,c) + (\alpha-\beta)b/(2a)}}{\sigma^2}) \tag{A.14}$$

$$\text{with} \quad g(r,c) = \frac{1}{2}((1+r^2)(\frac{\alpha+\beta}{2} + c(1-r^2)\sqrt{\frac{\alpha\beta(1-t)}{2}}) \tag{A.15}$$

$$K = \frac{1}{32\pi\sigma^4a^2}\exp(-\frac{b^2+1}{2\sigma^2}) \tag{A.16}$$

$$b = b(n_+, n_-) = \frac{n_+ - n_- + E_p}{\sqrt{n_+ + n_- + E_p}} \tag{A.17}$$

$$\beta_l = \sigma^4 f^2(I_0(\frac{\sqrt{\alpha}}{\sigma^2})e^{-q}) \tag{A.18}$$

and $f(\cdot)$ and α_l are defined in Section A.1. Observe that the expression for the cdf of the M-PO message in (A.3) is a special case of equation (A.13) for $n_+ = n_- = 0$, after explicitly integrating with respect to β.

APPENDIX B

PROOF OF CAPACITY RESULTS

In this appendix we give the proofs of the results in Chapter III. Sections are generally self containing and have few references to the Chapter III.

B.1 Simplification of the Mutual Information Expression

In view of Lemma 3.1, for the optimization problem described in (3.3) it suffices to consider circularly symmetric input densities. It was mentioned earlier that such densities can be parameterized by a function of one real variable, that is the density of the input amplitude variable $a = \|\mathbf{x}\|/\sigma$. In this appendix, we derive the expression for the mutual information induced by the circularly symmetric input random vector, in terms of its amplitude density.

Let $p(\mathbf{x}) = f(\|\mathbf{x}\|/\sigma)/\sigma^{2N}$ be such an input density and observe that in this case the induced output density also possesses such a property, *i.e.*, $p(\mathbf{y}) = g(\|\mathbf{y}\|/\sigma)/\sigma^{2N}$ for some function $g : [0, \infty) \to [0, \infty)$. For a fixed nonzero $\mathbf{x}$, we first consider the integral with respect to $\mathbf{y}$ in (3.4). Let U be a unitary matrix[1] whose first row is $\mathbf{x}^H/\|\mathbf{x}\|$, and the rest $N - 1$ rows are any set of orthonormal vectors, that are

[1]An implicit assumption at this point is that $N > 1$. However, the end result (3.9) is valid for $N = 1$ as well.

also orthogonal to $\mathbf{x}$. By changing variables from $\mathbf{y}$ to $\mathbf{w} = [r_1 e^{j\theta_1}, \ldots r_N e^{j\theta_N}]^T = U\mathbf{y}/\sigma$ the integrand becomes a function of only r_1 and $r_1^2 + \ldots + r_N^2$, and thus the phase variables θ_i can be integrated out. Finally, by introducing new variable $r = \sqrt{r_1^2 + \ldots + r_N^2}$ and integrating out with respect to all the variables except r_1 and r, the integral over $\mathbf{y}$ in (3.3) is reduced to

$$\int_0^\infty \int_0^r \left(\frac{r^2 - r_1^2}{2}\right)^{N-2} \frac{r r_1}{(N-2)!} \exp(-\frac{r^2 + a^2}{2}) \times \\ \times I_0(ar_1) \log \frac{I_0(ar_1)}{g(r)} dr_1 dr - N \log(2\pi) - (N + a^2). \quad \text{(B.1)}$$

The expression in front of the logarithm under the integral sign is the conditional density $p(r, r_1|a)$. By a further manipulation (B.1) becomes

$$\int_0^\infty p_1(r_1|a) \log I_0(ar_1) dr_1 - \int_0^\infty p_N(r|a) \log g(r) dr - N \log(2\pi) - (N + a^2) \quad \text{(B.2)}$$

where the function $p_n(r|a)$ is defined by

$$p_n(r|a) = e^{-(r^2+a^2)/2} r \left(\frac{r}{a}\right)^{n-1} I_{n-1}(ar). \quad \text{(B.3)}$$

In a similar fashion, the integral over $\mathbf{x}$ in (3.4) can be reduced to an integral over a, which results in the final expression in (3.9).

B.2 The Kuhn-Tucker Condition

In this section we derive the necessary and sufficient conditions for an input amplitude density to be the maximum one. First, we eliminate the power constraint via the introduction of the Lagrange multiplier, and convert the constrained problem in (3.12) into unconstrained one

$$\mathcal{C} = \sup_{p_a(a)} \int_0^\infty (F_{p_a}(a) - \mu(a^2 - \gamma)) p_a(a) da \quad \text{(B.4)}$$

for some nonnegative μ. It is known [50] that if the supremum is achieved in the unconstrained problem by some input density, it is also achieved in the constrained problem by the same input density, and that this input density satisfies the power constraint with equality if μ is not zero.

Following along the lines of [82, 25] we can obtain the necessary and sufficient conditions for an input density to be the maximizing one. The result relies on the well-known fact of convex optimization (see [41, 50]) which says that a concave functional $J(p)$ on a convex set achieves its maximum at p^* if and only if

$$\frac{d}{dt}J(p^* + t(p - p^*))\Big|_{t=0^+} \leq 0 \qquad \text{for any } p. \tag{B.5}$$

Applying this result to the optimization problem in (B.4) we get the following: $p_a^*(a)$ achieves the maximum in (B.4) if and only if

$$\int_0^\infty [F_{p_a^*}(a) - \mu(a^2 - \gamma)](p_a(a) - p_a^*(a))da \leq 0 \tag{B.6}$$

for any density $p_a(a)$. Furthermore, applying Corollary 2 of [82] we get the necessary and sufficient condition (called Kuhn-Tucker condition):

$$F_{p_a^*}(a) \leq \lambda + \mu a^2 \tag{B.7}$$

with equality if a is the mass point of density $p_a^*(a)$, and where λ is

$$\lambda = \int_0^\infty F_{p_a^*}(a)p_a^*(a)da - \mu\gamma = \mathcal{C} - \mu\gamma. \tag{B.8}$$

B.3 Existence and Uniqueness of the Maximizing Density

In this section we proceed to prove that the optimization problem of (B.4) and hence the constrained problem of (3.12) has a unique solution, that is, there exists *the* input amplitude density for a that achieves the maximum. The methodology followed in doing so parallels the approach of [82, 25].

In the following we consider densities as being elements of the dual of the space of the bounded continuous functions, that is, as linear functionals, where a density $p(a)$ acts on a bounded continuous function $\phi(a)$ as $p(\phi) = \int \phi(a)p(a)da$. Therefore, we can think of a "mass point at a_0" as being a linear functional δ that acts on $\phi(a)$ as $\delta(\phi) = \phi(a_0)$. This approach allows us to include the discrete random variables into the consideration as well.

The results of this appendix rely on the fact that the strictly concave continuous functional on a convex compact set achieves its unique maximum. Here continuity of functionals as well as the compactness of the set is understood in the weak* topology (*e.g.*, see [41, Sect. IV-3.3]). The convergence in this topology is "pointwise convergence", that is the sequence p_k is said to converge to p if the real sequence $p_k(\phi)$ converges to $p(\phi)$ for any bounded continuous function $\phi(a)$.

In [25] it was shown that the set of second-moment-constrained input densities is convex and compact. In the following we establish that the mutual information is continuous and strictly concave, which by the statement above proves the existence and uniqueness of the maximizing density.

B.3.1 Weak-* continuity of mutual information

We can rewrite $I(p_a)$ in (3.9) as

$$I(p_a) = \int_0^\infty \left[\int_0^\infty p_1(r|a) \log \frac{p_1(r|a)}{r} dr - \int_0^\infty p_N(r|a) \log \frac{c_{N-1}}{r^{2N-1}} dr \right] p_a(a) da - \int_0^\infty p_N(r) \log p_N(r) dr \quad \text{(B.9)}$$

where $p_N(r)$ is induced by the input density $p_a(a)$. We first show that the second term is continuous. Let p_a^k be a sequence of input densities that converges (in weak* sense) to p_a (and satisfy the power constraint) . Observe that $p_N(r)$ is a continuous

function[2] of $p_a(a)$ for all $r \geq 0$, and so is $p_N(r) \log p_N(r)$. Therefore, we only need to find an integrable function $\zeta(r)$, so that

$$|p_N^k(r) \log p_N^k(r)| \leq \zeta(r) \tag{B.10}$$

where p_N^k is the output density induced by p_a^k, and then use the Lebesgue dominated convergence theorem. To that end observe that

$$p_N(r|a) \leq \exp(-\frac{(r - \sqrt{a^2 + 2N - 1})^2}{2}) \leq \begin{cases} 3^{\frac{a^2+2N-1}{r^2}} & r \geq 1 \\ 1 & r < 1 \end{cases} \tag{B.11}$$

and hence

$$p_N^k(r) \leq \min\{1, 3\frac{\gamma + 2N - 1}{r^2}\}, \tag{B.12}$$

which together with the simple inequality $|x \log x| \leq 2x^{3/4}$ for $x \in [0, 1]$ implies

$$|p_N^k(r) \log p_N^k(r)| \leq \zeta(r) = 2 \min\{1, \frac{3^{3/4}(\gamma + 2N - 1)^{3/4}}{r^{3/2}}\}. \tag{B.13}$$

Since $\zeta(r)$ is obviously an integrable function, the required result follows from the Lebesgue dominated convergence theorem.

To show that the first expression in (B.9) is continuous we need to show that

$$\lim_{k \to \infty} \int_0^\infty G(a) p_a^k(a) da = \int_0^\infty G(a) p_a(a) da \tag{B.14}$$

where $G(a)$ is the expression inside the square brackets in (B.9). The function $G(a)$, though continuous, is not bounded, and therefore (B.14) does not merely follow from the definition of the weak* convergence. However, using the inequality

$$\begin{aligned} 0 \leq G(a) &\leq (2N - 1) E(\log(r_N)|a) \leq (2N - 1) \log(E(r_N|a)) \\ &\leq (2N - 1) \log(\sqrt{a^2 + 2N - 1} + 1) \leq (2N - 1)\sqrt{a^2 + 2N - 1}, \end{aligned} \tag{B.15}$$

---

[2]This follows because $p_N(r|a)$ is a bounded continuous function of a.

the tails of the integrals in (B.14) can be bounded uniformly in k as[3]

$$\int_t^\infty G(a)p_a^k(a)da \le (2N-1)\int_t^\infty \sqrt{a^2+2N-1}\,p_a^k(a)da$$
$$\le \frac{2N-1}{\sqrt{t^2+2N-1}}\int_t^\infty (a^2+2N-1)p_a^k(a)da \le \frac{(2N-1)(\gamma+2N-1)}{\sqrt{t^2+2N-1}}. \quad \text{(B.16)}$$

Therefore, for any given $\varepsilon > 0$, we can choose a big enough t so that for all k the LHS of (B.16) is less than $\varepsilon/4$. The function (in a) which is equal to $G(a) - G(0)$ for $a < t$ and $G(t) - G(0)$ for $a \ge t$, is bounded and continuous, and therefore using the definition of the weak* convergence, for some k', the sequence in (B.14) will be within $\varepsilon/2$ of the limiting value for all $k > k'$ (with upper limits of integrals being t). Combining these two results with the fact that any constant function is bounded and continuous the result in (B.14) is established. This proves that the first term in (B.9) is continuous as well, and, therefore, so is the mutual information.

