

**LOWER BOUNDS ON THE ERROR
PROBABILITY OF A GIVEN BLOCK CODE**

ASAF COHEN

**LOWER BOUNDS ON THE ERROR PROBABILITY OF
A GIVEN BLOCK CODE**

RESEARCH THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE
IN ELECTRICAL ENGINEERING

ASAF COHEN

SUBMITTED TO THE SENATE OF THE TECHNION — ISRAEL INSTITUTE OF TECHNOLOGY

KISLEV, 5763

HAIFA

NOVEMBER, 2002

THIS RESEARCH THESIS WAS DONE UNDER THE SUPERVISION OF PROF.
NERI MERHAV IN THE FACULTY OF ELECTRICAL ENGINEERING

I wish to express my most sincere gratitude to my supervisor, Prof. Neri Merhav, for his devoted guidance and support throughout this research. Most of all, I wish to thank him for teaching me the principles and methods required in research.

I wish to thank Prof. Simon Litsyn, Prof. Ronny Roth and Dr. Igal Sason for several interesting discussions and fruitful comments.

Finally, I wish to thank my family and my partner and best friend Ayelet Schreiber for their encouragement and support.

THE GENEROUS FINANCIAL HELP OF THE TECHNION IS GRATEFULLY
ACKNOWLEDGED

Contents

Abstract	1
List of Symbols and Abbreviations	3
1 Introduction	5
1.1 Overview	5
1.2 Research Goals	6
1.3 Outline and Main Results	6
2 A New Lower Bound on the Probability of a Union of Events	8
2.1 Introduction	8
2.2 Analysis	9
2.3 Discussion	11
3 A Lower Bound on the Error Probability for Signals in Additive White Gaussian Noise	13
3.1 Introduction	13
3.2 Preliminaries	15
3.3 Analysis	17
3.3.1 New Lower Bounds for Any Signal Set	17
3.3.2 New Lower Bounds for Linear Codes	21
3.3.3 Lower Bounds Depending Only on the Subcode \mathcal{C}_d^*	23

3.3.4	Kounias' Bound	24
3.4	Results	27
3.4.1	Implementation Notes	27
3.4.2	Examples and Numerical Analysis Results	27
4	A Lower Bound on the Error Probability for the Binary Symmetric Channel	31
4.1	Introduction	31
4.2	Analysis	33
4.2.1	New Lower Bounds for Linear Codes	33
4.2.2	Approximations for $\text{deg}(\mathbf{x} \mathbf{c}_0)$	37
4.2.3	Lower Bounds Using the Subcode \mathcal{C}_i^* and the Code's Covering Radius	40
4.3	Results	44
5	Upper Bounds on the Error Exponent	46
5.1	Introduction	46
5.2	Preliminaries	48
5.3	Upper Bounds on the Error Exponent for the Binary Symmetric Channel .	49
5.3.1	Main Results	49
5.3.2	Analysis	51
5.4	Upper Bounds on the Error Exponent for the Additive White Gaussian Noise Channel	59
5.5	Examples	61
6	Discussion and Future Work	66
A	Computation of the Integrals Required for Proposition 3.2	69
A.1	Generalized Pairwise Error Probability Integral	69
A.2	Generalized Triplets Error Probability Integral	71

B Proofs of Propositions 4.1 and 4.5	73
B.1 Proof of Proposition 4.1	73
B.2 Proof of Proposition 4.5	75
C Proofs and computations for Chapter 5	77
C.1 Proof of Proposition 5.2	77
C.2 The minimization of $E_1^{\delta_i}(\delta_l, \delta_m, p)$ over \mathcal{D}_1	78
C.3 The minimization of $E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p)$ over \mathcal{D}_2	79
C.4 Proof of Proposition 5.4	82
C.5 Computation of the exponential rate of $\Psi(\frac{1}{2}, x, x)$	84
References	85
Hebrew Abstract	π

List of Figures

3.1	Bounds on the decoding error probability of BCH(63,24) code, AWGN channel. The new lower bounds <i>norm - whole code</i> and <i>dot product - subcode</i> \mathcal{C}_d^* are shown, together with Kounias' version (\mathcal{C}_d^*) for linear codes (lower bound). For reference, Poltyrev's upper bound and Shannon's and Seguin's lower bounds are also plotted.	29
3.2	Bounds on the decoding error probability of Golay(23,12) code, AWGN channel. The new lower bounds <i>norm - whole code</i> and <i>dot product - subcode</i> \mathcal{C}_d^* are shown. For reference, Poltyrev's upper bound and Shannon's and Seguin's lower bounds are also plotted.	30
4.1	Bounds on the decoding error probability of BCH(63,24) code, BSC. The new bound, based on the approximation given in (4.2.28) is given. For reference, three bounds are plotted: Poltyrev's upper bound, Keren and Litsyn's lower bound and the sphere packing lower bound.	45
5.1	A simple one dimensional illustration of the concept behind the proof of Proposition 5.4.	58
5.2	Bounds on the error exponent, Example 5.6. Top to bottom, at $R = 0.02$, the bounds are: the sphere packing bound, linear programming bound, the new bounds (which diverge for higher rates), the straight line bound and a lower bound on the error exponent.	64

5.3	Bounds on the error exponent, BSC, Example 5.7 - ABV codes and the subcode \mathcal{C}_d^* . The two upmost curves are the discussed bound, with trivial $E_\eta^{\delta_d}$ above and non trivial $E_\eta^{\delta_d}$ below. The horizontal line is the value of $E_B^{\delta_d}$. The lowermost curve is the condition on the code. It is clear that for values of p for which the condition is not satisfied with the trivial $E_\eta^{\delta_d}$, non trivial $E_B^{\delta_d}$ tightens the bound.	65
5.4	Bounds on the error exponent, AWGN, Example 5.8 - ABV codes and the subcode \mathcal{C}_d^* . The two upmost curves are the bound with $a = 0$ above and non trivial a below. The lowermost curve is the sphere packing bound. . . .	65
B.1	Example for $\mathbf{c}_i, \mathbf{c}_j, \mathbf{c}_{j'}$ and \mathbf{x} in order to evaluate the size of the set $V_{i_j'}(u, w+1) \cap V_{i_j^c}(u, w)$. The word \mathbf{x} is divided to six parts, each one includes the possible number of 1's. The inequalities for m and l define the summation bounds in (B.1.5).	74

Abstract

In this work, we propose new lower bounds on the error probability in coded communication when using a maximum likelihood decoder. The bounds are derived by improving on a recent lower bound on the probability of a union of events by de Caen, and applying it to this problem. The derived bound on the probability of a union includes a function to be optimized in order to achieve the tightest bound. Moreover, since the optimal choice of this function is known, though usually, mathematically intractable to use as is when bounds on the decoding error probability are considered, we may identify several approximations for it, each in the spirit of the relevant channel model and the type of the code. Thus, each approximation yields a new lower bound on the error probability. Furthermore, the new bound on the probability of a union includes de Caen's bound as a special case, thus the bounds derived from it can be made at least as good as de Caen-based bounds, as long as the optimization process is carried out over a properly chosen family of functions.

In the first part of the work, we apply the new bound to derive lower bounds on the decoding error probability on the Additive White Gaussian Noise (AWGN) channel. The bounds are further specialized for the case where Binary Phase Shift Keying (BPSK) modulation of a linear code is used. In this case, the only knowledge on the code used is its distance distribution. The second part of the work includes analogous derivations for the Binary Symmetric Channel (BSC). In both the AWGN channel and the BSC, the optimal choice of the optimization function and its behavior are stated. Several approximations for this function are proposed. The usage of these approximations is shown to yield tighter

bounds than the de Caen-base bounds. Namely, the resulting bounds improve on latest bounds appearing in the current literature such as Seguin's bound for linear codes over the AWGN channel and Keren and Litsyn's bound for linear codes over the BSC.

In the third part of the work, bounds on the error exponents of the BSC and AWGN channel, resulting from the new bounds on the error probability, are discussed. It is shown that these bounds may be tighter than the bounds resulting from the de Caen bound.

List of Symbols and Abbreviations

Symbol	Meaning	Page
AWGN	Additive White Gaussian Noise	1
BPSK	Binary Phase Shift Keying	1
BSC	Binary Symmetric Channel	1
SNR	Signal to Noise Ratio	7
$\{A_i\}_{i \in \mathcal{I}}$	A set of events in a probability space (Ω, \mathcal{F}, P)	8
$P(A)$	The probability of the event A	8
$p(x)$	A probability density function	9
R_c	The critical rate	14
\mathbb{R}	The set of real numbers	15
\mathbf{s}_i	A signal for the AWGN channel	15
N_0	One sided spectral density of white noise (AWGN channel)	15
ε_{0i}	An error event in which i has better maximum likelihood metric than 0	15
$Q(x)$	The error function	16
$\Psi(\rho, x, y)$	The bivariate normal distribution	16
E_N	Energy per sent bit	21
E_b	Energy per information bit	21
B_i	The i 'th element of the distance distribution of a code	21
(N, K)	A binary linear code of length N and dimension K	21

\mathbf{c}_i	A binary codeword	21
$w(\mathbf{x})$	The Hamming weight of the binary word \mathbf{x}	21
\mathcal{C}_i^*	The subcode of \mathcal{C} containing the all-zero codeword and all codewords of weight i	23
p	BSC crossover probability	31
$GF(2)$	Galois Field of order 2	32
$d_H(\mathbf{c}_i, \mathbf{c}_j)$	The Hamming distance between the binary codewords \mathbf{c}_i and \mathbf{c}_j	32
\mathcal{S}_i	The support of the codeword \mathbf{c}_i	33
$\mathbf{x}_{\mathcal{M}}$	The subword word of \mathbf{x} whose indexes in \mathbf{x} are defined by \mathcal{M}	33
\mathbb{Z}^+	The set of (strictly) positive integers	34
\mathbb{R}^+	The set of (strictly) positive real numbers	41
$E_B^{\delta_i}$	The exponent of the number of codewords of weight i	48
$H(x)$	The binary entropy function	51
$\mathcal{N}(\mu, \sigma^2)$	A Gaussian distribution of expectation μ and variance σ^2	71

Chapter 1

Introduction

1.1 Overview

Consider the classical case of transmitting one of M equally likely signals over a communication channel. The error probability of the maximum likelihood decoder, which is optimal for this case, is often complicated to evaluate. Thus, to estimate the performance of a given code, lower and upper bound on the decoding error probability are required.

Numerous bounds on the error probability of maximum likelihood decoding can be found in the current literature. The bounds are based on a wide variety of techniques, and are mostly specialized for given channel types and code structures. We include only a brief survey of the relevant bounds in this introduction. A more detailed survey, preceding each chapter, includes a description of the closely related works.

In Chapter 2, where bounds on the probability of a union are discussed, we mainly refer to de Caen's lower bound [1], Kuai, Alajaji and Takahara's lower bound [2], and Kounias' lower bound [3]. The new bound on the probability of a union, derived in this work, is an improvement of de Caen's bound. In Chapters 3 and 4, where bounds on the error probability in the AWGN channel and BSC are derived, we mainly refer to the de-Caen based bounds appearing in the current literature, i.e, Seguin's bound [4] for the AWGN

channel and Keren and Litsyn's bound [5] for the BSC. The bounds are compared to the best known upper and lower bounds, namely, the upper bounds by Poltyrev [6] and the lower bounds by Shannon [7]. When bounds on the error exponent are discussed, we mainly refer to the well known lower bounds by Gallager in [8], and the upper bounds by Shannon, Gallager and Berlekamp in [9], McEliece and Omura's upper bound [10], Litsyn's upper bound [11], and Burnashev's upper bound [12].

1.2 Research Goals

The main goal of this work was to apply a new lower bound on the probability of a union of events, derived herein, to the problem of lower bounding the error probability in coded communication. In [4], Seguin used de Caen's bounds to derive lower bounds on the error probability for linear codes in the AWGN channel. In [5], Keren and Litsyn used the same bound as a basis to derive a bound on the error probability for linear codes in the BSC. Since the bound on the probability of a union, presented here, is an improvement of de Caen's bound, we were interested in comparing the performance of the de Caen-based bounds with the new bounds derived herein. If the new bounds perform better than Seguin's bound for the AWGN channel and Keren and Litsyn's bound for the BSC, as it turns out to be, we may conclude that the new bound on the probability of a union is indeed useful and efficient as a basic structure for deriving tight bounds on the error probability. Finally, we were interested to see whether the new bounds can be exponentially tighter than de Caen-based bounds.

1.3 Outline and Main Results

In Chapter 2, the new bound on the probability of a union is derived. The bound is analyzed and discussed while comparing it to related bounds in the current literature.

In Chapter 3, the bound is applied to the case of uniform signaling¹ over the AWGN channel. The resulting bounds are specialized for BPSK modulation and linear codes. In this case, the only knowledge on the code required is its distance distribution. Numerical analysis results show significant enhancement in performance compared to known bounds in the literature. To the author's knowledge, for medium and high values of the Signal to Noise Ratio (SNR) the bounds are shown to yield the tightest results currently available. Chapter 3 also includes a derivation of a new bound based on Kounias' [3] lower bound on the probability of a union. The resulting bound is very simple and performs well for every SNR (superior to Seguin's bound).

In Chapter 4, the bound is applied to the case of uniform signaling over the BSC. The resulting bounds are, again, specialized for linear codes. Numerical analysis results show enhancement in performance compared to Keren and Litsyn's bound, though in this case the improvement is milder. However, in Chapter 5 it is shown that the new bounds presented in this work may be exponentially tighter than the de Caen-based bounds. A detailed derivation of the upper bounds on the error exponent resulting from the lower bounds on the error probability, as well as some examples, are given therein.

Chapter 6 includes a short discussion and suggestions for future work.

¹I.e., equiprobable signals.

Chapter 2

A New Lower Bound on the Probability of a Union of Events

In this chapter, we derive a new lower bound on the probability of a union of events. We mainly follow the method used by de Caen in [1], however, the new bound includes a function that can be optimized to yield tighter bounds. This bound will stand at the basis of our analysis tools.

2.1 Introduction

Consider a family of events $\{A_i\}_{i \in \mathcal{I}}$ in a probability space (Ω, \mathcal{F}, P) . The probability of the union $P(\cup_{i \in \mathcal{I}} A_i)$ rarely admits a simple form in terms of the events probabilities, $P(A_i)$, or pairwise probabilities, $P(A_i \cap A_j)$, hence its evaluation requires an accurate profile of \mathcal{F} . However, knowledge of the event probabilities and the pairwise probabilities can be used to derive lower or upper bounds on the probability of the union. The current literature includes numerous bounds on the probability of a union of events. The bounds closely related to this work are de Caen's lower bound [1], derived using Cauchy-Schwarz inequality; Kuai, Alajaji and Takahara's lower bound [2], derived by solving a minimization problem with

linear constraints, following the methodology of Dawson and Sankoff in [13]; Kounias' lower bound [3], which uses the method of indicators and Hunter's upper bound [14], which uses a basic equality for the union of sets followed by methods in graph theory.

2.2 Analysis

In this section, we derive the new lower bound on the probability of a union. For convenience, we follow de Caen's notations as they appear in [1]. Let $\{A_i\}_{i \in \mathcal{I}}$ be any finite family of events in a probability space (Ω, \mathcal{F}, P) . For each $x \in \Omega$ define

$$\text{deg}(x) \triangleq |\{i \in \mathcal{I} : x \in A_i\}|. \quad (2.2.1)$$

The new lower bound is given by the following theorem.