B.3.2 Strict concavity of mutual information

The first term in (B.9) is linear in $p_a(a)$. The second term is a strictly concave function of $p_N(r)$, and since the latter is linear in $p_a(a)$, the second term is concave in $p_a(a)$. To show that it is actually strictly concave in $p_a(a)$, we need the mapping from $p_a(a)$ to $p_N(r)$ to be injective. Since both a and r are nonnegative random variables, it suffices to prove the injectivity of the mapping from the density of $\alpha = a^2$ to the density of $\rho = r^2$. Observe that the conditional density $p(\rho|\alpha) = p_N(\sqrt{\rho}|\sqrt{\alpha})/(2\sqrt{\rho})$ has the characteristic function

$$\Phi_{\rho|\alpha}(s|\alpha) = \int_0^\infty p(\rho|\alpha)e^{s\rho}d\rho = \left(\frac{1}{1-2s}\right)^N \exp\left(\frac{2s\alpha}{1-2s}\right) \quad \text{(B.17)}$$

[3]Obviously this holds for $p(a)$ as well

which implies that

$$\Phi_\rho(s) = \int_0^\infty \left[\int_0^\infty p(\rho|\alpha)p(\alpha)d\alpha\right] e^{s\rho}d\rho = \int_0^\infty \left(\frac{1}{1-2s}\right)^N \exp(\frac{2s}{1-2s}\alpha)p(\alpha)d\alpha$$
$$= \left(\frac{1}{1-2s}\right)^N \Phi_\alpha(\frac{2s}{1-2s}) \quad \text{(B.18)}$$

where Φ_α is the characteristic function of α. Letting $z = \frac{2s}{1-2s}$ we can rewrite this formula as

$$\Phi_\rho(\frac{z}{2(1+z)}) = (1+z)^N \Phi_\alpha(z) \quad \text{(B.19)}$$

Now let p_α and p'_α be two densities with corresponding output densities p_ρ and p'_ρ. Assume p_ρ and p'_ρ are equal. Then, their corresponding characteristic functions Φ_ρ and Φ'_ρ are also equal. Equation (B.19) in turn implies that the characteristic functions of α and α' are equal for all values of z in the complex plane except at $z = -1$, and since the characteristic functions are continuous, they are equal everywhere. This implies that α and α' are equal in distribution, which proves our claim.

Since the mutual information in (B.9) is a difference of a linear and a strictly concave functional in p_a, it is also strictly concave in input density p_a.

B.4 Extension of $\Psi_R(z)$ to the Imaginary Axis

In this appendix we find an alternative representation for $\Psi_R(z)$ that is valid in the region $\arg(z) \in (\pi/4, 3\pi/4)$, and thus obtain the values of this function on the imaginary axis. Observe that the expression in (3.17b) is valid for all values of z such that $\mathrm{Re}(z) > 0$; in particular, it is valid for $\arg(z) \in (\pi/4, \pi/2)$. We will start with z in this region and transform (3.17b) into the form that is directly extended onto and beyond the imaginary axis.

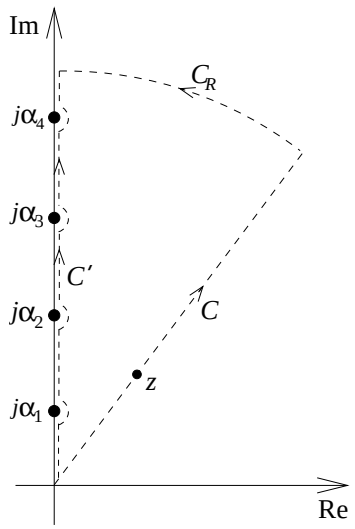


Figure B.1: Contour of integration for extending $\Psi_R(z)$ to the imaginary axis.

For that matter, let z be any nonzero complex number with $\arg(z) \in (\pi/4, \pi/2)$ and make a change of variables in (3.17b) $w = zr_1$ to get

$$\Psi_R(z) = \int_C e^{-(w/z)^2/2} \frac{w}{z^2} I_0(w) \log I_0(w) dw \tag{B.20}$$

where C is the induced contour of integration: the straight line from the origin to infinity in the direction of z (see Fig. B.1). The integrand is an analytic function in w which has logarithmic branching at the zeros of $I_0(w)$ on the imaginary axis (call them $j\alpha_k$, $k = 1, 2 \ldots$). Draw another contour C' along the positive imaginary axis from the origin to infinity which also goes around the zeros of $I_0(w)$ in the small semicircles of radius ε so that the zeros remain on the left of the contour. Connect the contours C and C' by the circular segment C_R (centered at the origin) of radius R going counter clockwise from C to C'. The integrand is analytic in the triangular shaped region surrounded by these three contours, and therefore we can apply the Cauchy theorem to obtain

$$\int_C X(w)dw + \int_{C_R} X(w)dw = \int_{C'} X(w)dw. \tag{B.21}$$

where $X(w)$ is the integrand in (B.20), and integration along C and C' is only taken up to the point of intersection with C_R. Observe that on C_R, $|w| = R$ and $\arg(w/z) \in [0, \pi/4)$ and, therefore,

$$\mathrm{Re}\Big(\frac{w^2}{2z^2}\Big) \geq \cos(2\alpha)R^2/|z|^2, \tag{B.22}$$

where $\alpha = \pi/2 - \arg(z) \in [0, \pi/4)$. Taking this into account we get

$$\begin{aligned}\left|\int_{C_R} X(w)\right| \leq \int_{C_R} |X(w)||dw| &\leq \alpha R \max_{w\in C_R} |X(w)| \\ &\leq \alpha R e^{-cos(2\alpha)/|z|^2R^2} \frac{R}{z^2} I_0(R)|R + jR| \end{aligned} \tag{B.23}$$

Since $\cos(2\alpha) > 0$, this expression goes to zero as R goes to infinity. Therefore, the integral along C is equal to the integral along C'. On the small semicircles of radius ε around the zeros of $I_0(w)$ on the positive imaginary axis, $I_0(w)\log I_0(w)$ goes to zero (uniformly for all values of w on the semicircle) as ε goes to zero. Hence, the integral over the semicircles goes to zero as well, and the final result becomes

$$\Psi_R(z) = -\int_0^\infty e^{r^2/(2z^2)} \frac{r}{z^2} J_0(r) \log|J_0(r)| dr - j\pi \int_0^\infty e^{r^2/(2z^2)} \frac{r}{z^2} J_0(r)\xi(r) dr \tag{B.24}$$

where $\xi(r) = 0$ if $r \in (0, \alpha_1)$ and $\xi(r) = k$ if $r \in (\alpha_k, \alpha_{k+1})$. The second term is due to the fact that the logarithm jumps in value by $j\pi$ every time a zero is passed. Observe that the last equation is valid for all z such that $\arg(z) \in (\pi/4, 3\pi/4)$. It is worth noting that $\Psi_R(z)$ is a many valued function, with infinite number of branches.

B.5 Nonzero Imaginary Part of $\Psi_R(z)$

In this appendix it is shown that $\Psi_R(z)$ is not purely real on the imaginary axis $z = jb$ by bounding the imaginary part of $\Psi_R(jb)$ away from zero. First observe

from (3.19) that

$$\frac{\mathrm{Im}(\Psi_R(jb))}{\pi} = \int_0^\infty e^{-r^2/(2b^2)} \frac{r}{b^2} J_0(r)\xi(r)dr = \sum_{k=1}^\infty k \int_{\alpha_k}^{\alpha_{k+1}} e^{-r^2/(2b^2)} \frac{r}{b^2} J_0(r)dr$$
$$= \sum_{k=1}^\infty \int_{\alpha_k}^\infty e^{-r^2/(2b^2)} \frac{r}{b^2} J_0(r)dr = \sum_{k=1}^\infty \int_{\alpha_k/b}^\infty e^{-r^2/2} r J_0(br)dr = \sum_{k=1}^\infty c_k. \quad \text{(B.25)}$$

Since $|J_0(r)| \leq 1$, a simple bound for $|c_k|$ is $|c_k| \leq \exp(-\alpha_k^2/(2b^2))$. Further, using the fact that the sequence α_k satisfies $\alpha_{k+1} - \alpha_k > \pi/2$ (actually, it can be seen from [45, Sect. VII-98] that $\alpha_k \simeq (3/4+k)\pi$) the following bound can be established

$$\sum_{k=3}^\infty |c_k| \leq \frac{2}{\pi} \sum_{k=3}^\infty e^{-\alpha_k^2/(2b^2)} (\alpha_k - \alpha_{k-1}) \leq \frac{2}{\pi} \int_{\alpha_2}^\infty e^{-x^2/(2b^2)} dx \leq \frac{2b^2}{\pi\alpha_2} e^{-\alpha_2^2/(2b^2)} \quad \text{(B.26)}$$

A lower bound on $|c_1|$ will now be obtained using integration by parts[4]. Let β be *the* zero of $J_1(r)$ that lies between α_1 and α_2 and consider

$$|c_1| = \Big| \int_{\alpha_1/b}^\infty e^{-r^2/2} r J_0(br)dr \Big| = b \int_{\alpha_1/b}^\infty e^{-r^2/2} J_1(br)dr$$
$$= b \int_{\alpha_1/b}^{\beta/b} e^{-r^2/2} J_1(br)dr + b \int_{\beta/b}^\infty e^{-r^2/2} J_1(br)dr$$
$$\geq b e^{-\beta^2/(2b^2)} \int_{\alpha_1/b}^{\beta/b} J_1(br)dr - \frac{b^2}{\beta} e^{-\beta^2/(2b^2)}$$
$$= (|J_0(\beta)| - \frac{b^2}{\beta}) e^{-\beta^2/(2b^2)}.$$

The imaginary part of $\Psi_R(jb)$ can now be bounded away from zero as

$$\frac{|\mathrm{Im}(\Psi_R(jb))|}{\pi} = \Big| \sum_{k=1}^\infty c_k \Big| \geq |c_1| - \sum_{k=2}^\infty |c_k|$$
$$\geq (|J_0(\beta)| - \frac{b^2}{\beta}) e^{-\beta^2/(2b^2)} - (1 + \frac{2b^2}{\pi\alpha_2}) e^{-\alpha_2^2/(2b^2)} \quad \text{(B.27)}$$

which is positive for small values of b (*e.g.*, for $0 < b < 1.23$).

---

[4]Actually the exact expression for c_k can be found in [73] as

$$c_k = e^{-b^2/2} - e^{-(\alpha_k/b)^2/2} \sum_{n=1}^\infty \Big(\frac{\alpha_k}{b^2}\Big)^n J_n(\alpha_k)$$

However, we will use simpler bounding techniques to get the desired result

B.6 Bound on Lagrange Multiplier

In this appendix we prove Lemma 3.4 which states that the Lagrange multiplier μ is strictly less than $1/2$.

Proof of Lemma 3.4: The left-hand side of (3.22) is zero at $a = a_k$ and positive otherwise. Therefore, its derivative is zero at $a = a_k$ as well. Using the formula

$$\frac{\partial}{\partial a} p_N(r|a) = a(p_{N+1}(r|a) - p_N(r|a))$$

and differentiating the left-hand side of (3.22) with respect to a (and evaluating at $a = a_k \neq 0$) we get

$$\int_0^\infty (p_2(r|a_k) - p_1(r|a_k)) \log \frac{p_1(r|a_k)}{r} dr - \\ - \int_0^\infty (p_{N+1}(r|a_k) - p_N(r|a_k)) \log \frac{c_{N-1} p_N(r)}{r^{2N-1}} dr = 2\mu \quad \text{(B.28)}$$

Noting that

$$\int_0^\infty p_N(r|a) r^2 dr = a^2 + 2N, \quad \text{(B.29)}$$

we can simplify the left-hand side of (B.28) to

$$\int_0^\infty (p_2(r|a_k) - p_1(r|a_k)) \log I_0(a_k r) dr - \\ - \int_0^\infty (p_{N+1}(r|a_k) - p_N(r|a_k)) \log(\sum_i p_i e^{-a_i^2} S_{N-1}(a_i r)) dr \quad \text{(B.30)}$$

Observe that for each N and a there exists $R > 0$ such that $p_{N+1}(r|a) - p_N(r|a)$ is negative for $r < R$ and positive for $r > R$. Therefore, for any increasing function $f(r)$, we have

$$\int_0^\infty (p_{N+1}(r|a) - p_N(r|a)) f(r) = \int_0^\infty (p_{N+1}(r|a) - p_N(r|a))(f(r) - f(R)) \geq 0 \quad \text{(B.31)}$$

where the equality follows from the fact that $\int (p_{N+1}(r|a) - p_N(r|a)) dr = 0$, and the inequality follows from the fact that the integrand is the product of two negative

terms to the left of R, and the product of two positive terms to the right of R. Similarly, for a decreasing function $f(r)$, the inequality in (B.31) is reversed.

Using this observation and noting that the expression inside the logarithm in the second integral in (B.30) is the sum of increasing functions we get that the second term is positive. Now observe that $I_0(x)\exp(-x)$ is a decreasing function and therefore, using (B.31), for the first term we have

$$\int_0^\infty (p_2(r|a_k) - p_1(r|a_k)) \log I_0(a_k r) dr \leq \int_0^\infty (p_2(r|a_k) - p_1(r|a_k)) a_k r dr$$
$$= a_k \sqrt{\frac{\pi}{2}} e^{-a_k^2/4} \frac{1}{2} \left(I_0(\frac{a_k^2}{4}) + I_1(\frac{a_k^2}{4}) \right) < 1 \quad \text{(B.32)}$$

where the last inequality follows by using the fact that $(I_0(x) + I_1(x))/2 < e^x/\sqrt{2\pi x}$. Combining these results we get that the left-hand side of (B.28) is less than 1 and hence $\mu < 1/2$, which we set out to prove. ■

B.7 Mass Point at Zero

In this section we prove Theorem 3.5, that there always exists a mass point at zero.

Let $\mathbf{a} = \{0, a_1, a_2 \ldots\}$ and $\mathbf{p} = \{p_0, p_1, p_2 \ldots\}$ be the locations and probabilities of the mass points, respectively, so that the pair $(\mathbf{a}, \mathbf{p})$ denotes a valid input amplitude density. For notational simplicity also assume that the sequence $\mathbf{a}$ is (strictly) increasing and let $I(\mathbf{a}, \mathbf{p})$ be the mutual information corresponding to this input density.

We start with the simple case of $N = 1$. In this case this fact is both very intuitive and easy to prove. In particular, for any given discrete input density with no mass point at zero, let $a_1 > 0$ be the smallest amplitude with nonzero probability. Evaluating the partial derivative of the mutual information expression with respect

to a_1 we get

$$p_1 a_1 \int_0^\infty (p_2(r|a_1) - p_1(r|a1)) \log \frac{p_1(r|a_1)}{p_1(r)} dr < 0, \tag{B.33}$$

where the inequality follows by observing that the ratio $\frac{p_1(r|a_1)}{p_1(r)}$ is a decreasing function of r, and then using a simple fact used in Section B.6 in the proof of Lemma 3.4. This in turn implies that reducing a_1 increases the mutual information and hence the original input density can not be the optimal one.

In the general case of $N > 1$, it is not always true that reducing the smallest mass point will increase the mutual information. However, it can be shown that it is always possible to alter the input density in a certain way that the mutual information indeed increases.

We start with a case when there is a single mass point. In this case, the strategy is to *introduce* the zero mass point of some positive probability and increase the amplitude of the nonzero mass point accordingly so as to balance the total power. More precisely, let a_1 be the amplitude of the single mass point, and consider for some positive x, the new two-mass-point input density pair

$$\mathbf{p}^*(x) = \{x, 1-x\} \tag{B.34a}$$

$$\mathbf{a}^*(x) = \{0, a\sqrt{\frac{1}{1-x}}\}, \tag{B.34b}$$

which corresponds to introducing a mass point at zero of probability x and accordingly increasing the nonzero amplitude to keep the average power the same. Observe that the case of $x = 0$ corresponds to the original input density pair. We will show that

$$\frac{d}{dx} I(\mathbf{a}^*(x), \mathbf{p}^*(x))\Big|_{x=0} > 0, \tag{B.35}$$

which implies that there is a $x_0 > 0$ so that $I(\mathbf{a}^*(x_0), \mathbf{p}^*(x_0)) > I(\{a_1\}, \{1\})$, proving our claim.

Observe that

$$\frac{d}{dx} I(\mathbf{a}^*(x), \mathbf{p}^*(x))\big|_{x=0} = \frac{\partial I}{\partial p_0} - \frac{\partial I}{\partial p_1} + \frac{1}{2}\frac{a_1}{p_1}\frac{\partial I}{\partial a_1}, \tag{B.36}$$

where the partial derivative with respect to a_k is the left-hand side of (B.28) in Section B.6 multiplied by $p_k a_k$, and the partial derivative with respect to p_k is given by $\frac{\partial I}{\partial p_k} = F_{p_a}(a_k) - 1$. Therefore, the derivative in (B.35) after some manipulation becomes

$$\frac{a_1^2}{2} + \int_0^\infty \left(\frac{a_1^2}{2} p_2(r|a_1) - (1 + \frac{a_1^2}{2}) p_1(r|a_1) \right) \log I_0(a_1 r) dr \tag{B.37}$$

$$- \int_0^\infty \left(p_N(r|0) - (1 + \frac{a_1^2}{2}) p_N(r|a_1) + \frac{a_1^2}{2} p_{N+1}(r|a_2) \right) \log \frac{p_N(r|a_1)}{r^{2N-1}} dr. \tag{B.38}$$

We first show that the integral in (B.38) is negative. For that purpose let us analyze the expression inside the brackets in (B.38), call it $f(r, a_1)$. Observe that for any $a_1 > 0$, there exists two positive numbers R_1 and R_2, so that $f(r, a_1) > 0$ for $r < R_1$ and $r > R_2$, and $f(r, a_1) < 0$ for $R_1 < r < R_2$. Furthermore, using the fact that $p_N(r|a)$ is a density function with the second moment given by (B.29), one can see that for any a, $f(r, a)$ satisfies

$$\int_0^\infty f(r, a) dr = 0 \tag{B.39}$$

$$\int_0^\infty f(r, a) r^2 dr = 0. \tag{B.40}$$

Therefore, the integral in (B.38) is equal to

$$\int_0^\infty f(r, a_1) \log \frac{p_N(r|a_1)}{r^{2N-1}} dr = \int_0^\infty f(r, a_1) \log S_{N-1}(a_1 r) e^{b + cr^2} dr \tag{B.41}$$

for any constants b and c. By realizing that $S_{N-1}(ar)e^{b+cr^2}$ is increasing up to some value of r and decreasing afterwards if $c < 0$, and choosing

$$b = \frac{R_2^2 \log S_{N-1}(aR_1) - R_1^2 \log S_{N-1}(aR_2)}{R_2^2 - R_1^2}$$

$$c = \frac{1}{R_2^2 - R_1^2} \log \frac{S_{N-1}(aR_1)}{S_{N-1}(aR_2)} < 0$$

we see that the sign of $\log S_{N-1}(a_1 r)e^{b+cr^2}$ is the same as the sign of $-f(r, a_1)$ for all values of r. Therefore, the integrand in the second integral in (B.41) is non-positive[5] for all values of r, which proves that the integral in (B.38) is negative.

We now show that the expression in (B.37) is positive. Rewrite this expression as

$$\frac{a_1^2}{2} - \int_0^\infty p_1(r|0) \log I_0(a_1 r) dr \tag{B.42}$$

$$+ \int_0^\infty \left(p_1(r|0) - (1 + \frac{a_1^2}{2}) p_1(r|a_1) + \frac{a_1^2}{2} p_2(r|a_1) \right) \log I_0(a_1 r) dr \tag{B.43}$$

and observe that the expression inside brackets in (B.43) has the same properties as the function $f(r, a_1)$ defined above (cf. (B.38)). Using the fact that the function

$$\log(I_0(ar)) - a^2 \frac{p_3(r|a_1)}{p_2(r|a_1)} = \log(I_0(ar)) - ar \frac{I_2(ar)}{I_1(ar)} \tag{B.44}$$

is negative and decreasing in r for any a, we can see that replacing $\log(I_0(ar))$ by $a^2 \frac{p_3(r|a_1)}{p_2(r|a_1)}$ in (B.42) and (B.43) will only decrease the resulting expression. Therefore, we have that

$$\frac{d}{dx} I(x)\big|_{x=0} \geq \frac{a_1^2}{2} + \int_0^\infty \left(\frac{a_1^2}{2} p_2(r|a_1) - (1 + \frac{a_1^2}{2}) p_1(r|a_1) \right) a_1^2 \frac{p_3(r|a_1)}{p_2(r|a_1)} dr \tag{B.45}$$

$$= a_1^2 (1 + \frac{a_1^2}{2}) \int_0^\infty (p_2(r|a_1) - p_1(r|a_1)) \frac{p_3(r|a_1)}{p_2(r|a_1)} dr - \frac{a_1^2}{2}. \tag{B.46}$$

Finally using the bound

$$\int_0^\infty (p_2(r|a_1) - p_1(r|a_1)) \frac{p_3(r|a_1)}{p_2(r|a_1)} dr \geq \frac{1}{a_1^2 + 2} \tag{B.47}$$

we get that (B.37) is positive as well, proving the case of a single mass point.

---

[5]The only values of r at which this expression is zero are 0, R_1, and R_2. Therefore, the integral is strictly negative.