Theorem 2.1 *Let $\{A_i\}_{i \in \mathcal{I}}$ be any finite family of events in a probability space (Ω, \mathcal{F}, P) . The probability of the union $P(\cup_{i \in \mathcal{I}} A_i)$ is lower bounded by*

$$P\left(\bigcup_{i \in \mathcal{I}} A_i\right) \geq \sum_{i \in \mathcal{I}} \frac{(\sum_{x \in A_i} p(x) m_i(x))^2}{\sum_{j \in \mathcal{I}} \sum_{x \in A_i \cap A_j} p(x) m_i^2(x)} \quad (2.2.2)$$

where $m_i(x) \geq 0$ is any real function on Ω such that the sums on the right hand side (r.h.s.) of (2.2.2) converge. Equality in (2.2.2) is achieved when

$$m_i(x) = m^*(x) = \frac{1}{\text{deg}(x)}, \quad \forall i \in \mathcal{I}. \quad (2.2.3)$$

Proof. We first consider the case where Ω is finite. Using a simple counting argument, we have

$$P\left(\bigcup_{i \in \mathcal{I}} A_i\right) = \sum_{i \in \mathcal{I}} \sum_{x \in A_i} \frac{p(x)}{\text{deg}(x)}. \quad (2.2.4)$$

Let $m_i(x) \geq 0$ be any real function on Ω . From the Cauchy-Schwarz inequality, it follows that

$$\begin{aligned} \left(\sum_{x \in A_i} \frac{p(x)}{\deg(x)} \right) \left(\sum_{x \in A_i} p(x) m_i^2(x) \deg(x) \right) &\geq \left(\sum_{x \in A_i} \sqrt{\frac{p(x)}{\deg(x)}} \sqrt{p(x) m_i^2(x) \deg(x)} \right)^2 \\ &= \left(\sum_{x \in A_i} p(x) m_i(x) \right)^2, \end{aligned} \quad (2.2.5)$$

provided that the sums in (2.2.5) converge. Therefore, from (2.2.4) and (2.2.5),

$$\begin{aligned} P \left(\bigcup_{i \in \mathcal{I}} A_i \right) &\geq \sum_{i \in \mathcal{I}} \frac{(\sum_{x \in A_i} p(x) m_i(x))^2}{\sum_{x \in A_i} p(x) m_i^2(x) \deg(x)} \\ &= \sum_{i \in \mathcal{I}} \frac{(\sum_{x \in A_i} p(x) m_i(x))^2}{\sum_{j \in \mathcal{I}} \sum_{x \in A_i \cap A_j} p(x) m_i^2(x)}. \end{aligned} \quad (2.2.6)$$

Note that $m_i(x)$ may be different for each i in the sum over all $i \in \mathcal{I}$. However, in order to achieve equality in (2.2.5) (which results in a trivial bound) we should define

$$m_i(x) = \frac{1}{\deg(x)}, \quad \forall i \in \mathcal{I}. \quad (2.2.7)$$

For a general probability space, as noted in [1] and [2], since there are only finitely many A_i 's, the number of Boolean atoms defined by the A_i 's unions and intersections is also finite. By considering each atom as a point in a new probability space, with a probability measurement inferred by P , the general space is reduced to a finite probability space. \square

Definition 2.2 Denote the choice of $m_i(x) \equiv 1$ as the trivial choice of $m_i(x)$.

By choosing the trivial choice for $m_i(x)$, we have

$$P \left(\bigcup_{i \in \mathcal{I}} A_i \right) \geq \sum_{i \in \mathcal{I}} \frac{P(A_i)^2}{\sum_{j \in \mathcal{I}} P(A_i \cap A_j)} \quad (2.2.8)$$

which is de Caen's bound [1]. Thus, de Caen's bound is a special case of the bound suggested in Theorem 2.1.

Remark 2.3 Clearly, de Caen’s bound can be written as

$$P\left(\bigcup_{i \in \mathcal{I}} A_i\right) \geq \sum_{i \in \mathcal{I}} \frac{P(A_i)^2}{P(A_i) + \sum_{j \neq i} P(A_i \cap A_j)}. \quad (2.2.9)$$

The expression $\sum_{j \neq i} P(A_i \cap A_j)$ can be considered as a “second order fitting”, without which, the well know union bound is obtained. Thus, when $\sum_{j \neq i} P(A_i \cap A_j)$ is negligible compared to $P(A_i)$, de Caen’s bound is tight. This remark will be elaborated on in the following chapters, where the limits of negligible channel noise level are discussed.

Remark 2.4 One can show ([15]) that both de Caen’s bound and Kuai, Alajaji and Takahara’s bound are derived by minimizing the expression $\sum_{k=1}^N \frac{1}{k} p_i(k)$, where $p_i(k)$ is a probability distribution on the positive integers $k = 1, 2, \dots, N$. While the first bound is obtained by applying Jensen’s inequality, the latter is obtained by a stronger method, which considers the fact that $p_i(k)$ is a probability distribution on the integers. In this way, however, one can show that Kuai, Alajaji and Takahara’s bound improves on de Caen’s bound by at most $9/8$.

2.3 Discussion

The essence of the bound given in Theorem 2.1 is the ability to choose an appropriate function $m_i(x)$. To define a proper strategy for choosing $m_i(x)$, first note that any constant multiplier of $m_i(x)$ factors out in equation (2.2.2). Hence, $m_i(x)$ should only define an *essence of behavior*, and not necessarily exact values. When seeking such a behavior, we remember that the optimal value of $m_i(x)$ is $1/\deg(x)$. While the function $\deg(x)$ is complex to evaluate, usually requires more than the available information on the sets $\{A_i\}_{i \in \mathcal{I}}$, and leads to a trivial identity in equation (2.2.2), its behavior possesses the guidelines for choosing a competent family of approximations. By requiring that any such family of approximations includes the trivial choice for $m_i(x)$ and optimizing the bound over this family, one can assure that the resulting bound is always at least as tight as de Caen’s. Additionally, it is important to require that the chosen family of approximations is mathematically

endurable so the sums in (2.2.2) are feasible.

It is clear that bound given in Theorem 2.1 does not depend only on the $P(A_i)$'s and $P(A_i \cap A_j)$'s. However, a proper choice of the function $m_i(x)$ may result in the same computational complexity while improving on de Caen's bound [1], achieved by choosing $m_i(x) \equiv 1$. This technique is used in Chapter 3 when bounds for the AWGN channel are discussed. When the computational complexity is of less importance, $m_i(x)$ may be chosen to be constant on *subsets* of the $P(A_i)$'s, yielding a more accurate result with only slightly higher computational complexity. This technique is used in Chapter 4 when bounds for the BSC are discussed.

Chapter 3

A Lower Bound on the Error Probability for Signals in Additive White Gaussian Noise

In this chapter, we apply the bound of Theorem 2.1 to derive new lower bounds on the decoding error probability of the maximum likelihood decoder when uniform signaling over an additive white Gaussian noise channel is used. The bounds are specialized for linear codes and binary phase shift keying modulation. Several suggestions for the optimizing function are discussed, each of which results in a new bound on the error probability.

3.1 Introduction

We consider the case of transmitting one of M equally likely signals $\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{M-1}$ over an AWGN channel. For this case, numerous upper and lower bounds on the error probability were derived. We mention here only several bounds closely related to this work.

In [7] Shannon derived lower and upper bounds on the error probability. We are mainly interested in the lower bound, known as the “sphere packing” bound. Roughly speaking,

the sphere packing bound is derived by noting that the error probability of any code whose codewords lie on a given sphere must be greater than the error probability of a code with the same rate and whose codewords are uniformly distributed over that sphere. The sphere packing bound is known to be exponentially tight for rates higher than a certain critical rate, R_c . For rates lower than R_c several bounds were offered, among them are the minimum distance bound (Shannon, [7]), which considers the error caused by pairs of closest codewords, bounds on the error exponent by Shannon, Gallager and Berlekamp [9] (which will be discussed later in this work), and Swaszek lower bound, [16], derived by reducing the complex decision regions to rectangular slabs. As for upper bounds, the tightest known is due to Poltyrev ([6]). Poltyrev elaborates on previous techniques used by Hughes ([17]) and Berlekamp ([18]), and considers only a subset of the space in which the error probability is relatively precisely evaluated, given this subset. Vectors in the rest of the space are considered erroneous.

The preceding bounds mainly use geometrical arguments in order to evaluate the error probability. An alternative approach is to define the probability of error as a probability of a union of events, and use known bound on this probability. When this method is used, the basic events, whose probabilities are to be evaluated directly, are usually the error events when only two or three codewords are involved, hence their probability evaluation is simple. A recent lower bound which uses this method is the bound given by Seguin [4]. Seguin uses de Caen's lower bound [1] on the probability of a union to derive a lower bound on the error probability for signals transmitted across the AWGN channel. Seguin's technique will be discussed throughout later in this chapter. In [19], Kuai, Alajaji and Takahara derive upper and lower bounds using the same method. Their work includes a bound by the same authors on the probability of a union, [2], together with simple algorithms for Kounias' and Hunter's bounds, [3] and [14], respectively. However, Kuai, Alajaji and Takahara consider uncoded communication and nonuniform signaling.

In this work, we use the new bound on the probability of a union to derive lower bounds

on the error probability. We then specialize the bounds for binary linear codes and BPSK modulation. As in Seguin's work, the resulting bound's dependence on the code is only through its weight enumeration. Nevertheless, we show that the new bound is at least as tight as Seguin's bound for every value of E_b/N_0 (SNR).

3.2 Preliminaries

In this section, we state the basic assumptions relevant to this chapter and introduce the required notations. We mainly follow Seguin's notations as they appear in [4]. For the sake of easy reference, the main results of Seguin's work [4] are also presented.

We consider the case of uniform signaling over an AWGN channel and maximum likelihood decoding. The transmitted signal is one of M equiprobable signals $\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{M-1}$ of length K . A discrete time channel is assumed since whether the signals are continuous time signals of dimension K , or discrete time signals of length N and dimension K , the Gram-Schmidt orthogonalization procedure can be used (see, for example, [20, Section 2.1]). If \mathbf{s}_0 is transmitted, the received signal is $\mathbf{r} = \mathbf{s}_0 + \mathbf{n}$, where \mathbf{n} is a vector of K independent Gaussian random variables¹ with zero mean and variance $\frac{N_0}{2}$ (i.e., the noise double sided spectral density is $\frac{N_0}{2}$). The maximum likelihood decoder, which is optimal for this case, chooses the closest of the M signals to \mathbf{r} , in the Euclidean sense. Thus, the probability of error given that \mathbf{s}_0 was sent is

$$P(\varepsilon|\mathbf{s}_0) = P(\cup_{i \neq 0} \varepsilon_{0i} | \mathbf{s}_0), \quad (3.2.1)$$

where

$$\varepsilon_{0i} \triangleq \{\mathbf{r} \in \mathbb{R}^K : \|\mathbf{r} - \mathbf{s}_i\| < \|\mathbf{r} - \mathbf{s}_0\|\} \quad (3.2.2)$$

and $\|\cdot\|$ is the Euclidean norm. Note the strict inequality in (3.2.2), a consequence of the assumption that ties are solved in favor of the correct signal. Generally speaking, this

¹Again, this means that for continuous channel, only the noise projection on the signals space is considered. The remainder is irrelevant to the decision.

assumption is essential when lower bounds on the error probability are discussed. When a continuous probability space is at hand it is of lesser importance.

To derive bounds on the error probability, we are interested in the following entities. The pairwise error probability, given by

$$P(\varepsilon_{0i}|\mathbf{s}_0) = P(\{\mathbf{r} \in \mathbb{R}^K : \|\mathbf{r} - \mathbf{s}_i\| < \|\mathbf{r} - \mathbf{s}_0\|\}|\mathbf{s}_0) = Q\left(\frac{\|\mathbf{s}_0 - \mathbf{s}_i\|}{\sqrt{2N_0}}\right), \quad (3.2.3)$$

where

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-y^2/2) dy \quad (3.2.4)$$

is the error function. The last equality in (3.2.3) is since we may consider only the noise component in the direction of the line between \mathbf{s}_0 and \mathbf{s}_i . The error probability for triplets is given by ([4],[19])

$$\begin{aligned} P(\varepsilon_{0i} \cap \varepsilon_{0j}|\mathbf{s}_0, i \neq j) &= P(\{\mathbf{r} \in \mathbb{R}^K : \|\mathbf{r} - \mathbf{s}_i\| < \|\mathbf{r} - \mathbf{s}_0\|, \|\mathbf{r} - \mathbf{s}_j\| < \|\mathbf{r} - \mathbf{s}_0\|\}|\mathbf{s}_0) \\ &= \Psi\left(\rho_{ij}, \frac{\|\mathbf{s}_0 - \mathbf{s}_i\|}{\sqrt{2N_0}}, \frac{\|\mathbf{s}_0 - \mathbf{s}_j\|}{\sqrt{2N_0}}\right), \end{aligned} \quad (3.2.5)$$

where

$$\rho_{ij} = \frac{\langle \mathbf{s}_i - \mathbf{s}_0, \mathbf{s}_j - \mathbf{s}_0 \rangle}{\|\mathbf{s}_i - \mathbf{s}_0\| \|\mathbf{s}_j - \mathbf{s}_0\|} \quad (3.2.6)$$

is the correlation between $\mathbf{s}_i - \mathbf{s}_0$ and $\mathbf{s}_j - \mathbf{s}_0$, $\langle \cdot, \cdot \rangle$ stands for the usual dot product and $\Psi(\cdot, \cdot, \cdot)$ is the bivariate normal distribution

$$\Psi(\rho, x', y') \triangleq \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{x'}^\infty \int_{y'}^\infty \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} dx dy. \quad (3.2.7)$$

All this said, one can apply de Caen's lower bound (2.2.8) to (3.2.1). We set

$$\mathcal{I} = \{1, \dots, M-1\}, \quad (3.2.8)$$

$$A_i = \varepsilon_{0i}, \quad (3.2.9)$$

resulting in Seguin's bound for any signal set

$$P(\varepsilon|\mathbf{s}_0) \geq \sum_{i=1}^{M-1} \frac{Q^2\left(\frac{\|\mathbf{s}_0 - \mathbf{s}_i\|}{\sqrt{2N_0}}\right)}{\sum_{j=1}^{M-1} \Psi\left(\rho_{ij}, \frac{\|\mathbf{s}_0 - \mathbf{s}_i\|}{\sqrt{2N_0}}, \frac{\|\mathbf{s}_0 - \mathbf{s}_j\|}{\sqrt{2N_0}}\right)}, \quad (3.2.10)$$

where for $\rho_{ij} = 1$ we have

$$\Psi(1, x, x) = Q(x). \quad (3.2.11)$$

3.3 Analysis

In this section, we derive the new lower bounds on the error probability for signals in the AWGN channel.

3.3.1 New Lower Bounds for Any Signal Set

In order to use the bound in Theorem 2.1, remember that

$$p(\mathbf{r}|\mathbf{s}_0) = (\pi N_0)^{-\frac{K}{2}} \exp \left\{ -\frac{1}{N_0} \|\mathbf{r} - \mathbf{s}_0\|^2 \right\}. \quad (3.3.1)$$

Technically speaking, the essence of the computation of the bound is the evaluation of the following integrals.

Definition 3.1 *Denote the following two integrals*

$$\int_{\varepsilon_{0i}} p(\mathbf{r}|\mathbf{s}_0) m(\mathbf{r}|\mathbf{s}_0) d\mathbf{r} \quad (3.3.2)$$

$$\int_{\varepsilon_{0i} \cap \varepsilon_{0j}} p(\mathbf{r}|\mathbf{s}_0) m^2(\mathbf{r}|\mathbf{s}_0) d\mathbf{r}, \quad (3.3.3)$$

as the generalized pairwise error probability and the generalized triplets error probability respectively.

Note that for the trivial choice of $m(\mathbf{r}|\mathbf{s}_0)$, i.e. $m(\mathbf{r}|\mathbf{s}_0) \equiv 1$, these are simply the pairwise error probability and triplets error probability as defined in (3.2.3) and (3.2.5) respectively.

To derive a lower bound on the error probability, a proper $m(\mathbf{r}|\mathbf{s}_0)$ should be chosen. Remember that the optimal value of $m(\mathbf{r}|\mathbf{s}_0)$ was given in Theorem 2.1,

$$m^*(\mathbf{r}|\mathbf{s}_0) = \frac{1}{deg(\mathbf{r}|\mathbf{s}_0)}, \quad (3.3.4)$$

where $deg(\mathbf{r}|\mathbf{s}_0)$ is the number of signals which are closer to \mathbf{r} than \mathbf{s}_0 , i.e.,

$$deg(\mathbf{r}|\mathbf{s}_0) = |\{i \neq 0 : \|\mathbf{r} - \mathbf{s}_i\| < \|\mathbf{r} - \mathbf{s}_0\|\}|. \quad (3.3.5)$$

However, the evaluation of $deg(\mathbf{r}|\mathbf{s}_0)$ is usually very complex and, in many practical cases, infeasible when only the weight enumeration of the code is to be used. Moreover, $m(\mathbf{r}|\mathbf{s}_0)$

should be mathematically endurable so the integrals in (3.3.2) and (3.3.3) can be easily computed². Nevertheless, we will see that suitable approximations can be found.