Now we deal with the general case when the original distribution has more than one mass point. In this case the scheme similar to (B.34) is utilized

$$\mathbf{p}^*(x) = \{x, p_1 - x, p_2, \ldots\} \tag{B.48a}$$

$$\mathbf{a}^*(x) = \{0, a_1\sqrt{\frac{p_1}{p_1 - x}}, a_2 \ldots\}, \tag{B.48b}$$

which results in

$$\frac{d}{dx} I(\mathbf{a}^*(x), \mathbf{p}^*(x))\big|_{x=0} =$$

$$\frac{a_1^2}{2} + \int_0^\infty \left(\frac{a_1^2}{2} p_2(r|a_1) - (1 + \frac{a_1^2}{2}) p_1(r|a_1)\right) \log I_0(a_1 r) dr \tag{B.49}$$

$$- \int_0^\infty \left(p_N(r|0) - (1 + \frac{a_1^2}{2}) p_N(r|a_1) + \frac{a_1^2}{2} p_{N+1}(r|a_2)\right) \log \frac{p_N(r|a_1)}{r^{2N-1}} dr. \tag{B.50}$$

$$- \int_0^\infty \left(p_N(r|0) - (1 + \frac{a_1^2}{2}) p_N(r|a_1) + \frac{a_1^2}{2} p_{N+1}(r|a_2)\right) \log \frac{p_N(r)}{p_N(r|a_1)} dr. \tag{B.51}$$

The expressions in (B.49) and (B.50) are exactly the same as (B.37) and (B.38), respectively. The integral in (B.51) can be positive, however, by calling the expression inside the brackets in (B.51) $g(r, a_1)$ we have

$$\int_0^\infty g(r, a_1) \log \frac{p_N(r)}{p_N(r|a_1)} dr \le \int_0^\infty g(r, a_1) \log\big(1 + A \frac{p_N(r|a')}{p_N(r|a_1)}\big) dr \tag{B.52}$$

$$\le \sqrt{\log A} \int_0^\infty g(r, a_1) \log \frac{1}{r} dr, \tag{B.53}$$

where $A = \frac{1-p_1}{p_1}$, and $a' = a_1 + \sqrt{\log A}$. For $\log A < N$ the expression in (B.53) is less than the total value in (B.49) and (B.50). On the other hand, when left hand side of (B.52) is bigger than the total value in (B.49) and (B.50), we have that

$$\frac{1}{a_1} \frac{\partial I}{\partial a_1} < \frac{1}{a_k} \frac{\partial I}{\partial a_k} \quad \text{for some } k, \tag{B.54}$$

which implies that the scheme

$$\mathbf{p}'(x) = \mathbf{p} \tag{B.55a}$$

$$\mathbf{a}'(x) = \{0, \sqrt{a_1^2 - \frac{x}{p_1}}, a_2, \ldots, a_{k-1}, \sqrt{a_k^2 + \frac{x}{p_k}}, a_{k+1}, \ldots\}, \tag{B.55b}$$

improves mutual information as can be observed from

$$\frac{d}{dx} I(\mathbf{a}'(x), \mathbf{p}'(x))\big|_{x=0} = \frac{1}{a_k}\frac{\partial I}{\partial a_k} - \frac{1}{a_1}\frac{\partial I}{\partial a_1} > 0. \tag{B.56}$$

Therefore, in this pathological case the scheme in (B.55) should be followed before applying the general scheme of (B.48). It should also be noted that when the smallest amplitude satisfies $a_1 > N-1$, then simply introducing a mass point at zero, without increasing a_1 (*i.e.*, using only (B.48a) and not (B.48b)), increases mutual information.

APPENDIX C

POWERFUL ERROR-CORRECTING CODES

As we have seen in the code design problem, binary error-correcting codes play a major role in the overall coding scheme. In this section, we give a brief introduction to widely used binary error-correcting codes, that are utilized throughout the chapters.

Recently, several practical coding schemes have been developed for the coherent AWGN channel, that come very close to its information-theoretic limit. Of these, we briefly mention the two most powerful ones, namely, (i) concatenated schemes, collectively known as turbo codes, and (ii) the recently rediscovered low-density parity-check (LDPC) codes. Turbo codes, first introduced in [10], use relatively short block or convolutional codes in a concatenated manner (either in parallel or serially) together with large pseudorandom interleavers to effectively obtain a large memory span. Although the theoretical performance of turbo codes is very close to the capacity of the AWGN channel, it is the existence of practical suboptimal powerful iterative decoders that makes them attractive. Several extensions to turbo codes have been suggested with most notable being the serially concatenated turbo codes (the original turbo codes were parallel concatenated) and turbo trellis coded modulation (TCM) schemes [4].

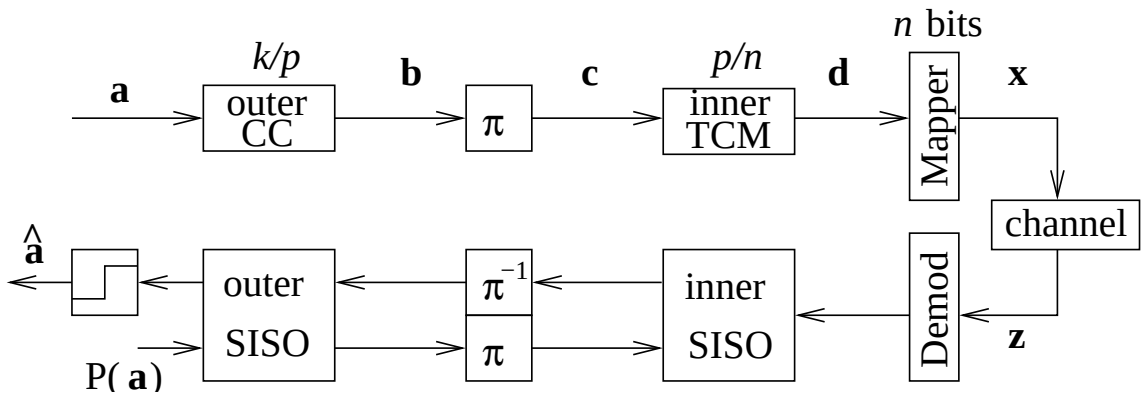


Figure C.1: SCTCM encoder and iterative decoder structure

In this appendix we discuss two coding schemes in detail, namely, serially concatenated TCM codes (SCTCM), and LDPC codes. Their general structure is presented, which is followed by the discussion of the design and analysis tools available for each coding scheme.

C.1 Serially Concatenated Turbo Codes

A typical SCTCM, shown in Fig. C.1, consists of a rate k/p outer convolutional code (CC), the output of which is fed into an inner TCM code of rate p/n, after symbol-wise or bit-wise interleaving (denoted by π). These n bits are then mapped onto a transmitted symbol, resulting in an overall throughput of k bits/symbol. For a length Lp (bits) interleaver the overall code can be thought of as an equivalent block code which takes Lk bits and outputs L symbols. The number of states in the equivalent code is large and thus the complexity of the maximum likelihood sequence detection (MLSD) decoder is very high. However, a practical sub-optimal decoder for such a code is the iterative decoder [4], which consists of two soft-input soft-output (SISO) modules [6] and an interleaver/deinterleaver pair, as shown in Fig. C.1. The two SISO modules evaluate the bit likelihoods and exchange them in an iterative

fashion to converge to the final decision after several iterations. Each SISO module has complexity proportional to the number of states of the corresponding constituent code.

rewrite below to make clear. explain exponential thing. The important performance figure of the turbo codes is the interleaving gain. This gain is obtained as a result of having negative powers of the interleaver size as coefficients in front of pairwise probability of error terms in the union bound expression. The pairwise probability of error is, in turn, exponentially decreasing in the squared output distance. It can be shown that for large values of interleaver size, L, and for a recursive inner code, the dominant error event in the inner code is the one that consists of concatenation of simple error events all corresponding to input Hamming distance of two bits. Furthermore, for bit-error probability the coefficient in front of such term is $O(L^{-\beta})$, where $\beta = \lfloor (d_o + 1)/2 \rfloor$, with d_o being the minimum Hamming distance of the outer code. Based on these the following design guidelines are implied

1. Make the inner code recursive

2. Maximize the effective free distance, $d^2_{f,\text{eff}}$, of the inner code. This parameter is defined as the minimum output distance between inner-code input sequences that are Hamming distance 2 apart, *i.e.*,

$$d_{f,\text{eff}} = \min_{d^H(\mathbf{c},\mathbf{c}')=2} d(\mathbf{x}(\mathbf{c}), \mathbf{x}(\mathbf{c}')). \tag{C.1}$$

3. Maximize the minimum Hamming distance of the outer code.

For the coherent AWGN channel the output distance is taken to be the Euclidean distance, while for the noncoherent channel, that distance can be taken as $\sqrt{1-\rho}$, where ρ is the magnitude of cross correlation.

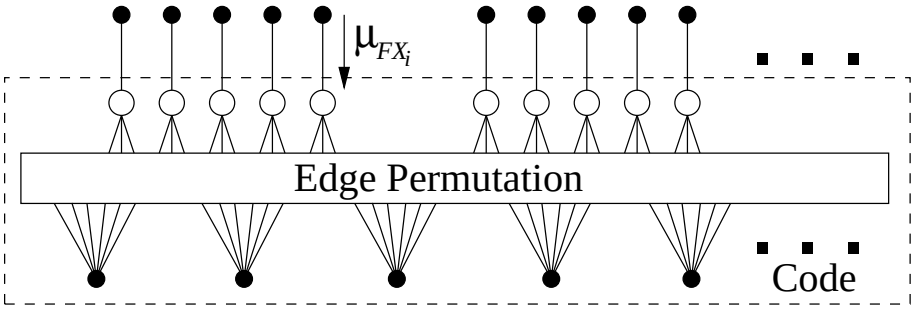


Figure C.2: Factor graph corresponding to the regular LDPC code matched with memoryless channel.

C.2 Low-Density Parity-Check Codes

LDPC codes, introduced as early as 1962 in [29, 30], are large binary linear codes with sparse parity check matrices. These codes have not received much attention until recently, when it was established that their performance is comparable to that of turbo codes [51].

Due to the sparse structure of the parity check matrix, LDPC codes can be decoded with practical complexity using a suboptimal iterative decoding based on factor graphs. Furthermore, a newly developed graph-based technique, called density evolution, enables the precise analysis of the LDPC codes; this technique is also very useful from code design point view.

Because of these properties LDPC codes are represented by factor graphs. A factor graph consists of two kind of nodes, variable and functional nodes. These nodes are connected by the edges, with the property that no two nodes of the same kind are connected by an edge. A generic factor graph corresponding to a binary LDPC code is given in Fig. C.2. Here empty circles (variable nodes) represent the binary variables that make up the codeword, and the bottom filled circles (functional nodes) represent the parity-checks in the code. Every parity check is connected to the variables whose parity it controls. The sparse structure of the parity check matrix is guaranteed

by restricting a number of edges connected to every node. The top filled circles connected to each variable node are used for decoding purposes, and represent the function nodes that give the reliability at the receiver of the corresponding variable nodes given the channel output.

A given word is a valid codeword if all the parity checks are satisfied. The suboptimal decoding algorithm is a message passing algorithm called Sum-Product Algorithm. The nodes in the graph iteratively update and exchange the messages along the edges with the neighboring nodes. A typical message represents the log-likelihood of the particular variable. The algorithm initializes messages with log-likelihoods obtained from the channel output, denoted by μ_{FX_i} in Fig. C.2. The messages exchanged by the sum-product algorithm at the variable and check nodes are given, in the logarithmic domain, by (see [77])

$$\mu_{XC} = \sum_{i=1}^{d_v-1} \mu_{C_iX} + \mu_{FX} \tag{C.2}$$

$$\mu_{CX} = 2\tanh^{-1}\left(\prod_{j=1}^{d_c-1} \tanh\left(\frac{\mu_{X_jC}}{2}\right)\right), \tag{C.3}$$

where μ_{XC} and μ_{CX} represent variable-to-check and check-to-variable node messages, with d_v+1 and d_c being the number of edges connected to variable and check nodes, respectively, and $\{C_i\}_{i=1}^{d_v-1}$ ($\{X_j\}_{j=1}^{d_c-1}$) is the set of check (variable) nodes connected to variable node X (check node C), other than the check node C (variable node X). The algorithm proceeds by iteratively updating messages at the variable and check nodes for a certain number of iterations. The final decisions are made on the combined message at the variable node.

The message update equations above assume the number of edges connected to variable and check nodes is fixed. LDPC codes with such a restriction are called regular. However, it was recently discovered that considerable gains can be obtained

by allowing these to vary. The LDPC codes that allow to have varying variable and check degrees are called irregular. For irregular LDPC codes, the design parameter is the distribution of the variable and check degrees.

A very useful method for design purposes is a recently discovered analysis tool, called density evolution. The idea is to keep track of the message densities that are exchanged by the sum-product algorithm. This way the probability density of a single message is evaluated through the iterations to obtain the density of the final message on which the hard decision will be made. The main theorem that justifies this approach is called Concentration Theorem, which states that with increasing code length and iterations the empirical behaviour of all the exchanged messages converges to the density of a single message propagated through the graph.

The density evolution is mainly used to find an SNR threshold of an LDPC code, that is the value of SNR below which probability of error does not converge to zero with increasing iterations. Using this method irregular LDPC code can be optimized in terms of minimizing the threshold for a given rate. Several optimization algorithms have been devised and successfully used in the literature. In fact, for large blocklengths, the best irregular LDPC codes, optimized using such techniques, outperform turbo codes, and come as close as 0.04 dB to the capacity of the AWGN channel [15].

APPENDIX D

LOWER BOUND ON THE PERFORMANCE OF SCTCM

D.1 Introduction

Serially concatenated (SC) trellis coded modulation (TCM) coding schemes were introduced in [4] as alternatives to the well known "turbo" coding schemes [10, 3], which are parallel concatenated codes. A typical SCTCM, consists of a rate k/p outer convolutional code, the output of which is interleaved[1], and fed into an inner rate p bits/symbol TCM code with 2^n-point constellation, resulting in an overall code rate of k bits/symbol (see Fig. C.1 of Appendix C). For a length N interleaver the overall code can be thought of as an equivalent block code which takes Nk bits and outputs N symbols.

Performance analysis of concatenated codes is usually based on upper bounds on the symbol error probability, averaged over all uniform interleavers [4, 3, 8, 6]. These *random coding* bounds, can be used to provide a proof for the existence of good codes, as well as to aid the design process of SCTCM. On the other hand, the MLSD

[1]The interleaving can be done symbol-wise as well as bit-wise.

performance – let alone the performance of the iterative receiver – of a particular code can be worse than the bound itself. Recently, true, interleaver specific, upper bounds have been derived [46], but the complexity of these bounds is quite substantial. It is noted that lower bounds on the performance of SCTCM, or any other "turbo-like" code, have not been derived, at least to the authors' knowledge.

In this appendix we derive a simple, interleaver independent lower bound that applies to a certain class of SC codes. Specifically, a symbol-wise interleaver is required, and the inner code is required to have parallel transitions, for the bound to apply. The latter is a quite reasonable assumption, especially in the regime of low-complexity inner codes (*e.g.*, 2 to 4 states) and high rate inner codes (*e.g.*, 8-ary, 16-ary, or 32-ary inner-code input alphabets). Although the suggested bound does not predict the well known "interleaver gain" [4, 3], it provides a valuable design tool. In particular, by utilizing the information provided by the bound, we are able to improve existing SCTCM codes with only moderate complexity increase. A series of design examples and simulations confirms our analysis and provides improved SCTCM schemes with moderate complexity.

D.2 Symbol Error Probability Lower Bound for SCTCM

Let $\mathbf{a} = (a_1, a_2, a_3, \ldots a_N)$ be the input sequence[2] to the encoder, with symbols from an M-ary input alphabet $\mathcal{A}$, and let $\mathbf{x}(\mathbf{a}) = (x_1, x_2, x_3, \ldots x_N)$ be the corresponding output of the modulator. Likewise, let $\mathbf{z}$ be the output of the channel at the receiver input:

$$\mathbf{z} = \mathbf{x}(\mathbf{a}) + \mathbf{w} \tag{D.1}$$

[2]In the following all the bold characters denote a sequence.

where $\mathbf{w}$ is an i.i.d Gaussian noise sequence, with variance σ^2 per dimension. We are interested in finding a lower bound for the symbol error probability of the maximum aposteriori symbol detection (MAPSyD) rule. The MAPSyD receiver minimizes the probability of symbol error (for the ith symbol)

$$P_{S,i}(\mathcal{E}) = P(a_i \neq \hat{a}_i) \tag{D.2}$$

where $\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2, \ldots \hat{a}_N)$ is the receiver output sequence, and $\hat{a}_i$ is the corresponding ith symbol. Note that since the MAPSyD receiver is the optimal receiver in the sense of yielding the smallest symbol error probability, a lower bound on (D.2) is a valid lower bound for other receivers as well, *e.g.*, for the maximum likelihood sequence detection (MLSD) receiver.

Since the purpose of this appendix is not the derivation of the bound itself, but rather its application to SCTCM design, we state the bound without proof and refer the interested reader to the relevant literature. Specifically, it can be shown that for equally likely input sequences the symbol error probability can be lower bounded by

$$P_{S,i}(\mathcal{E}) \geq \max_{d} P(A'_{d,i}) Q\left(\frac{d}{2\sigma}\right) \tag{D.3}$$

where $A'_{d,i} \subset A_{d,i}$ and $A_{d,i}$ is defined as the set of all sequences $\mathbf{a}'$ that have at least one neighboring sequence $\mathbf{a}$ with Euclidean distance $||\mathbf{x}(\mathbf{a}) - \mathbf{x}(\mathbf{a}')|| = d$, and $a_i \neq a'_i$. Several observations are in order at this point regarding the selection of the set $A'_{d,i}$.

In the original derivation [26] (for the similar problem of sequence detection in intersymbol interference (ISI) channels), it was assumed that $A'_{d,i} = A_{d,i}$. Problems with this selection arise from the use of a receiver aided by a "genie." As pointed out in [87], such reasoning is flawed if the side-information provides some bias to the aided receiver before observing the channel output. In fact, it was recently demonstrated in [14] (by means of a simple counter-example) that the lower bound

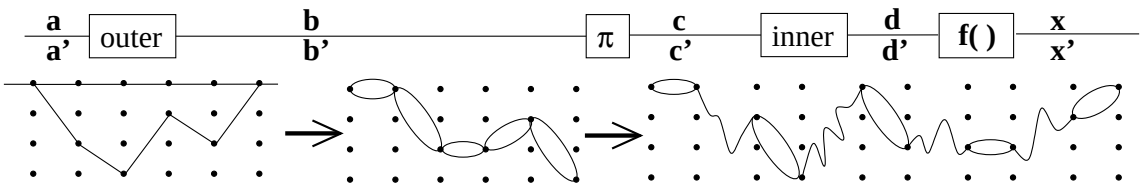


Figure D.1: Parallel transitions in SCTCM. The intermediate trellis corresponds to the inner code with input sequences $\mathbf{b}$ and $\mathbf{b}'$ (without interleaving).

in (D.3) is invalid if $A'_{d,i} = A_{d,i}$. Furthermore, a valid bound in the form of (D.3) can be derived by appropriately stripping the set $A_{d,i}$, such that the remaining set $A'_{d,i}$ satisfies an additional property, which was called "uniform side information" (USI) property [14], and represents a generalization of the looser, but valid, lower bound derived in [54].

The problem with applying the bound in (D.3) arises when one tries to find sequence pairs at a distance d, as well as the corresponding sets $A'_{d,i}$ for an SCTCM scheme with an interleaver size on the order of thousand symbols. Furthermore, this procedure should be repeated every time the interleaver is changed. The main idea about the proposed bound is illustrated in Fig. D.1. Consider an input sequence $\mathbf{a}$ and the corresponding outer codeword $\mathbf{b}$, interleaved codeword $\mathbf{c}$, and inner codeword $\mathbf{x}$. Also consider an input sequence $\mathbf{a}'$, and the corresponding sequences $\mathbf{b}', \mathbf{c}', \mathbf{x}'$. The sequences $\mathbf{a}$ and $\mathbf{a}'$ are such that their corresponding inner codewords $\mathbf{x}$ and $\mathbf{x}'$ represent sequences in the inner trellis having only parallel transitions, that is differing symbols between $\mathbf{b}$ and $\mathbf{b}'$ result in parallel transitions in the inner code. Consequently, permuting the outer codeword will not alter the output distance between $\mathbf{x}$ and $\mathbf{x}'$. Observe that this result holds for *any* interleaver, as mentioned earlier. Thus, calculating the lower bound reduces to finding pairs of input sequences $\mathbf{a}$, $\mathbf{a}'$ that result in inner code sequences with parallel transitions.

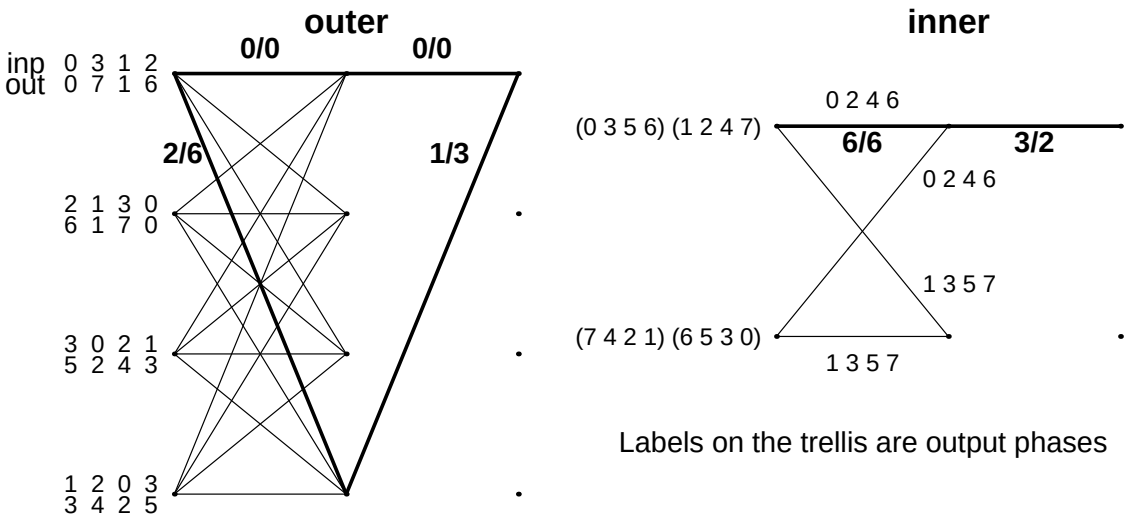


Figure D.2: Code A.1 description and the error sequence used in the lower bound.