A first approximation is derived directly from equation (3.3.5). Since $deg(\mathbf{r}|\mathbf{s}_0)$ is the number of signals in the interior of a sphere of radius $\|\mathbf{r} - \mathbf{s}_0\|$ centered at \mathbf{r} , one might suggest that the larger the volume of the sphere, the higher $deg(\mathbf{r}|\mathbf{s}_0)$ is. Namely, $deg(\mathbf{r}|\mathbf{s}_0)$ is monotonically increasing in $\|\mathbf{r} - \mathbf{s}_0\|$. Thus, $m(\mathbf{r}|\mathbf{s}_0)$ might be chosen as

$$m(\mathbf{r}|\mathbf{s}_0) = \exp \{-a\|\mathbf{s}_0 - \mathbf{r}\|^2\}, \quad (3.3.6)$$

where $a \geq 0$ is a parameter to be optimized in order to achieve the tightest bound. An exponential behavior was chosen to facilitate the computation of (3.3.2) and (3.3.3). A drawback of this approximation however, is that it is implicitly based on the infeasible assumption that the signals are uniformly distributed in \mathbb{R}^K . Nevertheless, this choice does improve on the trivial choice of $m(\mathbf{r}|\mathbf{s}_0)$ (corresponding to $a = 0$) in terms of performance, as we will see in Section 3.4.

Fortunately, for equal-energy signals, a more realistic approximation can be derived in a similar fashion. Since for all i , $\|\mathbf{s}_i\| = \|\mathbf{s}_0\| = \sqrt{E}$, we have

$$deg(\mathbf{r}|\mathbf{s}_0) = |\{i \neq 0 : \langle \mathbf{s}_i, \mathbf{r} \rangle > \langle \mathbf{s}_0, \mathbf{r} \rangle\}| = |\{i \neq 0 : \theta_{\mathbf{r}i} < \theta_{\mathbf{r}0}\}|. \quad (3.3.7)$$

where

$$\theta_{\mathbf{r}i} \triangleq \cos^{-1} \left\{ \frac{\langle \mathbf{s}_i, \mathbf{r} \rangle}{\|\mathbf{s}_i\| \|\mathbf{r}\|} \right\}, \quad 0 \leq \theta_{\mathbf{r}i} < \pi. \quad (3.3.8)$$

Assuming the signals are uniformly distributed on the *surface* of a sphere of radius \sqrt{E} centered at the origin, equation (3.3.7) implies that $deg(\mathbf{r}|\mathbf{s}_0)$ is monotonically increasing with respect to the *absolute value of the angle* between \mathbf{r} and \mathbf{s}_0 . Thus, $m(\mathbf{r}|\mathbf{s}_0)$ might be chosen as

$$m(\mathbf{r}|\mathbf{s}_0) = \exp \{a\langle \mathbf{s}_0, \mathbf{r} \rangle\}, \quad (3.3.9)$$

²For example, the pairwise error probability, i.e., when $m(\mathbf{r}|\mathbf{s}_0) \equiv 1$, can be computed by a one dimensional integral, regardless of the signal set's dimension. Fortunately, as we will soon see, there exist non trivial choices of $m(\mathbf{r}|\mathbf{s}_0)$ with the same property.

where, again, $a \geq 0$ is a parameter to be optimized. Clearly, when BPSK modulation of a binary code is used, which is the case drawing our main attention, the signals are of equal energy. However, equal-energy signals are worth considering anyhow. It is well known (see, for example, [7]) that the requirement for equal-energy signals does not cause any degradation in performance, in terms of the error exponent, compared to an equivalent (a sphere of the same radius) peak-energy requirement³. An average-energy constraint, however, results in a slightly altered lower bound on the error probability. The assumption that the signals are uniformly distributed on the surface of the sphere cannot, of course, be justified in general. However, it is important to note that since this assumption is at the basis of the sphere packing bound [7, Section 3], which is asymptotically tight for rates higher than R_c , we know that good codes of high rate do have approximately uniform distribution of codewords on the surface of the sphere. Thus, we expect (3.3.9) to be useful for long codes and rates higher than R_c .

Both suggestions for $m(\mathbf{r}|\mathbf{s}_0)$, defined in (3.3.6) and (3.3.9), are members of a wider family, characterized by three parameters a, b and c ,

$$m(\mathbf{r}|\mathbf{s}_0) = \exp \left\{ - \left(a\|\mathbf{r}\|^2 + b\langle \mathbf{r}, \mathbf{s}_0 \rangle + c\|\mathbf{s}_0\|^2 \right) \right\}. \quad (3.3.10)$$

Although more suggestions for $m(\mathbf{r}|\mathbf{s}_0)$ can be given, we choose to focus on (3.3.10) only. The following proposition introduces the new bound on the error probability for any signal set, using this suggestion. The simpler suggestions discussed earlier easily follow.

Proposition 3.2 *Let $\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{M-1}$ be a set of M signals of dimension K for the AWGN*

³This is easily explained ([7]) by adding a coordinate to the code with the required value to achieve equal-energy signals.

channel with spectral density $\frac{N_0}{2}$. The conditional probability of error of a maximum likelihood decoder is lower bounded by

$$P(\varepsilon|\mathbf{s}_0) \geq \exp\{(\beta' - 2\beta)\|\mathbf{s}_0\|^2\} \left(\frac{N_0'}{\sqrt{N_0 N_0''}} \right)^K \cdot \sum_{i=1}^{M-1} \frac{Q^2 \left(\frac{\|\alpha \mathbf{s}_0 - \mathbf{s}_i\|^2 - (\alpha-1)^2 \|\mathbf{s}_0\|^2}{\sqrt{2N_0'} \|\mathbf{s}_0 - \mathbf{s}_i\|} \right)}{\sum_{j=1}^{M-1} \Psi \left(\rho_{ij}, \frac{\|\alpha' \mathbf{s}_0 - \mathbf{s}_i\|^2 - (\alpha'-1)^2 \|\mathbf{s}_0\|^2}{\sqrt{2N_0''} \|\mathbf{s}_0 - \mathbf{s}_i\|}, \frac{\|\alpha' \mathbf{s}_0 - \mathbf{s}_j\|^2 - (\alpha'-1)^2 \|\mathbf{s}_0\|^2}{\sqrt{2N_0''} \|\mathbf{s}_0 - \mathbf{s}_j\|} \right)} \quad (3.3.11)$$

where

$$\begin{aligned} N_0' &= \frac{N_0}{1 + aN_0} & N_0'' &= \frac{N_0}{1 + 2aN_0} \\ \alpha &= \left(\frac{\frac{1}{N_0} - \frac{b}{2}}{a + \frac{1}{N_0}} \right) & \alpha' &= \left(\frac{\frac{1}{N_0} - b}{2a + \frac{1}{N_0}} \right) \\ \beta &= \frac{\left(\frac{1}{N_0} + a \right) \left(\frac{1}{N_0} + c \right) - \left(\frac{1}{N_0} - \frac{b}{2} \right)^2}{\frac{1}{N_0} + a} & \beta' &= \frac{\left(\frac{1}{N_0} + 2a \right) \left(\frac{1}{N_0} + 2c \right) - \left(\frac{1}{N_0} - b \right)^2}{\frac{1}{N_0} + 2a}, \end{aligned} \quad (3.3.12)$$

and where $a > -\frac{1}{2N_0}$, b , and c are arbitrary constants.

Proof . We apply the lower bound on the probability of a union given in (2.2.2), using equations (3.2.8), (3.2.9), (3.3.1) and (3.3.10). Equation (3.3.11) easily follows after computing the integrals (3.3.2) and (3.3.3). Since the computation of these integrals is rather cumbersome, it is included in Appendix A. \square

Clearly, choosing $a = b = c = 0$ results in Seguin's bound (3.2.10), hence the bound in (3.3.11) is at least as tight as Seguin's. To restrict ourselves to simpler bounds, when only one parameter can be optimized, we may choose $a = c = a'$, $b = -2a'$, which results in the *norm* bound (i.e., using (3.3.6)), or $a = c = 0$, $b = -a'$, which results in the *dot product* bound (i.e., using (3.3.9)).

3.3.2 New Lower Bounds for Linear Codes

The bound given in Proposition 3.2 requires two nested summations over the entire signal set. Thus, the bound is of very little use for large codes. In this section, we specialize the bound given in Proposition 3.2 for linear codes and BPSK modulation. In this case, a bound depending on the code only through its weight enumeration is achieved, hence the nested summations are only over the code length, N .

Analogously to Seguin's work [4], we assume a binary (N, K) linear code \mathcal{C} is used. \mathbf{s}_i is obtained by replacing the zeroes and ones in \mathbf{c}_i with $\sqrt{E_N}$ and $-\sqrt{E_N}$ respectively⁴ (BPSK modulation). The energy per bit in this case is $E_b = \frac{NE_N}{K}$. Denote by $w(\mathbf{c})$ the Hamming weight of the codeword \mathbf{c} and by $\mathcal{B} = \{B_0, B_1, \dots, B_N\}$ the weight enumeration of the code, i.e., B_i is the number of codewords of Hamming weight i . Assuming \mathbf{c}_0 is the all-zero codeword, we have

$$\|\mathbf{s}_0\|^2 = NE_N, \quad (3.3.13)$$

$$\|\alpha\mathbf{s}_0 - \mathbf{s}_i\|^2 = (\alpha - 1)^2 NE_N + 4\alpha E_N w(\mathbf{c}_i). \quad (3.3.14)$$

Hence,

$$Q\left(\frac{\|\alpha\mathbf{s}_0 - \mathbf{s}_i\|^2 - (\alpha - 1)^2 \|\mathbf{s}_0\|^2}{\sqrt{2N'_0} \|\mathbf{s}_0 - \mathbf{s}_i\|}\right) = Q\left(\sqrt{\frac{\alpha^2 E_N w(\mathbf{c}_i)}{N'_0/2}}\right) \quad (3.3.15)$$

and

$$\begin{aligned} \Psi\left(\rho_{ij}, \frac{\|\alpha'\mathbf{s}_0 - \mathbf{s}_i\|^2 - (\alpha' - 1)^2 \|\mathbf{s}_0\|^2}{\sqrt{2N''_0} \|\mathbf{s}_0 - \mathbf{s}_i\|}, \frac{\|\alpha'\mathbf{s}_0 - \mathbf{s}_j\|^2 - (\alpha' - 1)^2 \|\mathbf{s}_0\|^2}{\sqrt{2N''_0} \|\mathbf{s}_0 - \mathbf{s}_j\|}\right) \\ = \Psi\left(\rho_{ij}, \sqrt{\frac{\alpha'^2 E_N w(\mathbf{c}_i)}{N''_0/2}}, \sqrt{\frac{\alpha'^2 E_N w(\mathbf{c}_j)}{N''_0/2}}\right), \end{aligned} \quad (3.3.16)$$

where

$$\rho_{ij} = \frac{w(\mathbf{c}_i \mathbf{c}_j)}{\sqrt{w(\mathbf{c}_i)w(\mathbf{c}_j)}}. \quad (3.3.17)$$

⁴Referring to Section 3.2, note that the signals in this case are of length N and dimension K . Signals of length K are achieved by the Gram-Schmidt procedure and the projection of the signals on the new base. The computation of the signals' energy and distance spectrum is, however, clearer when the original signals are treated.

The expressions in (3.3.15) and (3.3.16) can be substituted into (3.3.11) directly. However, ρ_{ij} does not depend solely on the code's weight enumeration. In [4], Seguin proved that $\Psi(\rho, x, y)$ is monotonically increasing in ρ . Thus, to derive a bound which depends only on the weight enumeration of the code, and is thus much easier to evaluate, ρ_{ij} can be upper bounded in terms of the weight enumeration alone. Denote by d the minimum distance of the code, for $\mathbf{s}_i \neq \mathbf{s}_j$, an upper bound on ρ_{ij} , derived in [4], is given by

$$\rho_{ij} \leq \varrho(i, j) \triangleq \min \left\{ \sqrt{\frac{w(\mathbf{c}_i)}{w(\mathbf{c}_j)}}, \sqrt{\frac{w(\mathbf{c}_j)}{w(\mathbf{c}_i)}}, \frac{w(\mathbf{c}_i) + w(\mathbf{c}_j) - d}{2\sqrt{w(\mathbf{c}_i)w(\mathbf{c}_j)}} \right\}, \quad (3.3.18)$$

which can be substituted into (3.3.16) to achieve the desired expression. All this said, the following proposition introduces the new bound on the error probability for linear codes in the AWGN channel.

Proposition 3.3 *Let \mathcal{C} be a binary (N, K) linear code used over the AWGN channel with BPSK modulation. The probability of error of the maximum likelihood decoder is lower bounded by*

$$P(\varepsilon) \geq \exp\{(\beta' - 2\beta)NE_N\} \left(\frac{N'_0}{\sqrt{N_0 N''_0}} \right)^K \sum_{i \neq 0} \frac{B_i Q^2 \left(\sqrt{\frac{\alpha^2 E_N i}{N''_0/2}} \right)}{Q \left(\sqrt{\frac{\alpha'^2 E_N i}{N''_0/2}} \right) + (B_i - 1) \Psi \left(\varrho_{ii}, \sqrt{\frac{\alpha'^2 E_N i}{N''_0/2}}, \sqrt{\frac{\alpha'^2 E_N i}{N''_0/2}} \right) + \sum_{j \neq 0, i} B_j \Psi \left(\varrho_{ij}, \sqrt{\frac{\alpha'^2 E_N i}{N''_0/2}}, \sqrt{\frac{\alpha'^2 E_N j}{N''_0/2}} \right)} \quad (3.3.19)$$

where

$$\varrho_{ij} = \min \left\{ \sqrt{\frac{i}{j}}, \sqrt{\frac{j}{i}}, \frac{i + j - d}{2\sqrt{ij}} \right\}, \quad (3.3.20)$$

$\alpha, \alpha', \beta, \beta', N'_0$ and N''_0 are as defined in (3.3.12), and $a > -\frac{1}{2N_0}$, b , and c are parameters to be optimized in order to achieve the tightest bound.

Proof. Substitute expressions (3.3.15), (3.3.16) and (3.3.18) into (3.3.11). Since the resulting summands depend on the code only through the weight enumeration, the summation

can be carried out on the possible codewords weights. Finally, when linear codes are used on a binary-input output-symmetric channel with maximum likelihood decoding, the probability of error is independent of the codeword sent (see [20, pp. 86]). Hence, we assume the all-zero codeword was sent and $P(\varepsilon|\mathbf{s}_0) = P(\varepsilon)$. \square

Remark 3.4 Consider the limiting cases of $\frac{E_N}{N_0} \rightarrow \infty$ and $\frac{E_N}{N_0} \rightarrow 0$. While non-trivial values of the parameters a , b and c yield strictly tighter bounds for intermediate values of $\frac{E_N}{N_0}$, it is not so in these cases. When $\frac{E_N}{N_0} \rightarrow \infty$, Seguin's bound is optimal ([4, Section 5]), in the sense that the ratio with the union bound tends to unity (referring to Remark 2.3, in this case the error probability for triplets is negligible compared to the pairwise error probability). Therefore, no non-trivial values of the parameters a , b and c yield tighter results. Note that, however, the rate of convergence may be faster with non-trivial parameters. When $\frac{E_N}{N_0} \rightarrow 0$, since the bound cannot be represented as a function of $\frac{E_N}{N_0}$, unwieldy limit-computations are required. We will not include these computations in this work, only mention the key steps. The first step is to note that in the denominator of the r.h.s. of (3.3.19), ϱ_{ij} is always smaller than unity, thus when $\sqrt{\frac{\alpha'^2 E_N i}{N_0' / 2}} \rightarrow \infty$, Ψ tends to zero faster than Q and hence is negligible. Continuing with the following bounds ([20, pp. 63])

$$\left(1 - \frac{1}{\beta^2}\right) \frac{e^{-\beta^2/2}}{\sqrt{2\pi}\beta} < Q(\beta) < \frac{e^{-\beta^2/2}}{\sqrt{2\pi}\beta}, \quad \beta > 0 \quad (3.3.21)$$

and straightforward computations it is easy to see that the optimal values for the parameters in this case are the trivial ones.