It is required that b_i and b_i' result in parallel transitions departing from any state in the inner code. This induces the property on the inner code that there is a collection of disjoint sets of input symbols to the inner code, where any two symbols in the same set result in the parallel transitions regardless of the initial state. This is true in two important cases: 1) A TCM with a linear encoder (*i.e.*, the mapping from $\mathbf{c}$ to $\mathbf{d}$ is linear), and 2) for all codes designed following the rules in [4, 6, 86]. We denote this collection of inner code input symbols by

$$\mathcal{C} = \{L_0, L_1, \ldots L_m\} \tag{D.4}$$

The problem now reduces to finding pairs of codewords of outer code whose disagreeing symbols b_i and b_i' fall into the same set L_{k_i}. An example of such a partition is shown in Fig. D.2, for which the inner code has 2 states and the obtained partition consists of the sets $L_0 = \{0, 3, 5, 6\}$ and $L_1 = \{1, 2, 4, 7\}$.

A great amount of simplification occurs when both encoders are linear and the mapping $f(\cdot)$ of the inner encoded bits d to constellation points x possesses the "distance linearity" property, that is: $d(f(d), f(d \oplus d'))$ is independent of d for all

encoder output symbols d and d' such that $f(d)$ and $f(d \oplus d')$ fall within the same sub-constellation in the set partitioning of the TCM constellation. This property is, for example, true for 8-PSK and 16-QAM constellations with natural mapping and Ungerboeck set partitioning. In this case we only need to perform the above described procedure for the all zero input sequence, and only need to find one sequence at a distance d, say $\mathbf{e}$ to obtain $P(A'_{d,i}) = 1$.

Similar arguments can be used to bound the probability of bit error. It is desirable, however, to derive a lower bound for the bit error probability of the optimal MAP bit detector using *exactly* the same side information revealing mechanism used in deriving the symbol error probability. It can be shown that we can always bound the performance of MAP bit detector with the same bound found for MAPSyD except for a constant factor $1/k$. Moreover, by appropriately shifting the error sequence $\mathbf{e}$ (symbol-wise), it is sometimes possible to increase the factor $1/k$.

We conclude by emphasizing that aside from its value as an analytical tool, the proposed lower bound can be a useful design tool. Specifically, if a given SCTCM code is modified such that the conditions for evaluating the bound do not hold, then it is possible that the resulting code is more powerful. This process was successfully applied and is demonstrated in the next Section.

D.3 SCTCM Design Examples and Numerical Results

The first code is a construction described in [20] and uses a rate 2/3, 4 state outer convolutional code with a 2 state, rate 1 (3/3) inner code and 8-PSK constellation, resulting in a total rate of 2 bits per 8-PSK symbol. The code description is shown in Fig. D.2.

Both encoders are linear, and although the mapping is not natural, the uti-

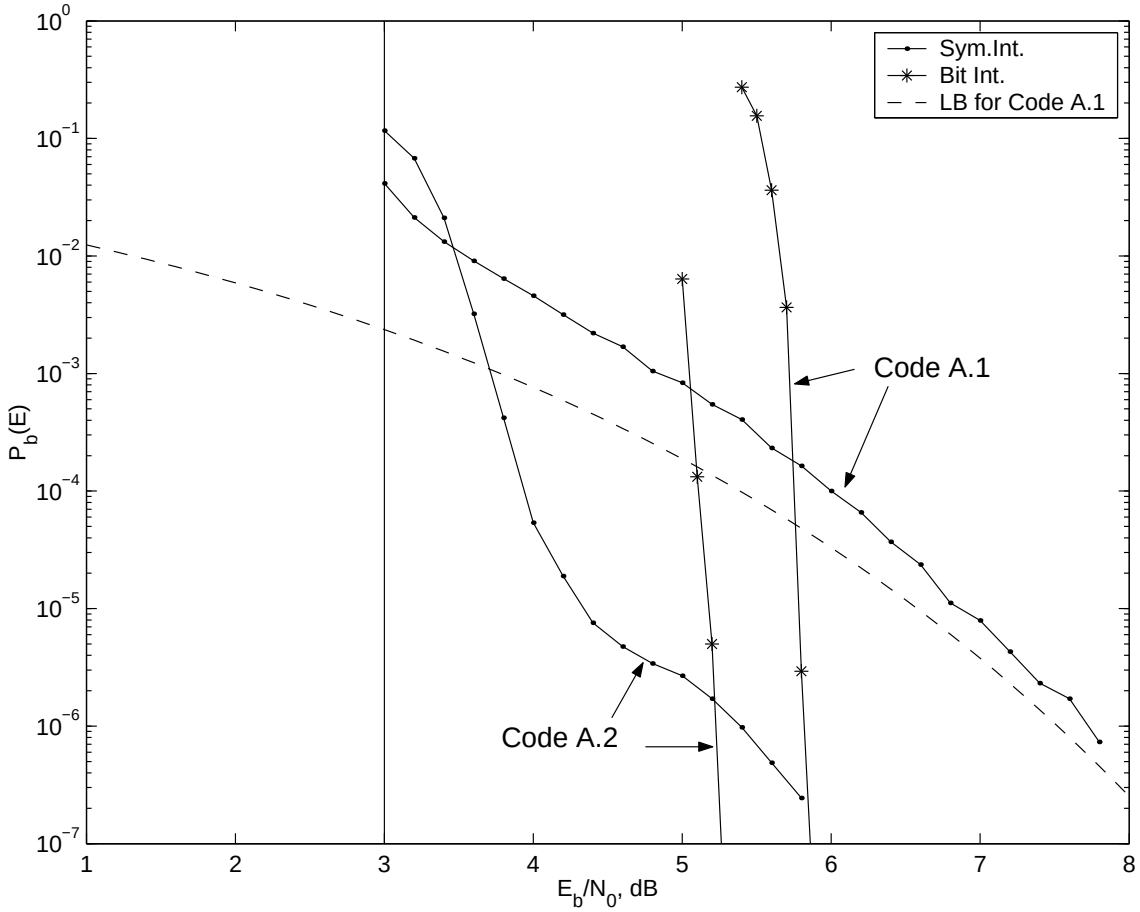


Figure D.3: Simulation Results for Codes A.1 and A.2 ($2/3 \times 3/3 \rightarrow$ 8PSK).

lized mapping does possess a "distance linearity" property, which results in considerable simplification. The partitioning of the input alphabet is $\mathcal{C} = \{L_0 = \{000, 011, 101, 110\}, L_1 = \{001, 010, 100, 111\}\}$, which consists of the set of codewords of the (3,2,1) parity check code and its coset. We will only need L_0 for our calculations, for which the involved nonzero squared Euclidean distances are $d^2 = 2$ and $d^2 = 4$. Calculating the bound reduces to finding a codeword $\mathbf{b}$ of the outer code with symbols from L_0 and having the smallest output Euclidean distance to the all zero sequence. A straightforward lookup quickly reveals one such sequence whose only nonzero symbols are $b_i = 6(110)$ and $b_{i+1} = 3(011)$ which is the re-

sult of encoding the input sequence $\{0\ldots0\ 10\ 01\ 0\ldots0\}$. After interleaving and inner encoding the sequence $\mathbf{b} = \{0\ldots0\ 110\ 011\ 0\ldots0\}$ we get an output binary sequence which has only two nonzero terms 4(100) and 2(010) which are mapped onto phases 2 and 6 of the 8-PSK constellation,[3] and yield a total output squared distance of $d^2 = d^2(0,2) + d^2(0,6) = 4$. The obtained lower bound for this code is $Q(\sqrt{4E_b/N_0})$. It is noted that the above bound is also valid for the average bit error rate. Simulations, shown in Fig. D.3 confirm our analysis. In the same plot, the simulated performance of the SCTCM scheme using bit-wise, instead of symbol-wise, interleaving is also shown for comparison.

As an example of how the lower bound can be utilized in improving code performance, we constructed Code A.2 from Code A.1 by utilizing a 4 state inner code that eliminates the parallel transitions enabling the derivation of the bound. The resulting performance is shown in Fig. D.3, where a gain of almost 3 dB at 10^{-5} compared to the symbol-wise interleaved Code A.1. Furthermore, Code A.2 outperforms both bit-wise interleaved codes A.1 and A.2 by 1.5 dB and 1 dB, respectively.

The second code (Code B.1) consists of a rate 2/3, 16 state outer convolutional code with a rate 3/4, 2 state inner encoder followed by mapping onto 16-QAM constellation (as designed in [4]). The obtained bound for bit error rate is $0.5Q(\sqrt{6.4E_b/N_0})$. This bound was utilized to design the 4 state inner code which results in the Code B.2 with performance gain of more than 2 dB at 10^{-5} (symbol-wise interleaving). Observe in this case, that bit-wise interleaving of the original Code B.1 results in the best performance.

As a last example, consider Codes 3 and 3.1 of Chapter V. Both these codes

---

[3]Symbol i corresponds to the 8PSK constellation point $\exp(j2\pi i/8)$.