3.3.3 Lower Bounds Depending Only on the Subcode \mathcal{C}_d^*

The bound given in Proposition 3.3 requires the weight enumeration of the entire code to be evaluated. However, for large codes, this weight enumeration is rarely available. Clearly, the error probability of a given code \mathcal{C} , $P_{\mathcal{C}}(\varepsilon)$, satisfies $P_{\mathcal{C}}(\varepsilon) \geq P_{\mathcal{C}^*}(\varepsilon)$, where \mathcal{C}^* is any subset of the code \mathcal{C} . Hence, any lower bound on $P_{\mathcal{C}^*}(\varepsilon)$ is a lower bound on $P_{\mathcal{C}}(\varepsilon)$. This

technique is widely used when lower bounds for low rates are discussed (see, for example, [20, pp. 174]). However, when the code \mathcal{C} is linear, \mathcal{C}^* is not necessarily linear. Nevertheless, we have

$$P_{\mathcal{C}}(\varepsilon) = P_{\mathcal{C}}(\varepsilon|\mathbf{c}_0) \geq P_{\mathcal{C}^*}(\varepsilon|\mathbf{c}_0) \quad (3.3.22)$$

for any subcode \mathcal{C}^* of a linear code \mathcal{C} . As in [5], we choose

$$\mathcal{C}^* = \mathcal{C}_d^* \triangleq \{\mathbf{c} \in \mathcal{C} : w(\mathbf{c}) = d\} \cup \{\mathbf{c}_0\}. \quad (3.3.23)$$

The resulting lower bound is given in the following proposition.

Proposition 3.5 *Let \mathcal{C} be a binary (N, K) linear code used over the AWGN channel with BPSK modulation. The probability of error of a maximum likelihood decoder is lower bounded by*

$$P(\varepsilon) \geq \frac{\exp\{(\beta' - 2\beta)NE_N\} \left(\frac{N'_0}{\sqrt{N_0 N''_0}}\right)^K B_d Q^2\left(\sqrt{\frac{\alpha^2 E_N d}{N'_0/2}}\right)}{Q\left(\sqrt{\frac{\alpha'^2 E_N d}{N''_0/2}}\right) + (B_d - 1)\Psi\left(\frac{1}{2}, \sqrt{\frac{\alpha'^2 E_N d}{N''_0/2}}, \sqrt{\frac{\alpha^2 E_N d}{N'_0/2}}\right)} \quad (3.3.24)$$

where $\alpha, \alpha', \beta, \beta', N'_0$ and N''_0 are as defined in (3.3.12), and $a > -\frac{1}{2N_0}$, b , and c are parameters to be optimized in order to achieve the tightest bound.

Proof. Based on the preceding discussion, substitute $B_i = 0$ for every $i \neq d$ in (3.3.19). The one-half appearing as the first argument of Ψ results from the substitution of $i = j = d$ in (3.3.20). \square

3.3.4 Kounias' Bound

We apply Kounias' lower bound [3] to derive a new lower bound, analogously to the preceding derivations in this chapter. Although Kounias' bound was used by Kuai, Alajaji and Takahara in [19] to derive a lower bound for the AWGN channel, their work included bounds for non-uniform signaling, hence no specialization of the bound for linear codes was

possible. In this section, in addition to the straightforward specialization for linear codes, we further develop the bound by using only the subcode \mathcal{C}_d^* . In this case, the customarily tedious optimization required in Kounias' bound is direct and can be done analytically. The resulting bound is very simple to evaluate and performs better than Seguin's bound (yet, inferior to (3.3.24)) for every value of E_b/N_0 .

Kounias Bound for Any Signal Set

Under the notations of Chapter 2, Kounias' bound is given by

$$P\left(\bigcup_{i \in \mathcal{I}} A_i\right) \geq \max_{\mathcal{J} \subseteq \mathcal{I}} \left\{ \sum_{i \in \mathcal{J}} P(A_i) - \sum_{i, j \in \mathcal{J}, i < j} P(A_i \cap A_j) \right\}. \quad (3.3.25)$$

Refereing to our problem, utilization of this bound yields

$$P\left(\bigcup_{i \neq 0} \varepsilon_{0i} | \mathbf{s}_0\right) \geq \max_{\mathcal{J} \subseteq \mathcal{M} \setminus \{0\}} \left\{ \sum_{i \in \mathcal{J}} Q\left(\frac{\|\mathbf{s}_0 - \mathbf{s}_i\|}{\sqrt{2N_0}}\right) - \sum_{i, j \in \mathcal{J}, i < j} \Psi\left(\rho_{ij}, \frac{\|\mathbf{s}_0 - \mathbf{s}_i\|}{\sqrt{2N_0}}, \frac{\|\mathbf{s}_0 - \mathbf{s}_j\|}{\sqrt{2N_0}}\right) \right\}, \quad (3.3.26)$$

where $\mathcal{M} = \{0, 1, \dots, M-1\}$. Note that the fact that Kounias' bound allows us to use any subset $\mathcal{J} \subseteq \mathcal{M} \setminus \{0\}$ is insignificant since when lower bounds on the error probability are considered, this step is straightforward (refer to Section 3.3.3). Hence, in this case Kounias' bound is equivalent to the well known Bonferroni's inclusion-exclusion lower bound ([21]).

Kounias Bound for Linear Codes

So far, the preceding derivations were identical to those of Kuai, Alajaji and Takahara in [19]. To specialize the bound for linear codes, note that the r.h.s. of (3.3.26) is a decreasing function of Ψ , therefore, we can use ρ_{ij} as in (3.3.19), resulting in a bound depending only on the weight enumeration. Remember that \mathcal{B} is the weight enumeration of the code, i.e., B_i is the number of codewords of weight i . However, Kounias' bound allows us to use any subset $\mathcal{J} \subseteq \mathcal{M} \setminus \{0\}$ of the code (in this case, \mathcal{J} is a subset of *indices*). Let \mathcal{J} be any such subset. We denote the weight enumeration of the corresponding subcode by $\mathcal{B}^{\mathcal{J}}$, i.e.,

$B_i^{\mathcal{J}} \leq B_i$ is the number of codewords of weight i in this subcode. Note, however, that this is not the distance distribution of the subcode. Thus, when linear codes are considered, the maximum is over all possible $\mathcal{B}^{\mathcal{J}}$'s. Since several subcodes may have the same weight enumeration, the optimization process is much faster than that of (3.3.26). The resulting lower bound is

$$p(\varepsilon) \geq \max_{\mathcal{B}^{\mathcal{J}}} \left\{ \sum_{B_i^{\mathcal{J}} \neq 0} B_i^{\mathcal{J}} Q \left(\sqrt{\frac{2E_N i}{N_0}} \right) - \sum_{B_i^{\mathcal{J}} \neq 0} \left(\binom{B_i^{\mathcal{J}}}{2} \Psi \left(\varrho_{ii}, \sqrt{\frac{2E_N i}{N_0}}, \sqrt{\frac{2E_N i}{N_0}} \right) + \sum_{B_j^{\mathcal{J}} \neq 0, j > i} B_i^{\mathcal{J}} B_j^{\mathcal{J}} \Psi \left(\varrho_{ij}, \sqrt{\frac{2E_N i}{N_0}}, \sqrt{\frac{2E_N j}{N_0}} \right) \right) \right\}, \quad (3.3.27)$$

with the understanding, relevant throughout this document, that $\binom{n}{k}$ is 0 if $k > n$.

The bound in (3.3.27) is still tedious to evaluate for large codes, even when the stepwise algorithm suggested in [19] is used. Thus, analogously to the derivations of Section 3.3.3, we limit the search to weight enumerations of *subsets of the subcode* \mathcal{C}_d^* . Hence, we have

$$p(\varepsilon) \geq \max_{1 \leq b \leq B_d} \left\{ bQ \left(\sqrt{\frac{2E_N d}{N_0}} \right) - \binom{b}{2} \Psi \left(\frac{1}{2}, \sqrt{\frac{2E_N d}{N_0}}, \sqrt{\frac{2E_N d}{N_0}} \right) \right\}. \quad (3.3.28)$$

Since the r.h.s. of (3.3.28) is a concave (\cap) function of b and its *second* derivative with respect to b is constant, the maximum is achieved by comparing the first derivative to zero and taking the closest integer value to the result, provided that it is in the range $\{1, 2, \dots, B_d\}$. Thus, the maximum is achieved with

$$b^* = \min \left\{ \left\lceil \frac{1}{2} + \frac{Q \left(\sqrt{\frac{2E_N d}{N_0}} \right)}{\Psi \left(\frac{1}{2}, \sqrt{\frac{2E_N d}{N_0}}, \sqrt{\frac{2E_N d}{N_0}} \right)} \right\rceil, B_d \right\} \quad (3.3.29)$$

where $\lceil x \rceil$ is closest integer to x . Consequently, we have

$$p(\varepsilon) \geq b^* Q \left(\sqrt{\frac{2E_N d}{N_0}} \right) - \binom{b^*}{2} \Psi \left(\frac{1}{2}, \sqrt{\frac{2E_N d}{N_0}}, \sqrt{\frac{2E_N d}{N_0}} \right). \quad (3.3.30)$$

3.4 Results

In this section, several examples with well known codes are given and the results of the numerical analysis are shown.

3.4.1 Implementation Notes

Before the numerical results for the lower bounds are introduced, we address several practical issues. First, the definition of $Q(\cdot)$ as given in (3.2.4) requires an integration over an infinite set. Instead, an alternative form by Craig⁵ [23] was implemented⁶

$$Q(x) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{x^2}{2 \sin^2 \theta}\right) d\theta \quad x \geq 0. \quad (3.4.1)$$

As for $\Psi(\cdot, \cdot, \cdot)$, an expression given by Simon and Divsalar in [25] was used

$$\begin{aligned} \Psi(\rho, x, y) = & \frac{1}{2\pi} \int_0^{\pi/2 - \tan^{-1}(y/x)} \frac{\sqrt{1 - \rho^2}}{1 - \rho \sin 2\theta} \exp\left\{-\frac{x^2}{2} \frac{1 - \rho \sin 2\theta}{(1 - \rho^2) \sin^2 \theta}\right\} d\theta \\ & + \frac{1}{2\pi} \int_0^{\tan^{-1}(y/x)} \frac{\sqrt{1 - \rho^2}}{1 - \rho \sin 2\theta} \exp\left\{-\frac{y^2}{2} \frac{1 - \rho \sin 2\theta}{(1 - \rho^2) \sin^2 \theta}\right\} d\theta. \end{aligned} \quad (3.4.2)$$

3.4.2 Examples and Numerical Analysis Results

We compare the new lower bounds for linear codes, presented in this chapter, with several known bounds in the current literature. For the sake of simplicity, only three new bounds are discussed. The first is the *norm bound - whole code*, i.e., the bound given in (3.3.19) with $a = c = a'$ and $b = -2a'$. The second is the *dot product bound - subcode \mathcal{C}_d^** , i.e., the bound given in (3.3.24) with $a = c = 0$ and $b = -a'$. The *norm bound* using only the subcode \mathcal{C}_d^* and the *dot product bound* using the whole code are not given since these bounds do not further improve the results or give any different insight. The third bound is Kounias' lower bound as given in (3.3.30). The new bounds are compared to Seguin's lower bound [4], Shannon's lower bound [7] and Poltyrev's upper bound [6], all appearing

⁵also appearing in [22], with a simpler proof.

⁶A simple series for the computation of $Q(\cdot)$ by Beaulieu [24] was also implemented as a reference.

also in [4, Section 6] as bounds for linear codes. Swaszek's lower bound [16] is not given since it appears to be inferior to Seguin's ([4]). The results for the codes BCH(63,24) and Golay(23,12) are given in Figures 3.1 and 3.2, respectively. For the sake of clarity, Figure 3.2 does not include Kounias' bound. It is only slightly superior to Seguin's.

It is clear that the new bounds perform better than Seguin's for any value of E_b/N_0 (for BCH(63,24), the improvement is up to 21dB where the bounds are inferior to Shannon's bound and up to 12dB where the bounds are superior). This can be seen both for the bound using the whole code (as in Seguin's bound) and for the bound using the subcode \mathcal{C}_d^* . Referring to Remark 3.4, the fact that non trivial values of the parameters cannot improve performance for very high values of E_b/N_0 is clear. This behavior is repeated for low values of E_b/N_0 , although clear only for $E_b/N_0 < -3dB$, a range not included in the presented graphs. To the author's knowledge, for high values of E_b/N_0 , where the new bounds are superior to Shannon's lower bound, they establish the best know results in the current literature. Furthermore, in Chapter 5, we calculate the upper bound on the error exponent resulting from the *dot product* bound. It is shown therein that this upper bound can be tighter than the de Caen-based bound, and the optimal value of the parameter a' is given.

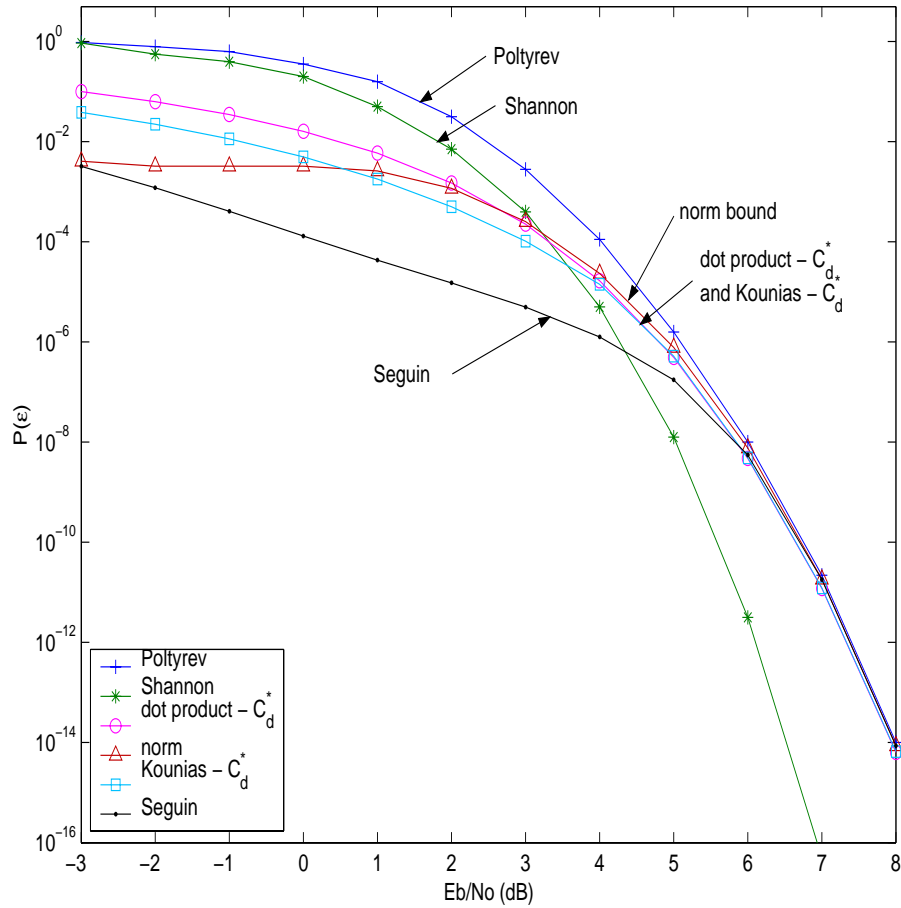


Figure 3.1: Bounds on the decoding error probability of BCH(63,24) code, AWGN channel. The new lower bounds *norm* - whole code and *dot product* - subcode C_d^* are shown, together with Kounias' version (C_d^*) for linear codes (lower bound). For reference, Poltyrev's upper bound and Shannon's and Seguin's lower bounds are also plotted.

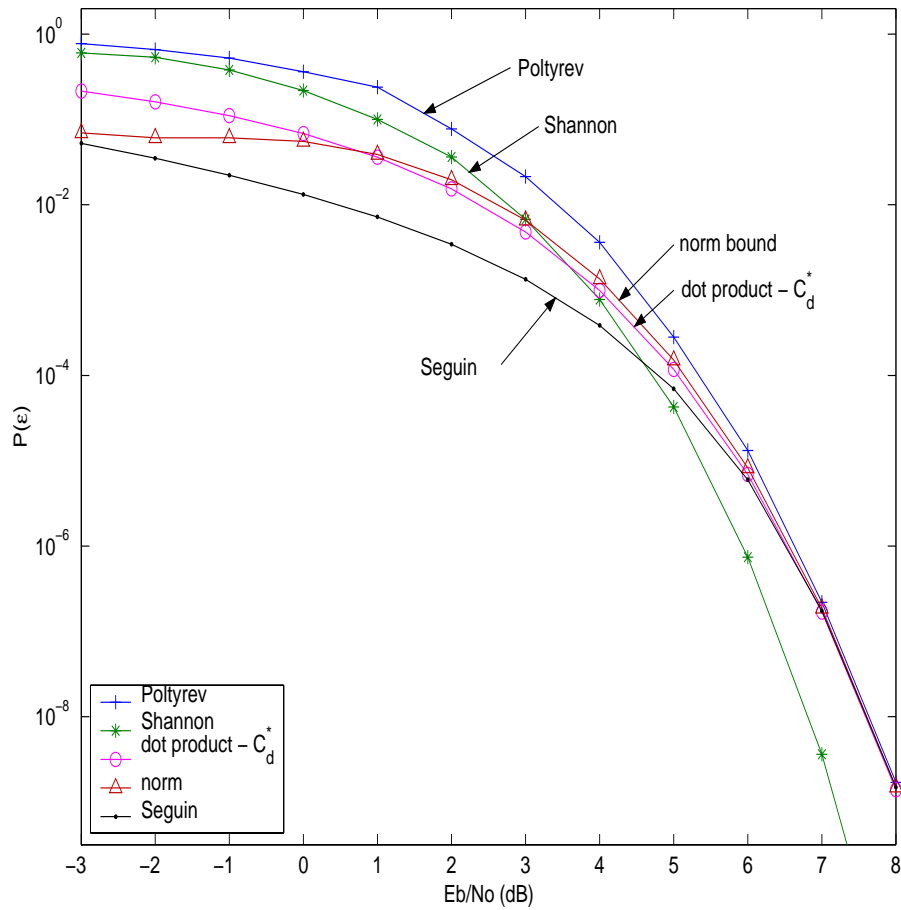


Figure 3.2: Bounds on the decoding error probability of Golay(23,12) code, AWGN channel. The new lower bounds *norm* - whole code and *dot product* - subcode C_d^* are shown. For reference, Poltyrev's upper bound and Shannon's and Seguin's lower bounds are also plotted.

Chapter 4

A Lower Bound on the Error Probability for the Binary Symmetric Channel

Analogously to the previous chapter, in this chapter, we apply the bound in Theorem 2.1 to derive new lower bounds for uniform signaling and maximum likelihood decoding over the BSC. In a similar fashion, the bounds are specialized for linear codes and several suggestions for the optimizing function are discussed.

4.1 Introduction

In this section, we briefly introduce the channel model and several related works.

We consider the case of uniform signaling over a BSC channel and maximum likelihood decoding. The transmitted codeword is one of $M = 2^K$ equiprobable binary codewords $\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_{M-1}$ of length N . Denote by p the channel crossover probability, i.e., the probability that the channel's output bit is different from the channel's input bit. Let \mathbf{x} be the received word. The optimal decoder, namely, the maximum likelihood decoder, chooses the

closest of the M codewords to \mathbf{x} in the Hamming sense, i.e., $\hat{i} = \arg \min_i d_H(\mathbf{c}_i, \mathbf{x})$. Thus, the probability of error given that \mathbf{c}_0 was sent is

$$P(\varepsilon|\mathbf{c}_0) = P(\cup_{i \neq 0} \varepsilon_{0i} | \mathbf{c}_0), \quad (4.1.1)$$

where in this case

$$\begin{aligned} \varepsilon_{0i} &= \{\mathbf{x} \in GF(2)^N : d_H(\mathbf{x}, \mathbf{c}_i) < d_H(\mathbf{x}, \mathbf{c}_0)\} \\ &= \{\mathbf{x} \in GF(2)^N : w(\mathbf{x} + \mathbf{c}_i) < w(\mathbf{x})\}, \end{aligned} \quad (4.1.2)$$

assuming \mathbf{c}_0 is the all-zero codeword.

Our goal is to lower bound the error probability given in (4.1.1). Again, when the code used is a binary (N, K) linear code \mathcal{C} , we wish to express the bound in terms of the code's weight enumeration and the channel's crossover probability alone. To the author's knowledge, the best known lower bounds, in this context, are those due to Keren and Litsyn, appearing in [5]¹ and the sphere packing bound (see, for example, [6]). Keren and Litsyn's technique is closely related to the one presented here since de Caen's lower bound on the probability of a union is a key element in their derivations. We discuss their technique thoroughly in the proceeding section. The sphere packing bound, citing [26], states that the probability of error is greater than that of a perfect code. Namely, the best probability of error is achieved when spheres of equal radii centered at the codewords cover the entire space. The best known upper bound is due to Poltyrev ([6]). Poltyrev's bound is based on the following inequality, which was also in the basis of the bound for the AWGN channel,

$$P(\varepsilon) \leq P(\varepsilon, \mathcal{A}) + P(\mathcal{A}^c). \quad (4.1.3)$$

for any $\mathcal{A} \subseteq GF(2)^N$. Poltyrev chose the set \mathcal{A} as the set of all binary words of weight higher than some threshold m , which is later optimized to yield the tightest bound. The value of $P(\varepsilon, \mathcal{A})$ was bounded by the well known union bound. We further discuss inequality (4.1.3) in the proceeding section.

¹A more detailed paper, unpublished though, is also available ([26]).

4.2 Analysis

In this section, we derive the new lower bounds on the error probability over the BSC. For the sake of simplicity, we consider only linear codes. Bounds for any block code can be derived in the same fashion.

4.2.1 New Lower Bounds for Linear Codes

We wish to use the bound in Theorem 2.1 to derive the new bounds for the BSC. Since the method developed in Chapter 3 is general, and can be used in any case where the error probability admits a union form, we focus only on the channel-specific derivations.

Let \mathcal{C} be any linear code. Assume \mathbf{c}_0 , the all-zero codeword, was sent. We have

$$\mathcal{I} = \{1, \dots, M-1\}, \quad (4.2.1)$$

$$A_i = \varepsilon_{0i}, \quad (4.2.2)$$

$$\begin{aligned} p(\mathbf{x}|\mathbf{c}_0) &= p^{d_H(\mathbf{x}, \mathbf{c}_0)}(1-p)^{N-d_H(\mathbf{x}, \mathbf{c}_0)} \\ &= p^{w(\mathbf{x})}(1-p)^{N-w(\mathbf{x})}, \end{aligned} \quad (4.2.3)$$

$$\begin{aligned} \text{deg}(\mathbf{x}|\mathbf{c}_0) &= |\{\mathbf{c}_i \in \mathcal{C}, i \neq 0 : d_H(\mathbf{x}, \mathbf{c}_i) < d_H(\mathbf{x}, \mathbf{c}_0)\}| \\ &= |\{\mathbf{c}_i \in \mathcal{C}, i \neq 0 : w(\mathbf{x} + \mathbf{c}_i) < w(\mathbf{x})\}|. \end{aligned} \quad (4.2.4)$$

Define the set $\mathcal{N} = \{1, 2, \dots, N\}$ and for every $\mathcal{M} \subseteq \mathcal{N}$ and $\mathbf{x} \in GF(2)^N$ define $\mathbf{x}_{\mathcal{M}}$ to be the subword $(\mathbf{x}(m_1), \dots, \mathbf{x}(m_{|\mathcal{M}|}))$ of \mathbf{x} . Let the *support* of \mathbf{c}_i be the set $\mathcal{S}_i = \{j : \mathbf{c}_i(j) = 1\}$ and denote by \mathcal{S}_i^c the set $\mathcal{N} \setminus \mathcal{S}_i$. Hence, $\mathbf{x}_{\mathcal{S}_i}$ is the subword of \mathbf{x} consisting of \mathbf{x} in the places \mathbf{c}_i equals 1 and $\mathbf{x}_{\mathcal{S}_i^c}$ is the subword of \mathbf{x} in the places \mathbf{c}_i equals 0.

In the computation of (2.2.2) according to equations (4.2.1) to (4.2.3), we encounter the summations over $\mathbf{x} \in \varepsilon_{0i}$ and $\mathbf{x} \in \varepsilon_{0i} \cap \varepsilon_{0j}$. The summands are $p(\mathbf{x}|\mathbf{c}_0)m_i(\mathbf{x})$ and $p(\mathbf{x}|\mathbf{c}_0)m_i^2(\mathbf{x})$, respectively. While the dependence of $p(\mathbf{x}|\mathbf{c}_0)$ on \mathbf{x} is only through $w(\mathbf{x})$, $m_i(\mathbf{x})$ is a function to be optimized and hence might, in general, be chosen to be different for each \mathbf{x} . However, to avoid a tedious evaluation of the considered sums, we prefer to

reduce the degrees of freedom in choosing $m_i(\mathbf{x})$ by the restriction

$$m_i(\mathbf{x}) = \eta_i(w(\mathbf{x})), \quad \eta_i : \mathbb{Z}^+ \mapsto \mathbb{R}. \quad (4.2.5)$$

Clearly, since $\deg(\mathbf{x}|\mathbf{c}_0)$ is not likely to depend only on $w(\mathbf{x})$ when non-trivial codes are discussed, we may assume that the optimal value for $m_i(\mathbf{x})$ cannot be achieved by any function η_i . Nevertheless, we will discover that the function η_i may still be chosen to yield tighter bounds than the one achieved with the trivial choice of $m_i(\mathbf{x})$. To conclude, according to (2.2.2), we may write

$$P(\varepsilon) \geq \sum_{i=1}^M \frac{(\sum_{\mathbf{x} \in \varepsilon_{0i}} p^{w(\mathbf{x})} (1-p)^{N-w(\mathbf{x})} \eta_i(w(\mathbf{x})))^2}{\sum_{j=1}^M \sum_{\mathbf{x} \in \varepsilon_{0i} \cap \varepsilon_{0j}} p^{w(\mathbf{x})} (1-p)^{N-w(\mathbf{x})} \eta_i^2(w(\mathbf{x}))} \quad (4.2.6)$$

where $\eta_i : \mathbb{Z}^+ \mapsto \mathbb{R}$ is any function to be optimized. In the proceeding subsection we estimate the value of η_i , in the spirit of equation (2.2.3). For the time being, we evaluate (4.2.6) for any η_i whose dependence on \mathbf{x} is only through $w(\mathbf{x})$.

To evaluate (4.2.6), we start by calculating the sum over $\mathbf{x} \in \varepsilon_{0i}$ in the numerator. Referring to (4.1.2), a word $\mathbf{x} \in GF(2)^N$ satisfies $\mathbf{x} \in \varepsilon_{0i}$ if under \mathbf{c}_i 's support it has more 1's than 0's. The number of 1's or 0's out of \mathbf{c}_i 's support is irrelevant. Thus

$$\mathbf{x} \in \varepsilon_{0i} \quad \text{iff} \quad w(\mathbf{x}_{S_i}) \geq \left\lfloor \frac{w(\mathbf{c}_i)}{2} \right\rfloor + 1. \quad (4.2.7)$$

Accordingly, we have

$$\begin{aligned} P_{num}(w(\mathbf{c}_i)) &\triangleq \sum_{\mathbf{x} \in \varepsilon_{0i}} p^{w(\mathbf{x})} (1-p)^{N-w(\mathbf{x})} \eta_i(w(\mathbf{x})) \\ &= \sum_{l=\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor + 1}^{w(\mathbf{c}_i)} \sum_{m=0}^{N-w(\mathbf{c}_i)} \binom{w(\mathbf{c}_i)}{l} \binom{N-w(\mathbf{c}_i)}{m} p^{l+m} (1-p)^{N-l-m} \eta_i(l+m). \end{aligned} \quad (4.2.8)$$

To avoid cumbersome notations, the notation for $P_{num}(w(\mathbf{c}_i))$ does not reflect its dependence on p and the parameter N . The first sum in (4.2.8) is over $l = w(\mathbf{x}_{S_i})$. Since

$$\left\lfloor \frac{w(\mathbf{c}_i)}{2} \right\rfloor + 1 \leq l \leq w(\mathbf{c}_i) \quad (4.2.9)$$

it is clear (4.2.7) is satisfied. The second sum is over $m = w(\mathbf{x}_{\mathcal{S}_i^c})$, counting all the possibilities for 1's under \mathcal{S}_i^c , where the received word has no restrictions. Note that (4.2.8) is the *generalized pairwise error probability* for the BSC, analogously to Definition 3.1. On the more technical side, choosing a non-trivial η_i prevents us from using the binomial formula to evaluate the second sum and thus increases the computational complexity. For the codes tested in this work, this tradeoff was worthwhile.

The evaluation of the sum in the denominator is carried out in the same fashion. In this case

$$\mathbf{x} \in \varepsilon_{0i} \cap \varepsilon_{0j} \quad \text{iff} \quad w(\mathbf{x}_{\mathcal{S}_i}) \geq \left\lfloor \frac{w(\mathbf{c}_i)}{2} \right\rfloor + 1 \quad \text{and} \quad w(\mathbf{x}_{\mathcal{S}_j}) \geq \left\lfloor \frac{w(\mathbf{c}_j)}{2} \right\rfloor + 1. \quad (4.2.10)$$

Thus, when $\mathbf{c}_i = \mathbf{c}_j$ we have

$$\begin{aligned} P_{den}(w(\mathbf{c}_i)) &\triangleq \sum_{\mathbf{x} \in \varepsilon_{0i}} p^{w(\mathbf{x})} (1-p)^{N-w(\mathbf{x})} \eta_i^2(w(\mathbf{x})) \\ &= \sum_{l=\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor + 1}^{w(\mathbf{c}_i)} \sum_{m=0}^{N-w(\mathbf{c}_i)} \binom{w(\mathbf{c}_i)}{l} \binom{N-w(\mathbf{c}_i)}{m} p^{l+m} (1-p)^{N-l-m} \eta_i^2(l+m). \end{aligned} \quad (4.2.11)$$

When $\mathbf{c}_i \neq \mathbf{c}_j$ we have

$$\begin{aligned} \tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j) &\triangleq \sum_{\mathbf{x} \in \varepsilon_{0i} \cap \varepsilon_{0j}, i \neq j} p^{w(\mathbf{x})} (1-p)^{N-w(\mathbf{x})} \eta_i^2(w(\mathbf{x})) \\ &= \sum_{l=0}^{w(\mathbf{c}_i, \mathbf{c}_j)} \sum_{m=\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor + 1 - l}^{w(\mathbf{c}_i) - w(\mathbf{c}_i, \mathbf{c}_j)} \sum_{n=\lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor + 1 - l}^{w(\mathbf{c}_j) - w(\mathbf{c}_i, \mathbf{c}_j)} \sum_{k=0}^{N-w(\mathbf{c}_i) - w(\mathbf{c}_j) + w(\mathbf{c}_i, \mathbf{c}_j)} \binom{w(\mathbf{c}_i, \mathbf{c}_j)}{l} \binom{w(\mathbf{c}_i) - w(\mathbf{c}_i, \mathbf{c}_j)}{m} \\ &\cdot \binom{w(\mathbf{c}_j) - w(\mathbf{c}_i, \mathbf{c}_j)}{n} \binom{N - w(\mathbf{c}_i) - w(\mathbf{c}_j) + w(\mathbf{c}_i, \mathbf{c}_j)}{k} p^{l+m+n+k} (1-p)^{N-l-m-n-k} \eta_i^2(l+m+n+k) \end{aligned} \quad (4.2.12)$$

where the first sum in (4.2.12) is over the intersection of \mathbf{c}_i 's and \mathbf{c}_j 's supports - \mathcal{S}_{ij} , the second is over $\mathcal{S}_i \setminus \mathcal{S}_{ij}$, the third is over $\mathcal{S}_j \setminus \mathcal{S}_{ij}$ and the fourth is over $\mathcal{N} \setminus \mathcal{S}_i \setminus \mathcal{S}_j$. Again, the expression in (4.2.12) is denoted as the *generalized triplets error probability* for the BSC. Clearly, $\tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j)$ does not depend on \mathbf{c}_i and \mathbf{c}_j solely through $w(\mathbf{c}_i)$ and $w(\mathbf{c}_j)$

since it includes the expression $w(\mathbf{c}_i \mathbf{c}_j)$. Thus, its evaluation requires more than the weight enumeration of the code. We recall dealing with an equivalent problem in the AWGN channel. In [4] Seguin has proved that

$$P(\varepsilon_{ui} \cap \varepsilon_{uj} | \mathbf{s}_0, i \neq j) = \Psi \left(\rho_{ij}, \frac{\|\mathbf{s}_0 - \mathbf{s}_i\|}{\sqrt{2N_0}}, \frac{\|\mathbf{s}_0 - \mathbf{s}_j\|}{\sqrt{2N_0}} \right) \quad (4.2.13)$$

is monotonically increasing in ρ_{ij} in the interval $(-1, 1)$ and therefore substituting an upper bound on ρ_{ij} yields an upper bound on Ψ . This proof removed the obstacle in specializing the bound for linear codes. The following proposition aims at the same target, for the BSC.