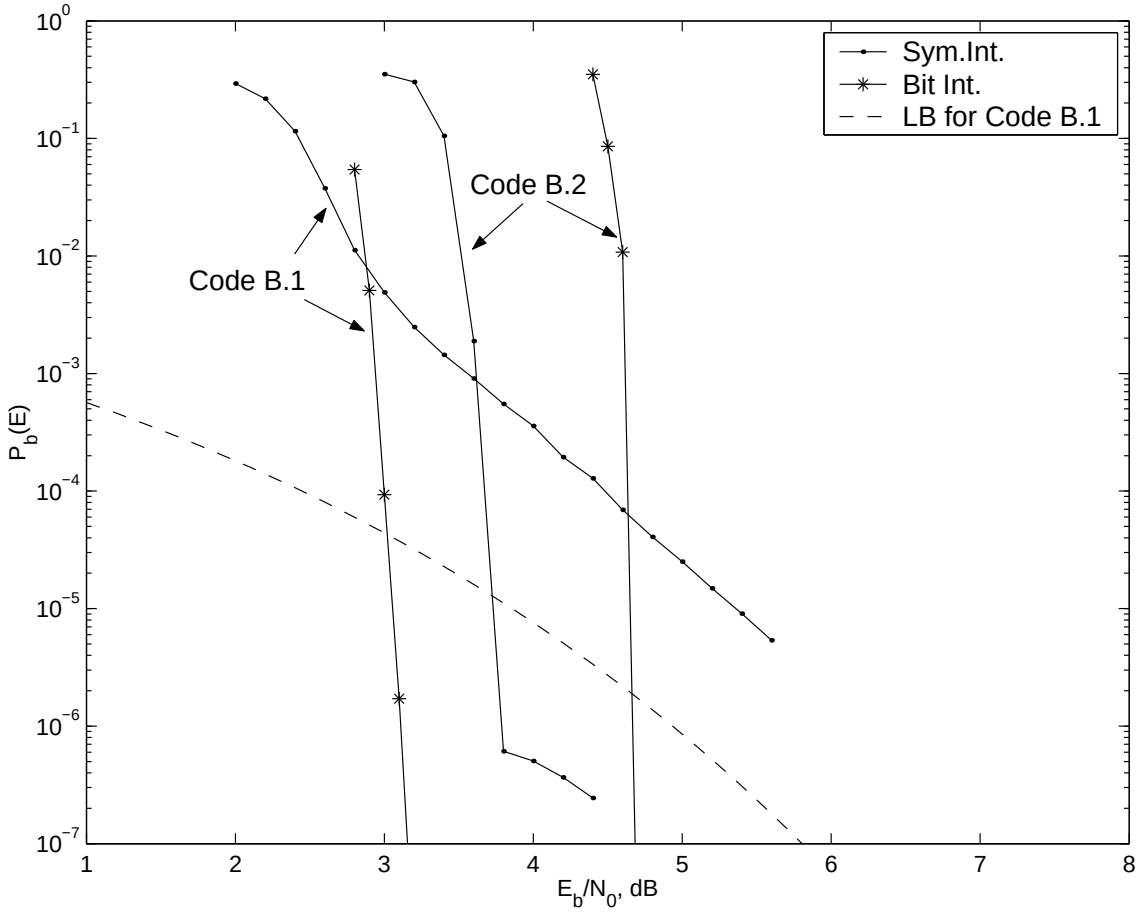


Figure D.4: Simulation Results for Codes B.1 and B.2 ($2/3 \times 3/4 \rightarrow$ 16QAM).

are rate 3 codes over 16-QAM constellation, which consist of a rate 3/4 16-state outer convolutional code followed by a rate 1 4-state inner convolutional code. The corresponding BER curves are presented in Fig. 5.5. The obtained lower bound on BER for both codes is $Q(\sqrt{4.8E_b/N_0})$, which is shown on the graph as a dashed line. It can be seen from the graph that the lower bound accurately predicts the knee effect demonstrated by the symbolwise interleaved versions of these codes, thus suggesting that the errors due to the parallel transitions dominate the probability of error in the high SNR region.

D.4 Conclusion

In this appendix, a simple interleaver independent lower bound that applies to a certain class of SCTCM codes was derived. The validity of this bound was verified via numerous examples. The result found herein were partially used in Chapter V. This lower bound implies useful design criteria, which were utilized to derive improved SCTCM codes. An extensive comparison between bit-wise and symbol-wise interleaving (not reported here) showed that the choice between the two options depends on the target BER and operating SNR region. For most of the codes, and for moderate bit error rates symbol-wise interleaving is preferred, whereas at very low desired error rates bit-wise interleaving is the choice of preference.

Derivation of lower bounds that predict the interleaving gain, is clearly an exciting future research direction. Such generalization seems feasible, due to the generality of the developed underlying theoretical framework for the lower bound in (D.3).

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] A. Anastasopoulos and K. M. Chugg. Adaptive iterative detection for phase tracking in turbo coded systems. *IEEE Trans. Communications*, 49(12):2135–2144, December 2001.

[2] A. Ben-Zur and D. Raphaeli. Noncoherent trellis coded amplitude-phase modulation. In *Proc. Globecom Conf.*, pages 2518–2522, Rio de Janeiro, Brazil, December 1999.

[3] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. Parallel concatenated trellis coded modulation. In *Proc. International Conf. Communications*, pages 974–978, Dallas, Texas, June 1996.

[4] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. Serial concatenated trellis coded modulation with iterative decoding. In *Proc. International Symposium on Information Theory*, page 8, Ulm, Germany, June 1997.

[5] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding. *IEEE Trans. Information Theory*, 44(3):909–926, May 1998.

[6] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. Soft-input soft-output modules for the construction and distributed iterative decoding of code networks. *European Trans. Telecommun.*, 9(2):155–172, March/April 1998.

[7] S. Benedetto, R. Garello, M. Mondin, and M. D. Trott. Rotational invariance of trellis codes – Part II: Group codes and decoders. *IEEE Trans. Inform. Theory*, 42(3):766–778, May 1996.

[8] S. Benedetto and G. Montorsi. Performance evaluation of turbo codes. *IEE Electronics Letters*, 31(3):163–165, February 1995.

[9] C. Berrou and A. Glavieux. Near optimum error correcting coding and decoding: Turbo-codes. *IEEE Trans. Communications*, 44(10):1261–1271, October 1996.

[10] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting coding and decoding: turbo-codes. In *Proc. International Conf. Communications*, pages 1064–1070, Geneva, Switzerland, May 1993.

[11] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon. *Introduction to Trellis-Coded Modulation with Applications*. Macmillan Publishing Company, New York, N. Y., 1991.

[12] S.A. Butman, I. Bar-David, B. K. Levitt, R. F. Lyon, and M. J. Klass. Design cirteria for noncoherent Gaussian channels with MFSK signaling and coding. *IEEE Trans. Communications*, 24:1078–1088, October 1976.

[13] R.-R. Chen, D. Agrawal, and U. Madhow. Noncoherent detection of factor-graph codes over fading channels. In *Proc. Conference on Information Sciences and Systems (CISS)*, March 2000.

[14] K. M. Chugg and A. Anastasopoulos. On symbol error probability bounds for ISI-like channels. *IEEE Trans. Communications*, 49:1704–1709, October 2001.

[15] S.-Y. Chung, G. D. Forney, Jr., T. J. Richardson, and R. L. Urbanke. On the design of of low-density parity-check codes within 0.0045 dB of the Shannon limit. *IEEE Commun. Lett.*, 5(2):58–60, February 2001.

[16] G. Colavolpe, G. Ferrari, and R. Raheli. Noncoherent iterative (turbo) detection. *IEEE Trans. Communications*, 48(9):1488–1498, September 2000.

[17] G. Colavolpe and R. Raheli. Noncoherent sequence detection. *IEEE Trans. Communications*, 47(9):1376–1385, September 1999.

[18] G. Colavolpe and R. Raheli. Capacity of the noncoherent AWGN channel. *European Trans. Telecommun.*, 12(4):289–296, 2001.

[19] V. A. Ditkin and A. P. Prudnikov. *Integral Transforms and Operational Calculus*. Pergamon Press, 1st edition, 1965.

[20] D. Divsalar, S. Dolinar, and F. Pollara. Serial concatenated trellis coded modulation with rate-1 inner code. In *Proc. Globecom Conf.*, volume 2, pages 777–782, San Francisco, CA, November 2000.

[21] D. Divsalar and M. Simon. Multiple-symbol differential detection of MPSK. *IEEE Trans. Communications*, 38:300–1308, March 1990.

[22] F. Edbauer. Bit error rate of binary and quaternary DPSK signals with multiple differential feedback detection. *IEEE Trans. Commun.*, 40(3):457–460, March 1992.

[23] M. V. Eyuboğlu and G. D. Forney, Jr. Trellis precoding: Combined coding, precoding and shaping for intersymbol interference channels. *IEEE Trans. Information Theory*, 38:301–314, March 1992.

[24] P. Fan and X. Xia. A noncoherent coded modulation for 16QAM. *IEEE Trans. Commun.*, 49(4):260–262, 2001.

[25] I. C. A. Faycal, M. D. Trott, and S. Shamai (Shitz). The capacity of discrete-time memoryless Rayleigh fading channels. *IEEE Trans. Information Theory*, 47(4):1290–1301, May 2001.

[26] G. D. Forney, Jr. Lower bounds on error probability in the presence of large intersymbol interference. *IEEE Trans. Communications*, 20(1):76–77, February 1972.

[27] G. D. Forney, Jr. Trellis shaping. *IEEE Trans. Information Theory*, 38:281–300, March 1992.

[28] B. J. Frey. *Graphical models for machine learning and digital communications*. MIT Press, Cambridge, MA, 1998.

[29] R. G. Gallager. Low density parity check codes. *IEEE Trans. Information Theory*, 8:21–28, January 1962.

[30] R. G. Gallager. *Low-Density Parity-Check Codes*. MIT Press, Cambridge, MA, 1963.

[31] M. Gertsman and J. H. Lodge. Symbol-by-symbol MAP demodulation of CPM and PSK on Rayleigh flat-fading channels. *IEEE Trans. Communications*, 45(7):788–799, July 1997.

[32] B. Hassibi and Y. Jing. Unitary space-time codes and the Cayley transform. In *Proc. ICASSP*, volume 3, pages 2409 –2412, Orlando, FL, May 2002.

[33] M. V. Hedge and W. E. Stark. Capacity of frequency-hop spread-spectrum multiple-access communications systems. *IEEE Trans. Commun.*, 38(7):1050–1059, July 1990.

[34] B. M. Hochwald, T.L. Marzetta, T. J. Richardson, W. Sweldens, and R. L. Urbanke. Systematic design of unitary space–time constellations. *IEEE Trans. Information Theory*, 46(6):1962–1973, September 2000.

[35] J. Hou, P. H. Siegel, and L. B. Milstein. Performance analysis and code optimization of low-density parity-check codes on Raleigh fading channels. *IEEE J. Select. Areas Commun.*, 19:924–934, May 2001.

[36] H. Jin and T. J. Richardson. Design of low-density parity-check codes for non-coherent MPSK communication. In *Proc. International Symposium on Information Theory*, page 169, Lausanne, Switzerland, June 2002.

[37] G. A. Kabatyanskii and V. I. Levenshtein. Bounds for packings on a sphere and in space. *Problems of Information Transmission*, 14(1):1–17, 1978.

[38] M. Katz and S. Shamai. On the capacity-achieving distribution of the discrete-time non-coherent additive Gaussian noise channel. In *Proc. International Symposium on Information Theory*, page 165, Lausanne, Switzerland, June 2002.

[39] M. Katz and S. Shamai. On the capacity-achieving distribution of the discrete-time non-coherent and partially-coherent awgn channels. *IEEE Trans. Information Theory*, September 2002. Submitted for publication. Also in EE Publication No. 1335 (CCIT Report 396), Electrical Engineering, Technion, Haifa, Israel.

[40] R. Knopp and H. Leib. M-ary phase coding for the noncoherent AWGN channel. *IEEE Trans. Information Theory*, 40(6):1968–1984, November 1994.

[41] A. N. Kolmogorov and S. V. Fomin. *Elementy teorii funktsiy i funktsianalnogo analiza (Elements of Theory of Functions and Functional Analysis)*. Nauka, Moscow, 6th edition, 1989. in russian.

[42] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Trans. Information Theory*, 47(2):498–519, February 2001.

[43] L. Lampe and R. Schober. Iterative decision–feedback differential demodulation of bit-interleaved coded MDPSK for flat rayleigh fading channels. *IEEE Trans. Commun.*, 49(7):1176–1187, July 2001.

[44] A. Lapidoth. On phase noise channels at high SNR. In *Proc. Information Theory Workshop*, pages 1–4, Bangalore, India, October 2002.

[45] M. A. Lavrentiev and B. V. Shabat. *Metody teorii funktsiy kompleksnogo peremennogo (Methods of theory of functions of complex variable)*. Nauka, Moscow, 5th edition, 1987. in russian.

[46] I. Lee. The effect of a precoder on serially concatenated systems with ISI channel. In *Proc. International Conf. Communications*, New Orleans, LA, 2000. (Comm. Theory Mini-Symp.).

[47] H. Leib and S. Pasupathy. The phase of a vector perturbed by Gaussian noise and differentially coherent receivers. *IEEE Trans. Inform. Theory*, 34(6):1491–1501, November 1988.

[48] J.M. Liebetreu. Joint carrier phase estimation and data detection algorithms for multi-hop CPM data transmission. *IEEE Trans. Communications*, pages 873–881, September 1986.

[49] W. Liu and S. G. Wilson. Rotationally-invariant concatenated (turbo) TCM codes. In *Proc. Asilomar Conf. Signals, Systems, Comp.*, volume 1, pages 32–36, 1999.

[50] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, 1969.

[51] D. J. C. MacKay and R. M. Neal. Near Shannon limit performance of low density parity check codes. *IEE Electronics Letters*, 32(18):1645–1646, August 1996.

[52] K. M. Mackenthun, Jr. Codes based on a trellis cut set transformation—parti II: codes for noncoherent detection. *IEEE Trans. Communications*, 42:998–1007, July 1999.

[53] T. L. Marzetta and B.M. Hochwald. Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading. *IEEE Trans. Information Theory*, 45:139–157, January 1999.

[54] J. E. Mazo. Faster-than-Nyquist signaling. *Bell System Tech. J.*, 54(8):1451–1462, October 1975.

[55] R. J. McEliece. The effectiveness of turbolike codes on nonstandard channel models. In *Proc. International Symposium on Information Theory*, Washington, D.C., June 2001. Plenary talk.

[56] R. J. McEliece and W. E. Stark. Channels with block interference. *IEEE Trans. Inform. Theory*, 30(1):44–53, January 1984.

[57] I. Motedayen and A. Anastasopoulos. Polynomial-complexity noncoherent symbol-by-symbol detection with application to adaptive iterative decoding of turbo-like codes. *IEEE Trans. Communications*, February 2003. (to appear).

[58] C. R. Nassar and M. R. Soleymani. Optimal data detection of MPSK in the presence of unknown carrier phase at low complexity. In *Proc. Allerton Conf. Commun., Control, Comp.*, September 1994.

[59] R. Nuriyev and A. Anastasopoulos. Design and robustness analysis of rotationally invariant SCTCM. In *Proc. International Conf. Communications*, pages 2226–2230, Helsinki, Finland, June 2001.

[60] R. Nuriyev and A. Anastasopoulos. Analysis and design of pilot-symbol-assisted codes, for the non-coherent AWGN channel, using density evolution. In *Proc. International Conf. Communications*, pages 1511–1515, New York, NY, April 2002.

[61] R. Nuriyev and A. Anastasopoulos. Analysis of joint iterative decoding and phase estimation for the non-coherent AWGN channel, using density evolution. In *Proc. International Symposium on Information Theory*, page 168, Lausanne, Switzerland, June 2002.

[62] R. Nuriyev and A. Anastasopoulos. Capacity and coding for the noncoherent block-independent AWGN channel. *IEEE Trans. Information Theory*, January 2003. (Submitted).

[63] R. Nuriyev and A. Anastasopoulos. Capacity-approaching code design for the noncoherent AWGN channel. In *Proc. Globecom Conf.*, San Francisco, USA, December 2003. (Submitted).

[64] R. Nuriyev and A. Anastasopoulos. Capacity characterization for the noncoherent block-independent AWGN channel. In *Proc. International Symposium on Information Theory*, Yokohama, Japan, June 2003. (Accepted for publication/presentation).

[65] R. Nuriyev and A. Anastasopoulos. Pilot-symbol-assisted coded transmission over the block-noncoherent AWGN channel. *IEEE Trans. Communications*, June 2003. (to appear).

[66] R. Nuriyev and A. Anastasopoulos. Rotationally invariant and rotationally robust codes for the AWGN and the noncoherent channel. *IEEE Trans. Communications*, April 2003. (Accepted for publication).

[67] J. K. Omura and D. Jackson. Cut off rates for using bandwidth efficient modulations. In *Proc. NTC*, pages 14.1.1–14.1.11, 1980.

[68] M. Peleg and S. Shamai (Shitz). On the capacity of the blockwise incoherent MPSK channel. *IEEE Trans. Communications*, 46:603–609, May 1998.

[69] M. Peleg and S. Shamai (Shitz). Efficient communication over memoryless Rayleigh fading channels with Turbo coding/decoding. In *Proc. International Conf. Communications*, pages 83–88, Vancouver, B.C., Canada, 1999. (Comm. Theory Mini-Conf.).

[70] M. Peleg, S. Shamai (Shitz), and S. Galán. Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channel. *IEE Proceedings-Communications*, 147:87–95, April 2000.

[71] H. V. Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 2nd edition, 1994.

[72] J. G. Proakis. *Digital Communications*. McGraw-Hill, New York, 3rd edition, 1995.

[73] A. P. Prudnikov, U. A. Brytzkov, and O. I. Marichev. *Integraly i ryady. Spezialniye funktsii. (Integrals and Series. Special Functions.)*. Nauka, Moscow, 1st edition, 1983. in russian.

[74] R. Raheli, A. Polydoros, and C.-K. Tzou. Per-survivor processing: A general approach to MLSE in uncertain environments. *IEEE Trans. Communications*, 43(2/3/4):354–364, Feb/Mar/Apr. 1995.

[75] D. Raphaeli. Noncoherent coded modulation. *IEEE Trans. Communications*, 44:172–183, February 1996.

[76] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. Information Theory*, 47(2):619–637, February 2001.

[77] T. J. Richardson and R. L. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans. Information Theory*, 47(2):599–618, February 2001.

[78] R. A. Scholtz. Signal design for phase-incoherent communications. *IEEE Trans. Information Theory*, 12:456–463, October 1966.

[79] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, July 1948.

[80] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:623–656, October 1948.

[81] A. Shokrollahi and R. Storn. Design of efficient erasure codes with differential evolution. In *Proc. IEEE Symposium on Information Theory*, page 5, Sorrento, Italy, June 2000.

[82] J. G. Smith. The information capacity of amplitude and variance-constained scalar gaussian channels. *Information and Control*, 18(3):203–219, April 1971.

[83] W. E. Stark and R. J. McEliece. On the capacity of channels with block memory. *IEEE Trans. Inform. Theory*, 34(2):322–324, March 1988.

[84] R. Storn and K. Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359, 1997.

[85] M. D. Trott, S. Benedetto, R. Garello, and M. Mondin. Rotational invariance of trellis codes – Part I: Encoders and precoders. *IEEE Trans. Information Theory*, 42(3):751–765, May 1996.

[86] G. Ungerboeck. Channel coding with multilevel/phase signals. *IEEE Trans. Information Theory*, 28(1):55–67, January 1982.

[87] S. Verdú. *Multiuser Detection*. Cambridge University Press, Cambridge, UK, 1998.

[88] D. Warrier and U. Madhow. Spectrally efficient noncoherent communication. *IEEE Trans. Information Theory*, 48:651–668, March 2002.

[89] L. F. Wei. Rotationaly invariant convolutional channel coding with expanded signal set–Part I: 180°. *IEEE J. Select. Areas Commun.*, SAC-2:659–671, September 1984.

[90] L. F. Wei. Rotationaly invariant convolutional channel coding with expanded signal set–Part II: Nonlinear codes. *IEEE J. Select. Areas Commun.*, SAC-2:672–686, September 1984.

[91] L. F. Wei. Rotationaly invariant trellis coded modulations with multidimensional M-PSK. *IEEE J. Select. Areas Commun.*, 7:1281–1295, December 1989.

[92] R. Wei and M. Lin. Noncoherent-coded modulation constructed from conventional trellis-coded modulation. *IEEE Commun. Lett.*, 2(9):260–262, 1998.

[93] L. R. Welch. Lower bounds on the maximum cross correlation of signals. *IEEE Trans. Inform. Theory*, 20:397–399, May 1974.

[94] N. Wiberg. *Codes and Decoding on General Graphs.* PhD thesis, Linköping University, Linköping, Sweden, 1996.

[95] S. G. Wilson and C. D. Hsu. Joint MAP data/phase sequence estimation for trellis phase codes. In *Proc. International Conf. Communications*, pages 26.1.1–26.1.5, 1980.

[96] A. P. Worthen and W. E. Stark. Unified design of iterative receivers using factor graphs. *IEEE Trans. Information Theory*, 47(2):843–849, February 2001.

ABSTRACT

COMMUNICATION OVER THE NONCOHERENT CHANNEL

by

Rza Nuriyev

Chair: Achilleas Anastasopoulos

Communication over the noncoherent additive white Gaussian noise (AWGN) channel is considered, where the transmitted signal undergoes a phase rotation, unknown to the transmitter and the receiver. The effects of phase dynamics are explicitly taken into account by considering a block-independent model for the phase process.

Two main problems regarding this channel are investigated in this work. The first is the design and analysis of practical powerful codes whose performance is close to the theoretical limit. The second, more theoretical problem, is finding the fundamental limits of communication over this channel.

Code design is initiated with a practical and intuitive coding scheme that uses pilot symbols to facilitate phase estimation and effectively translate the noncoherent channel into the coherent AWGN channel. This coding scheme is analyzed using a recently discovered technique, called density evolution, and the inherent trade-off associated with the pilot-power is quantified.

We consider a theoretical aspect of communication problem by analyzing the information capacity and the structure of the capacity achieving signaling scheme. In particular, the capacity achieving input distribution is characterized; it is shown

that the maximizing density has circular symmetry, is discrete in amplitude with infinite number of mass points and always has a mass point at zero. Furthermore, asymptotic expressions show that the probability of a mass point is decreasing double exponentially with its amplitude. Based on these results, the capacity is evaluated through numerical optimizations for unconstrained and modulation-constrained input distributions.

Inspired by the capacity results, two novel classes of coding and modulation schemes are proposed for fast and moderate phase dynamics. In the case of fast phase dynamics, optimized modulation alphabets are designed in conjunction with simple serially concatenated convolutional codes, and show close-to-capacity performance with reasonable overall complexity. In the case of moderate phase dynamics, specially designed modulation alphabets that have linear complexity with block length, are utilized together with optimized irregular low-density parity-check codes. Simulation results show that these codes can achieve close-to-capacity performance with moderate complexity, and outperform the best known codes so far.