Proposition 4.1 *The generalized triplets error probability for the BSC given in (4.2.12) is monotonically increasing in $w(\mathbf{c}_i \mathbf{c}_j)$ for any $0 \leq w(\mathbf{c}_i \mathbf{c}_j) \leq \min \{w(\mathbf{c}_i) - 1, w(\mathbf{c}_j) - 1\}$, $w(\mathbf{c}_i)$, $w(\mathbf{c}_j)$ and for any $\eta_i : \mathbb{Z}^+ \mapsto \mathbb{R}^+$.*

The complete proof is given in Appendix B.1. For some intuition, remember that Seguin's proof for the AWGN channel was based on proving that the derivative $\frac{\partial}{\partial \rho} \Psi(\rho, x, y)$ is positive. The proof of Proposition 4.1 resembles in the sense that discrete derivation is used. Referring to (4.2.4), it is clear that the demand for positive η_i is not restricting.

To utilize Proposition 4.1, define

$$\bar{w}(\mathbf{c}_i \mathbf{c}_j) \triangleq \min \left\{ w(\mathbf{c}_i), w(\mathbf{c}_j), \left\lfloor \frac{w(\mathbf{c}_i) + w(\mathbf{c}_j) - d}{2} \right\rfloor \right\}. \quad (4.2.14)$$

Since

$$w(\mathbf{c}_i \mathbf{c}_j) = \frac{w(\mathbf{c}_i) + w(\mathbf{c}_j) - d_H(\mathbf{c}_i, \mathbf{c}_j)}{2}, \quad (4.2.15)$$

it is clear that $w(\mathbf{c}_i \mathbf{c}_j) \leq \bar{w}(\mathbf{c}_i \mathbf{c}_j)$. Thus,

$$\begin{aligned} \tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j) &\leq P_{den}(w(\mathbf{c}_i), w(\mathbf{c}_j)) \\ &\triangleq \sum_{l=0}^{\bar{w}(\mathbf{c}_i \mathbf{c}_j)} \sum_{m=\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor + 1 - l}^{w(\mathbf{c}_i) - \bar{w}(\mathbf{c}_i \mathbf{c}_j)} \sum_{n=\lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor + 1 - l}^{w(\mathbf{c}_j) - \bar{w}(\mathbf{c}_i \mathbf{c}_j)} \sum_{k=0}^{N - w(\mathbf{c}_i) - w(\mathbf{c}_j) + \bar{w}(\mathbf{c}_i \mathbf{c}_j)} \binom{\bar{w}(\mathbf{c}_i \mathbf{c}_j)}{l} \binom{w(\mathbf{c}_i) - \bar{w}(\mathbf{c}_i \mathbf{c}_j)}{m} \\ &\cdot \binom{w(\mathbf{c}_j) - \bar{w}(\mathbf{c}_i \mathbf{c}_j)}{n} \binom{N - w(\mathbf{c}_i) - w(\mathbf{c}_j) + \bar{w}(\mathbf{c}_i \mathbf{c}_j)}{k} p^{l+m+n+k} (1-p)^{N-l-m-n-k} \eta_i^2(l+m+n+k). \end{aligned} \quad (4.2.16)$$

Now it is possible to derive a lower bound on the probability of error using only the code's weight spectrum. The summations can be carried out on the possible weights rather than the whole code. Thus, the new lower bound on the error probability of a linear code \mathcal{C} on the BSC is given by

$$P(\varepsilon) \geq \sum_{n=1}^N \frac{B_n P_{num}^2(n)}{P_{den}(n) + (B_n - 1)P_{den}(n, n) + \sum_{m=1, m \neq n}^N B_m P_{den}(n, m)}, \quad (4.2.17)$$

where $P_{num}(n)$, $P_{den}(n)$ and $P_{den}(n, m)$ include the function η_i , which can be optimized to yield the tightest bound.

4.2.2 Approximations for $deg(\mathbf{x}|\mathbf{c}_0)$

In a similar fashion to the analysis of Chapter 3, we seek a function η_i of the form

$$\eta_i(w(\mathbf{x})) = \frac{1}{\widetilde{deg}(w(\mathbf{x}))} \quad (4.2.18)$$

where $\widetilde{deg}(w(\mathbf{x}))$ is any approximation of $deg(\mathbf{x}|\mathbf{c}_0)$ whose dependance on \mathbf{x} is only through $w(\mathbf{x})$. The reason for preferring only this kind of approximation is not only to reduce the computational complexity, but also since the only knowledge available on the code is its weight enumeration, thus any knowledge on the *structure* of a specific word is rarely useful.

Referring to (4.2.4), $deg(\mathbf{x}|\mathbf{c}_0)$ is the number of words with Hamming weight less than $w(\mathbf{x})$ in the coset $\mathcal{C} + \mathbf{x}$. Thus, we are interested in the weight enumeration of this coset when the only knowledge on \mathbf{x} is $w(\mathbf{x})$. As a simple example, consider a 1-bit parity check code. Clearly, there are only two cosets in this case. The first is the code itself, i.e., the set of all even-weight words. The second is the set of all odd-weight words. Thus, given a received word \mathbf{x} , its weight is sufficient to identify the correct coset and $deg(\mathbf{x}|\mathbf{c}_0) = deg(w(\mathbf{x})|\mathbf{c}_0)$ is known exactly. However, the evaluation of the error probability for this code is trivial in the first place. There are several codes whose cosets weight enumeration can be found in the current literature. Yet, even for simple codes, the weight enumeration of $\mathcal{C} + \mathbf{x}$ cannot, in general, be evaluated using $w(\mathbf{x})$ alone. For example, in [27, pp. 170, Example 1] the

weight enumeration of Hamming(7,4) and its cosets is given. Words of weight 3 or 4 cannot be classified to a specific coset only by their weight, i.e., may belong to the code *or* one of its cosets.

We include here three possible approximations for $\text{deg}(\mathbf{x}|\mathbf{c}_0)$ using only the existing information on \mathbf{x} and the code's weight enumeration. The approximations given here characterize several of the methods which can be used to derive more approximations in future work.

Chernoff Bound for $\text{deg}(\mathbf{x}|\mathbf{c}_0)$

In this approximation, we view $\text{deg}(\mathbf{x}|\mathbf{c}_0)$ as $|\mathcal{C}|P(w(\mathbf{x} + \mathbf{c}) < w(\mathbf{x}))$. Let \mathbf{x} be fixed and \mathbf{c} a codeword chosen randomly with uniform distribution. Considering $w(\mathbf{x} + \mathbf{c})$ as a random variable, we have

$$P(w(\mathbf{x} + \mathbf{c}) < w(\mathbf{x})) \leq \mathbb{E} \left\{ e^{-a(w(\mathbf{x}+\mathbf{c})-w(\mathbf{x}))} \right\}, \quad (4.2.19)$$

where the expectation is over all possible codewords \mathbf{c} , and a is a non negative parameter to be optimized. Clearly,

$$\begin{aligned} \mathbb{E} \left\{ e^{-a(w(\mathbf{x}+\mathbf{c})-w(\mathbf{x}))} \right\} &= \sum_{\mathbf{c} \in \mathcal{C}} P(\mathbf{c}) e^{-a(w(\mathbf{x}+\mathbf{c})-w(\mathbf{x}))} \\ &= \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c} \in \mathcal{C}} e^{-a(w(\mathbf{x}+\mathbf{c})-w(\mathbf{x}))}. \end{aligned} \quad (4.2.20)$$

The expression in (4.2.20) does not depend solely on $w(\mathbf{x})$ and the weight enumeration of the code. Substituting the bound

$$w(\mathbf{x} + \mathbf{c}) \geq |w(\mathbf{x}) - w(\mathbf{c})| \quad (4.2.21)$$

in (4.2.20), we have

$$\begin{aligned} \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c} \in \mathcal{C}} e^{-a(w(\mathbf{x}+\mathbf{c})-w(\mathbf{x}))} &\leq \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c} \in \mathcal{C}} e^{-a(|w(\mathbf{x})-w(\mathbf{c})|-w(\mathbf{x}))} \\ &= \frac{e^{aw(\mathbf{x})}}{|\mathcal{C}|} \sum_{i=0}^N B_i e^{-a|w(\mathbf{x})-i|}. \end{aligned} \quad (4.2.22)$$

The approximation for $\text{deg}(\mathbf{x}|\mathbf{c}_0)$ is therefore

$$\widetilde{\text{deg}}(w(\mathbf{x})) = e^{aw(\mathbf{x})} \sum_{i=0}^N B_i e^{-a|w(\mathbf{x})-i|}, \quad (4.2.23)$$

where $a \geq 0$ is a parameter to be optimized.

Polynomial (Chernoff-Like) Bound for $\text{deg}(\mathbf{x}|\mathbf{c}_0)$

The previous subsection suggested an exponential bound for $P(w(\mathbf{x} + \mathbf{c}) < w(\mathbf{x}))$. In this subsection, we use the inequality

$$\mathbf{1}_{(x < w)} \leq \left(\frac{N-x}{N-w} \right)^a \quad w \in [0, N-1], \quad x \in [0, N], \quad (4.2.24)$$

where $\mathbf{1}_{(\cdot)}$ is the indicator function and $a \geq 0$ is a parameter to be optimized. Since $P(w(\mathbf{x} + \mathbf{c}) > N) = 0$, for any $w(\mathbf{x}) \neq N$ we have

$$\begin{aligned} P(w(\mathbf{x} + \mathbf{c}) < w(\mathbf{x})) &\leq \mathbb{E} \left\{ \left(\frac{N - w(\mathbf{x} + \mathbf{c})}{N - w(\mathbf{x})} \right)^a \right\} \\ &= \sum_{\mathbf{c} \in \mathcal{C}} P(\mathbf{c}) \left(\frac{N - w(\mathbf{x} + \mathbf{c})}{N - w(\mathbf{x})} \right)^a \\ &\stackrel{(a)}{\leq} \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c} \in \mathcal{C}} \left(\frac{N - |w(\mathbf{x}) - w(\mathbf{c})|}{N - w(\mathbf{x})} \right)^a \\ &= \frac{1}{|\mathcal{C}|} \sum_{i=0}^N B_i \left(\frac{N - |w(\mathbf{x}) - i|}{N - w(\mathbf{x})} \right)^a, \end{aligned} \quad (4.2.25)$$

where (a) is due to (4.2.21). Thus, a new approximation for $\text{deg}(\mathbf{x}|\mathbf{c}_0)$ is

$$\widetilde{\text{deg}}(w(\mathbf{x})) = \begin{cases} \sum_{i=0}^N B_i \left(\frac{N - |w(\mathbf{x}) - i|}{N - w(\mathbf{x})} \right)^a & w(\mathbf{x}) \neq N \\ |\mathcal{C}| & w(\mathbf{x}) = N, \end{cases} \quad (4.2.26)$$

where $a \geq 0$ is a parameter to be optimized.

Curve Approximation for $\text{deg}(\mathbf{x}|\mathbf{c}_0)$

In this subsection, we consider a different method to approximate $\text{deg}(\mathbf{x}|\mathbf{c}_0)$. Clearly,

$$\text{deg}(\mathbf{x}) = \begin{cases} 0 & w(\mathbf{x}) \leq \lfloor \frac{d-1}{2} \rfloor \\ \sum_{i=0}^{w(\mathbf{x})-1} B_i^{\mathbf{x}} & \lfloor \frac{d-1}{2} \rfloor < w(\mathbf{x}) < N \\ |\mathcal{C}| - 1 & w(\mathbf{x}) = N, \end{cases} \quad (4.2.27)$$

where $B_i^{\mathbf{x}}$ is the number of words of weight i in the coset $\mathcal{C} + \mathbf{x}$. Thus, $\text{deg}(\mathbf{x}|\mathbf{c}_0)$ is monotonically increasing in $w(\mathbf{x})$, with known values for $w(\mathbf{x}) = \lfloor \frac{d-1}{2} \rfloor$ and $w(\mathbf{x}) = N$. By choosing any concave or convex line between these two points we have the following approximation

$$\widetilde{\text{deg}}(w(\mathbf{x})) = \begin{cases} (|\mathcal{C}| - 1) \left(\frac{w(\mathbf{x}) - \lfloor \frac{d-1}{2} \rfloor}{N - \lfloor \frac{d-1}{2} \rfloor} \right)^a & w(\mathbf{x}) > \lfloor \frac{d-1}{2} \rfloor \\ 0 & \text{else,} \end{cases} \quad (4.2.28)$$

where $a \geq 0$ is a parameter to be optimized. This approximation is easier to evaluate than the previous ones since no summation is required. Moreover, only the size of the code, its length and its minimum distance are used.

4.2.3 Lower Bounds Using the Subcode \mathcal{C}_i^* and the Code's Covering Radius

In this section, we consider two variations on the bound given in (4.2.17). These two variations, suggested in [5], will both reduce the complexity of the bound and improve performance.

Definition 4.2 ([27], [5]) *Denote by t the covering radius of the code \mathcal{C}*

$$t = \max_{\mathbf{f} \in GF(2)^N} \min_{\mathbf{c} \in \mathcal{C}} d_H(\mathbf{f}, \mathbf{c}). \quad (4.2.29)$$

Namely, t is the maximum number of errors that can be corrected. More than t errors cannot be corrected.

Clearly,

$$P_{\mathcal{C}}(\varepsilon) = P_{\mathcal{C}}(\varepsilon, w(\mathbf{x}) \leq t) + P(w(\mathbf{x}) > t). \quad (4.2.30)$$

Thus, when evaluating a bound on the error probability, we may assume no more than t errors were made, then add the probability of *all* the words with weight higher than t . As noted in [5], when lower bounds on the error probability are discussed, an upper bound $M \geq t$ can be used if t is not known. Note the resemblance to Poltyrev's ([6]) technique, as

given in (4.1.3). Poltyrev chose \mathcal{A}^c as the set of all words with weight higher than a certain threshold. However, this threshold is optimized to yield the tightest *upper* bound, and the resulting value is always *smaller* than the covering radius ([6]).

Finally, as in Section 3.3.3, since

$$P_C(\varepsilon, w(\mathbf{x}) \leq M) \geq P_{C_i^*}(\varepsilon, w(\mathbf{x}) \leq M), \quad (4.2.31)$$

where $C_i^* = \{\mathbf{c} \in \mathcal{C} : w(\mathbf{c}) = i\} \cup \{\mathbf{c}_0\}$, we may compute the bound disregarding all codewords of weight other than i . Although the numerical analysis shows that best results are achieved with $i = d$, we prefer this general form for future reference. Consequently, we have the following proposition.

Proposition 4.3 *Let \mathcal{C} be any linear code over $GF(2)$ of length N and minimum distance d . Let B_i be the number of codewords of hamming weight i , $d \leq i \leq N - \lceil \frac{d}{2} \rceil$. Denote by $P(\varepsilon)$ the decoding error probability on a BSC with crossover probability p . We have*

$$P(\varepsilon) \geq LB_i(\eta_i, p) \triangleq \frac{B_i \bar{P}_{num}^2(i)}{\bar{P}_{den}(i) + (B_i - 1) \bar{P}_{den}(i, i)} + P_M, \quad (4.2.32)$$

where

$$\bar{P}_{num}(i) \triangleq \sum_{l=\lfloor \frac{i}{2} \rfloor + 1}^i \sum_{m=0}^{M-l} \binom{i}{l} \binom{N-i}{m} p^{l+m} (1-p)^{N-l-m} \eta_i(l+m), \quad (4.2.33)$$

$$\bar{P}_{den}(i) \triangleq \sum_{l=\lfloor \frac{i}{2} \rfloor + 1}^i \sum_{m=0}^{M-l} \binom{i}{l} \binom{N-i}{m} p^{l+m} (1-p)^{N-l-m} \eta_i^2(l+m), \quad (4.2.34)$$

$$\begin{aligned} \bar{P}_{den}(i, i) \triangleq & \sum_{l=0}^{i-\lceil \frac{d}{2} \rceil} \sum_{m=\lfloor \frac{i}{2} \rfloor + 1 - l}^{\lceil \frac{d}{2} \rceil} \sum_{n=\lfloor \frac{i}{2} \rfloor + 1 - l}^{\lceil \frac{d}{2} \rceil} \sum_{k=0}^{M-l-m-n} \binom{i-\lceil \frac{d}{2} \rceil}{l} \binom{\lceil \frac{d}{2} \rceil}{m} \binom{\lceil \frac{d}{2} \rceil}{n} \binom{N-i-\lceil \frac{d}{2} \rceil}{k} \\ & \cdot p^{l+m+n+k} (1-p)^{N-l-m-n-k} \eta_i^2(l+m+n+k), \end{aligned} \quad (4.2.35)$$

$$P_M \triangleq \sum_{l=M+1}^N \binom{N}{l} p^l (1-p)^{N-l} \quad (4.2.36)$$

and $\eta_i : \mathbb{Z}^+ \mapsto \mathbb{R}^+$ is any function to be optimized.

Note that the demand $i \leq N - \lceil \frac{d}{2} \rceil$ is not restricting. Since there is no more than one codeword with weight higher than $N - \lceil \frac{d}{2} \rceil$, a subcode \mathcal{C}_i^* for such i is degenerated.

Proof (Proposition 4.3). Based on the preceding discussion, substitute $B_n = 0$ for every $n \neq i$ in (4.2.17). To use the covering radius of the code, evading high values of $w(\mathbf{x})$, we may alter the expressions in (4.2.8), (4.2.11) and (4.2.16) to include only words \mathbf{x} with weight $w(\mathbf{x}) \leq M$ by changing the upper bound of the last summation in each expression. P_M is the probability of more than M bit errors. \square

Remark 4.4 Note that when $\lfloor \frac{i}{2} \rfloor + 1 - l > \lceil \frac{d}{2} \rceil$ the sums over m and n in (4.2.35) are empty. Thus, the value of $\bar{P}_{den}(i, i)$ is unchanged if we sum over $\lfloor \frac{i}{2} \rfloor + 1 - \lceil \frac{d}{2} \rceil \leq l \leq i - \lceil \frac{d}{2} \rceil$ instead of $0 \leq l \leq i - \lceil \frac{d}{2} \rceil$.

To choose a proper η_i , we return to Section 4.2.2. Although the approximations therein refer to the bound given in (4.2.17), i.e., when the whole code is used, we find them useful in (4.2.32) for two main reasons. First, since η_i defines only the essence of behavior, and not necessarily exact values (refer to Section 2.3), the approximations in Section 4.2.2 may be sufficient as are. Second, even if several variations are required, the *methods* suggested in Section 4.2.2 are still applicable. These facts are also supported by numerical results. For example, in the numerical analysis, where the bound given in Proposition 4.3 is utilized, the following variation of approximation (4.2.28) was used

$$\widetilde{deg}(w(\mathbf{x})) = \begin{cases} (B_d - 1) \left(\frac{w(\mathbf{x}) - \lfloor \frac{d-1}{2} \rfloor}{N - \lceil \frac{d-1}{2} \rceil} \right)^a & w(\mathbf{x}) > \lceil \frac{d-1}{2} \rceil \\ 0 & \text{else.} \end{cases} \quad (4.2.37)$$

Before we conclude, we refer to Keren and Litsyn's bound [5]. The bound presented in [5] is based on the same techniques, namely, de Caen's bound [1], the subset \mathcal{C}_d^* is used and words with weight higher than the covering radius are considered erroneous. However, even when $\eta_i(w) \equiv 1$ the bound given in Proposition 4.3 is not identical to [5]. The major difference is the fact that in [5] the set $\{\mathbf{x} \in GF(2)^N : \exists_i w(\mathbf{x} + \mathbf{c}_i) < w(\mathbf{x}), w(\mathbf{x}) \leq t\}$

is partitioned to constant weight subsets and de Caen's bound is employed to each subset separately. This partition simplifies several computations and instead of Proposition 4.1 a more ad hoc approach can be used. However, in this way the usage of Theorem 2.1 instead of de Caen's bound is burdensome. Remember that the approximations suggested in Section 4.2.2 depend only on $w(\mathbf{x})$, thus if used in [5], where de Caen's bound is employed on constant weight subsets, they immediately factor out as constant multipliers. Hence, the method suggested here is more appropriate when Theorem 2.1 is to be used. It is also important to note that the bound in [5] is easier to evaluate since the summations required are simpler and no optimization is needed.

Finally, analogous to the bound for the AWGN channel with high values of E_b/N_0 , and with compliance with Remark 2.3, we refer to the asymptotic tightness of the bound presented in this chapter. Let $UB(p)$ be the union bound for the BSC

$$P(\varepsilon) \leq UB(p) \triangleq \sum_{i=1}^N B_i P(i), \quad (4.2.38)$$

where $P(i)$ is the pairwise error probability, i.e., the probability that a codeword of weight i is closer to the received word than the all-zero codeword. $P(i)$ is given by

$$P(i) = \sum_{l=\lfloor \frac{i}{2} \rfloor + 1}^i \binom{i}{l} p^l (1-p)^{i-l}. \quad (4.2.39)$$

Under this definition, we have the following proposition.

Proposition 4.5 *Consider the bound in Proposition 4.3 when $i = d$. Then, for any function $\eta_d > 0$ which is independent of p we have*

$$\lim_{p \rightarrow 0} \frac{LB_d(\eta_d, p)}{UB(p)} = 1. \quad (4.2.40)$$

Namely, the bound is tight for $p \rightarrow 0$.

Note that the condition on η_d is not restricting in any way since we know that any reasonable η_d is independent of the channel's crossover probability and satisfies $\eta_d(l) \geq 1$ for each $l > \lfloor \frac{d-1}{2} \rfloor$ (refer to (4.2.4)). The proof of Proposition 4.5 can be found in Appendix B.2.

4.3 Results

We compare the new lower bound given in Proposition 4.3 with several known bounds in the current literature, when the linear code used is $BCH(63, 24)$. Figure 4.1 includes the new bound with $i = d$ and the approximation (4.2.37). Other approximations given in Section 4.2.2 or their variations yield slightly inferior results. For reference, three bounds are plotted: Poltyrev's upper bound [6], Keren and Litsyn's lower bound [5] and the sphere packing lower bound (as it is given in [6]). The new bound with the trivial choice of η_i is not plotted since the results are very similar to Keren and Litsyn's. The *two codewords* bound is not given since it is proved in [26] to be inferior to Keren and Litsyn's bound for every value of p , the channel's crossover probability. The BSC version of Kounias' bound, which is the analogue of (3.3.30), is not plotted since it is the same as the new bound and Keren and Litsyn's bound for low values of p , and inferior to both for high values.

The new bound is at least as good as Keren and Litsyn's bound for every value of p . The improvement is obvious for high values of p ($3.8dB@p = 0.1$). However, for lower, and more realistic, values of p , where Keren and Litsyn's bound is superior to the sphere packing bound, the improvement is scarce ($0.7dB@p = 0.03$). Nevertheless, it is clear that a non trivial choice of the optimizing function η_i yields better results. Moreover, in the next chapter we prove that a non trivial η_i may yields a strictly tighter *error exponent* and identify the optimal choice of η_i .

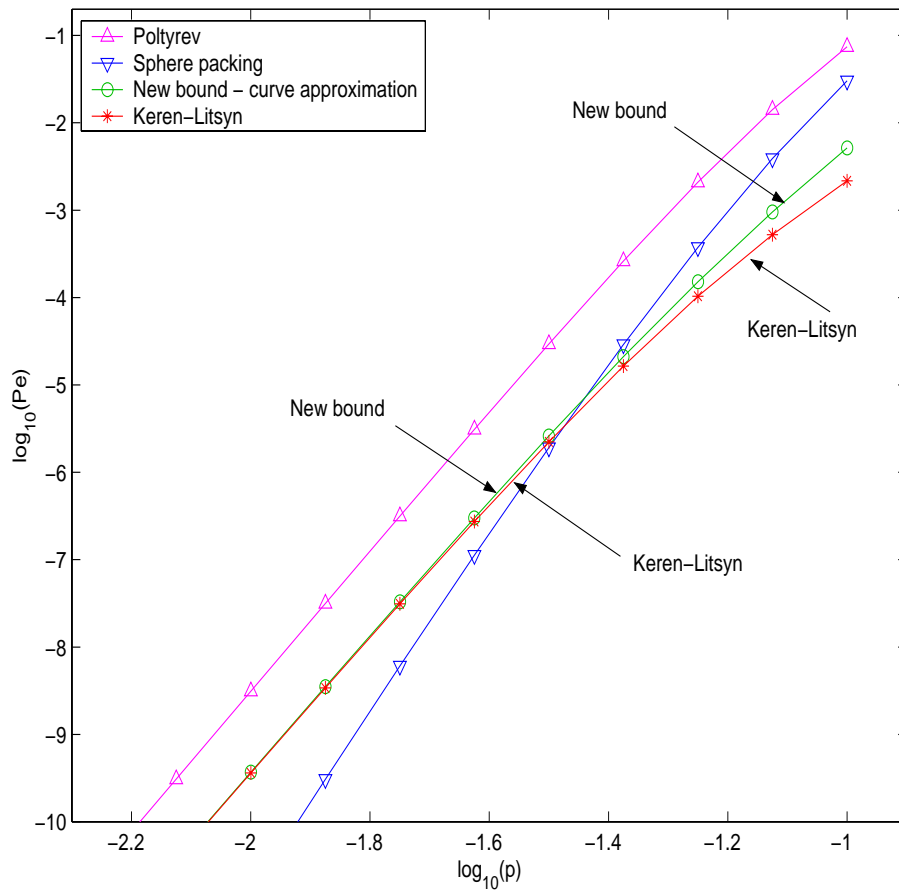


Figure 4.1: Bounds on the decoding error probability of BCH(63,24) code, BSC. The new bound, based on the approximation given in (4.2.28) is given. For reference, three bounds are plotted: Poltyrev's upper bound, Keren and Litsyn's lower bound and the sphere packing lower bound.

Chapter 5

Upper Bounds on the Error Exponent

In this chapter, we calculate an upper bound on the error exponent resulting from the bounds given in Propositions 3.5 and 4.3. We start with the bound in Proposition 4.3, and prove that a non trivial choice of η_i may result in a tighter bound on the error exponent. We also identify the optimal choice of η_i . Using similar analysis, we compute the bound on the error exponent resulting from the *dot product bound*, given in Proposition 3.5, and show that a non trivial choice of the parameter a may result in an exponentially tighter bound.

5.1 Introduction

In this section, a short literature survey is given. We focus on bounds on the error exponent of a Discrete Memoryless Channel (DMC). The bounds presented here apply to the BSC, and are easily extended to the AWGN channel as well.

Consider the case of uniform signaling over a DMC and maximum likelihood decoding. It is well known the the probability of error in this case decreases exponentially to zero, as long as the rate of the code does not exceed a certain threshold, C , the channel capacity.

Namely,

$$e^{-NE_L(R)+o(N)} \leq P(\varepsilon) \leq e^{-NE(R)+o(N)}, \quad (5.1.1)$$

where $o(N)/N$ tends to zero as N grows and both $E_L(R)$ and $E(R)$ are strictly positive for $R < C$. The exponential rate, however, is known exactly only for rates higher than a given threshold or zero rate, i.e., only in these cases $E_L(R) = E(R)$. For low rates, $E_L(R)$ and $E(R)$ differ, thus provide only upper and lower bounds on the exact error exponent.

We give here a small review of bounds on the error exponent. A more detailed survey of these bounds can be found in [11] (BSC), [12] (AWGN) and [20]. The best known lower bounds on the error exponent are due to Gallager, [8]. The lower bound for high rates, derived by methods of random coding, coincide with the sphere packing bound by Shannon, Gallager and Berlekamp, [9], for rates higher than R_c , thus yielding the exact error exponent for these rates. For low rates, a tighter lower bound than the random coding bound was derived by methods of expurgation, also appearing in [8]. As for upper bounds, tighter bounds than the sphere packing bound for low rates and zero rate were also derived in [9]. The main idea in improving the upper bounds for low rates was the fact, proved in [9], that any straight line between a low rate upper bound and the sphere packing bound is also an upper bound. As the low rate upper bound, the two codewords bound ([7], [9]) is usually used. For binary input channels¹, however, the upper bounds for low rates were further tightened by McEliece and Omura. In [10], McEliece and Omura used an improved upper bound on the minimum distance of codes, derived by McEliece, Rodemich, Rumsey and Welch ([27, pp. 559]) and the straight line bound, to tighten the bound in [9] for low rates. The latest upper on the error exponent of the BSC was derived by Litsyn in [11]. The essence of this bound is a new bound on the distance distribution of codes, and not an improvement of McEliece-Rodemich-Rumsey-Welch's bound, as might have been expected. The latest upper bound on the error exponent of the AWGN was derived by Burnashev in [12]. Burnashev showed that by extending the range in which the union bound analysis

¹This includes the BSC as well as the AWGN channel when any binary modulation is used.

applies, together with a bound on the distance distribution of codes, the bound on the error exponent can be tightened. We note here that for random codes, random linear codes and *typical* codes out of these ensembles the error exponent is known exactly [28].

In this chapter, we calculate the upper bounds on the error exponent resulting from the bounds in the previous chapters. We show that in certain cases, these bound may result in tighter bounds on the error exponent than the de Caen-based bounds. We give several examples for the bounds discussed here. These examples are based on weight distributions appearing in the current literature, namely, the weight distribution of a *typical linear code*, appearing in [28] and the weight distribution of codes with exponentially many minimum distance codewords, appearing in [29]. It is important to mention, however, that only bounds for specific codes are discussed, and not bounds on the error exponent of the BSC or AWGN in general. To use the bounds derived here as bounds on the error exponent, bounds on the weight distribution of any code are required².

5.2 Preliminaries

In this section, we introduce the required notations relevant for this chapter.

Let $\{\mathcal{C}_N\}$ be any sequence of codes, each of which is of length N and minimum distance d_N . For every $d_N < i \leq N$, denote by δ_i the ratio $\frac{i}{N}$. Let B_i^N be the number of codewords of weight i in each code. We consider only sequences of codes for which the limits $\lim_{N \rightarrow \infty} \frac{1}{N} \log B_i^N$ and $\lim_{N \rightarrow \infty} \frac{d_N}{N}$ exists, and denote their values by $E_B^{\delta_i}$ and δ_d , respectively. Hereafter, the base of the logarithm is 2.

Definition 5.1 *Let $F(N)$ and $G(N)$ be any two functions. If*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log F(N) = \lim_{N \rightarrow \infty} \frac{1}{N} \log G(N) \quad (5.2.1)$$

²In this case, the codes are not restricted to be linear. However, a subcode with a “linearity property” is used, i.e., a subcode in which every codeword has essentially the same distance spectrum. For examples, see [11] or [12].

we write $F(N) \doteq G(N)$. If

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log F(N) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \log G(N) \quad (5.2.2)$$

we write $\frac{1}{N} \log F(N) \preceq \frac{1}{N} \log G(N)$, namely, $F(N)$ is exponentially smaller than $G(N)$.

In this chapter, we are interested in upper bounds on the error exponent, in the form of (5.2.2), valid for any linear code \mathcal{C} of rate R and a given weight enumeration.

5.3 Upper Bounds on the Error Exponent for the Binary Symmetric Channel

In this section, we calculate the upper bound on the error exponent resulting from the bound in Proposition 4.3 and analyze the results. We substitute $M = N$ in equations (4.2.32) to (4.2.36), i.e., assume no knowledge on the covering radius is available. This assumption weakens the bound, especially at high values of p , however, it simplifies computations and allows us to focus on identifying the best choice of η_i . As long as the first summand in the r.h.s. of (4.2.32) dominates, a non trivial covering radius only causes a difference in the optimization ranges defined below, hence the essence of the technique proposed herein is not changed. Thus, for any $d_N \leq i \leq N - \left\lceil \frac{d_N}{2} \right\rceil$, the following bound is considered

$$P(\varepsilon) \geq \frac{B_i^N P_{num}^2(i)}{P_{den}(i) + (B_i^N - 1)P_{den}(i, i)}, \quad (5.3.1)$$

where $P_{num}(i)$, $P_{den}(i)$ and $P_{den}(i, i)$ were defined in (4.2.8), (4.2.11) and (4.2.16), respectively.

5.3.1 Main Results

For easy reference and facile understanding of this section, we briefly introduce the outline of the analysis and summarize the main results.

Consider the bound in (5.3.1). We wish to calculate the resulting bound on the error exponent, and to identify the optimal choice of the function η_i . Clearly, since the denominator of the r.h.s. of (5.3.1) is a sum of two expressions, the exponential behavior of the bound in (5.3.1) depends on which expression dominates. In the first part of the analysis, whose results are given by Proposition 5.2, we show that this observation translates to a *condition on the code*, which determines the value of the new bound on the error exponent in each case.

In the second part of the analysis, whose results are given by Corollary 5.3 and the discussion proceeding it, we analyze the condition on the code and the resulting bound on the error exponent when this condition is satisfied. It is shown therein, that if the difference between the triplets error exponent and the pairwise error exponent is not too small (i.e., the rate of the code is not too large), the condition on the code is satisfied, and the resulting bound on the error exponent is given by

$$-\frac{1}{N} \log P(\varepsilon) \preceq -\delta_i \log \sqrt{4p(1-p)} - E_B^{\delta_i}, \quad (5.3.2)$$

for any $\delta_d \leq \delta_i \leq 1 - \frac{1}{2}\delta_d$. Namely, one can use the union bound to derive a valid lower bound on the error probability (in this case, we say that the *union bound analysis applies*).

This far, we have not considered the choice of the function η_i . Our main result, given by Proposition 5.4, is that while it can be easily proved that when the condition on the code is satisfied, the trivial η_i is optimal, this is not the case when it is not satisfied. In this case, a non trivial η_i can extend the range of rates for which the union bound analysis applies, thus achieving a tighter bound on the error exponent. The optimal value of η_i , the range of rates for which the union bound analysis applies and a quantification of the improvement over the bound with trivial η_i are given in Proposition 5.4 and the discussion which follows.

5.3.2 Analysis

In this subsection, we include the detailed derivations required to achieve the results discussed earlier. We start with several definitions. For $t = \delta_i N$, define

$$E_\eta^{\delta_i}(\delta_t) = \lim_{N \rightarrow \infty} -\frac{1}{N} \log \eta_{\delta_i N}(t). \quad (5.3.3)$$

Since $\eta_i(t)$ is any function to be optimized, we may reduce the set of possible functions to assure that the limit in (5.3.3) exists. For $\eta_i(t) \equiv 1$ we have $E_\eta^{\delta_i}(\delta_t) \equiv 0$. Analogously to the previous chapters, we denote this case as the *trivial* choice of $E_\eta^{\delta_i}(\delta_t)$. For any δ_d and $\delta_d \leq \delta_i \leq 1 - \frac{1}{2}\delta_d$, define the following regions in $[0, 1]^2$ and $[0, 1]^4$, respectively

$$\mathcal{D}_1 = \left\{ (\delta_l, \delta_m) \in [0, 1]^2 : \frac{\delta_i}{2} \leq \delta_l \leq \delta_i, 0 \leq \delta_m \leq 1 - \delta_i \right\} \quad (5.3.4)$$

and

$$\mathcal{D}_2 = \left\{ (\delta_l, \delta_m, \delta_n, \delta_k) \in [0, 1]^4 : \frac{1}{2}(\delta_i - \delta_d) \leq \delta_l \leq \delta_i - \frac{\delta_d}{2}, \frac{\delta_i}{2} - \delta_l \leq \delta_m \leq \frac{\delta_d}{2}, \right. \\ \left. \frac{\delta_i}{2} - \delta_l \leq \delta_n \leq \frac{\delta_d}{2}, 0 \leq \delta_k \leq 1 - \delta_i - \frac{\delta_d}{2} \right\}. \quad (5.3.5)$$

Let $H(x)$ be the binary entropy function

$$H(x) = -x \log(x) - (1-x) \log(1-x). \quad (5.3.6)$$

Define

$$E_1^{\delta_i}(\delta_l, \delta_m, p) \triangleq -\delta_i H\left(\frac{\delta_l}{\delta_i}\right) - (1 - \delta_i) H\left(\frac{\delta_m}{1 - \delta_i}\right) + (\delta_l + \delta_m) \log\left(\frac{1-p}{p}\right) - \log(1-p) \quad (5.3.7)$$

and

$$E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p) \triangleq -(\delta_i - \delta_d/2) H\left(\frac{\delta_l}{\delta_i - \delta_d/2}\right) - \frac{\delta_d}{2} H\left(\frac{\delta_m}{\delta_d/2}\right) - \frac{\delta_d}{2} H\left(\frac{\delta_n}{\delta_d/2}\right) \\ - (1 - \delta_i - \delta_d/2) H\left(\frac{\delta_k}{1 - \delta_i - \delta_d/2}\right) + (\delta_l + \delta_m + \delta_n + \delta_k) \log\left(\frac{1-p}{p}\right) - \log(1-p). \quad (5.3.8)$$

Under these definitions, we have the following proposition.

Proposition 5.2 Let $\{\mathcal{C}_N\}$ be a sequence of codes for the BSC. Let p be the crossover probability of the channel. Then, for any $\delta_d \leq \delta_i \leq 1 - \frac{1}{2}\delta_d$, and for any piecewise continuous function $E_\eta^{\delta_i} : [0, 1] \mapsto \mathbb{R}^+$, we have

$$-\frac{1}{N} \log P(\varepsilon) \preceq 2 \min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(\delta_l, \delta_m, p) + E_\eta^{\delta_i}(\delta_l + \delta_m) \right\} - E_B^{\delta_i} \\ - \min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(\delta_l, \delta_m, p) + 2E_\eta^{\delta_i}(\delta_l + \delta_m) \right\} \quad (5.3.9)$$

if

$$E_B^{\delta_i} \leq \min_{\mathcal{D}_2} \left\{ E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p) + 2E_\eta^{\delta_i}(\delta_l + \delta_m + \delta_n + \delta_k) \right\} \\ - \min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(\delta_l, \delta_m, p) + 2E_\eta^{\delta_i}(\delta_l + \delta_m) \right\} \quad (5.3.10)$$

and

$$-\frac{1}{N} \log P(\varepsilon) \preceq 2 \min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(\delta_l, \delta_m, p) + E_\eta^{\delta_i}(\delta_l + \delta_m) \right\} \\ - \min_{\mathcal{D}_2} \left\{ E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p) + 2E_\eta^{\delta_i}(\delta_l + \delta_m + \delta_n + \delta_k) \right\} \quad (5.3.11)$$

otherwise.

The condition in (5.3.10) is a condition on the code's parameter³ $E_B^{\delta_i}$ (hereafter referred to as the *condition on the code*). The essence of Proposition 5.2, is the fact that the new bound on the error exponent is given by one of two different expressions, corresponding to the cases where condition (5.3.10) is satisfied or not. As mentioned earlier, this fact can easily be understood by observing that the denominator of (5.3.1) is the sum of two expressions. The resulting bound on the error exponent depends on which expression dominates. The complete proof of Proposition 5.2 is given in Appendix C.1.

We first analyze the case where the sequence of codes satisfies condition (5.3.10). In Appendix C.1, we show that $\min_{\mathcal{D}_1} E_1^{\delta_i}(\delta_l, \delta_m, p)$ is the error exponent for the pairwise

³We will see later that this is the condition for applying the union bound analysis. As noted in [12], this condition can be referred to as a condition on the code's parameter $E_B^{\delta_i}$, or a condition on the code's rate, R .

error probability $P(\varepsilon_i | \mathbf{c}_0)$, while $\min_{\mathcal{D}_2} E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p)$ is the error exponent for triplets $P(\varepsilon_i \cap \varepsilon_j | \mathbf{c}_0)$. Hence, we expect to have

$$\min_{\mathcal{D}_2} E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p) \geq \min_{\mathcal{D}_1} E_1^{\delta_i}(\delta_l, \delta_m, p), \quad (5.3.12)$$

which means that the set of possible functions $E_\eta^{\delta_i}$ for which the r.h.s. of (5.3.10) is nonnegative is not empty (it includes at least the trivial choice). In such a case, condition (5.3.10) is not trivial, and the following corollary is constructive.

Corollary 5.3 *Let $\delta_i, \delta_d \leq \delta_i \leq 1 - \frac{1}{2}\delta_d$, be fixed. Let \mathcal{S}_η be a set of functions $E_\eta^{\delta_i} : [0, 1] \mapsto \mathbb{R}$, indexed by η , which includes the trivial choice. Suppose that $\{\mathcal{C}_N\}$ is a sequence of codes for which condition (5.3.10) is satisfied for every value of $p \in \mathcal{P}$, for some $\mathcal{P} \subseteq (0, \frac{1}{2})$, and for every choice of $E_\eta^{\delta_i} \in \mathcal{S}_\eta$. Then, for every $p \in \mathcal{P}$, the trivial choice of $E_\eta^{\delta_i}$ minimizes the upper bound in (5.3.9) over all choices of $E_\eta^{\delta_i} \in \mathcal{S}_\eta$, and we have*

$$-\frac{1}{N} \log P(\varepsilon) \preceq -\delta_i \log \sqrt{4p(1-p)} - E_B^{\delta_i}. \quad (5.3.13)$$

Proof. When condition (5.3.10) is satisfied and $E_\eta^{\delta_i}(\delta_t) \equiv 0$, we have

$$-\frac{1}{N} \log P(\varepsilon) \preceq \min_{\mathcal{D}_1} E_1^{\delta_i}(\delta_l, \delta_m, p) - E_B^{\delta_i}. \quad (5.3.14)$$

Subtracting the r.h.s. of (5.3.14) from the r.h.s. of (5.3.9), we have

$$\begin{aligned} & 2 \min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(l, m, p) + E_\eta^{\delta_i}(l+m) \right\} - E_B^{\delta_i} - \min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(l, m, p) + 2E_\eta^{\delta_i}(l+m) \right\} \\ & \quad - \left(\min_{\mathcal{D}_1} E_1^{\delta_i}(l, m, p) - E_B^{\delta_i} \right) \\ & = \min_{\mathcal{D}_1} \left\{ 2E_1^{\delta_i}(l, m, p) + 2E_\eta^{\delta_i}(l+m) \right\} \\ & \quad - \left(\min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(l, m, p) + 2E_\eta^{\delta_i}(l+m) \right\} + \min_{\mathcal{D}_1} E_1^{\delta_i}(l, m, p) \right) \geq 0 \end{aligned} \quad (5.3.15)$$

for any $E_\eta^{\delta_i} \in \mathcal{S}_\eta$. Thus, when condition (5.3.10) is satisfied, $E_\eta^{\delta_i}(\delta_t) \equiv 0$ is the optimal choice and (5.3.14) is the resulting bound. Finally, in Appendix C.2 we show that

$$\min_{\mathcal{D}_1} E_1^{\delta_i}(\delta_l, \delta_m, p) = -\delta_i \log \sqrt{4p(1-p)}, \quad (5.3.16)$$

and (5.3.13) immediately follows. \square

At this point, several remarks are in order. The bound

$$-\frac{1}{N} \log P(\varepsilon) \preceq \min_{\mathcal{D}_1} E_1^{\delta_i}(\delta_l, \delta_m, p) = -\delta_i \log \sqrt{4p(1-p)} \quad (5.3.17)$$

is the well known *two codewords* bound, where $-\log \sqrt{4p(1-p)}$ is the *Bhattacharyya distance* for the BSC (see, for example, [20, pp. 88]). However, the bound in (5.3.13) implies that, under certain conditions, when there are exponentially many codewords of weight i , the exponent $E_B^{\delta_i}$ can be subtracted, yielding a tighter upper bound. This is to say that a union bound analysis results in a valid upper bound on the error exponent (a lower bound on the error probability). Thus, by optimizing the bound on δ_i , i.e., choosing the correct subcode, the union bound analysis gives the true error exponent for the code⁴. As noted in [28], the fact that union bound analysis yields the true error exponent for random codes is well known. In [30], using a lower bound of the form of $P_1 - P_2$, where P_1 is an upper bound, Gallager proved that union bound analysis gives the true error exponent for random codes by proving that P_2 decays exponentially faster than P_1 . Barg and Forney used this argument in [28] to derive the exact error exponents for typical codes from Shannon's random code ensemble as well as typical codes from a random linear code ensemble. Yet, the bound in (5.3.13) is valid for any given code, as long as condition (5.3.10) is satisfied.

Till this point, the results given by Proposition 5.2 were analyzed only as long as the condition on the code is satisfied. The main statement in Corollary 5.3 is that the union bound analysis is applicable in this case. However, note that the r.h.s. of (5.3.10) includes the function $E_\eta^{\delta_i}$, which can be optimized. The most important result of this chapter, as we will see below, is that by choosing a non trivial $E_\eta^{\delta_i}$, the range of rates for which the union bound analysis is applicable can be widened.

⁴The union bound, given by $P(\varepsilon|c_0) \leq \sum_{w=1}^N B_w P(\varepsilon_{0w}|c_0)$, has only polynomially many summands. Merely one of them determines the exponential behavior. Consequently, if we calculate a lower bound on the error probability using this subcode, and find out that the union bound analysis applies, this is the true exponential behavior. In this context, it is clear that if our choice of $E_\eta^{\delta_i}$ yields the true error exponent, no other $E_\eta^{\delta_i}$ is required.

To consider the case where the condition on the code is not satisfied, and discuss non trivial choices of $E_\eta^{\delta_i}$, the minimization $\min_{\mathcal{D}_2} E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p)$ should be discussed. In Appendix C.3, we show that

$$\min_{\mathcal{D}_2} E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p) = E_2^{\delta_i}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p), \quad (5.3.18)$$

where

$$\delta_{m_2} = \delta_i/2 - \delta_{l_2}, \quad (5.3.19)$$

$$\delta_{n_2} = \delta_i/2 - \delta_{l_2}, \quad (5.3.20)$$

$$\delta_{k_2} = p \left(1 - \delta_i - \frac{\delta_d}{2} \right), \quad (5.3.21)$$

and δ_{l_2} is the only root (with respect to δ_l) of the following cubic equation

$$\frac{\delta_l \left(\frac{\delta_d}{2} - \frac{\delta_i}{2} + \delta_l \right)^2}{\left(\frac{\delta_i}{2} - \delta_l \right)^2 \left(\delta_i - \frac{\delta_d}{2} - \delta_l \right)} = \frac{1-p}{p}, \quad (5.3.22)$$

such that $(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}) \in \mathcal{D}_2$. Since this solution is rather cumbersome to analyze, we handle here only the special case where $\delta_i = \delta_d$, namely, the subcode \mathcal{C}_d^* is used. In this case, equation (5.3.22) has a simple solution and our course of action and choice of $E_\eta^{\delta_i}$ becomes clearer. The general case is analogous, and yields similar results. We return to it at the end of this section.

When $\delta_i = \delta_d$, equation (5.3.22) simplifies to

$$\left(\frac{\delta_l}{\frac{\delta_d}{2} - \delta_l} \right)^3 = \frac{1-p}{p}, \quad (5.3.23)$$

yielding the following solution to the minimization of $E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p)$ over \mathcal{D}_2

$$\begin{aligned} \delta_{l_2} &= \frac{\frac{\delta_d}{2}}{1 + \sqrt[3]{\frac{p}{1-p}}}, & \delta_{m_2} &= \frac{\delta_d}{2} - \delta_{l_2} = \frac{\frac{\delta_d}{2} \sqrt[3]{\frac{p}{1-p}}}{1 + \sqrt[3]{\frac{p}{1-p}}}, \\ \delta_{k_2} &= p \left(N - \frac{3\delta_d}{2} \right), & \delta_{n_2} &= \frac{\delta_d}{2} - \delta_{l_2} = \frac{\frac{\delta_d}{2} \sqrt[3]{\frac{p}{1-p}}}{1 + \sqrt[3]{\frac{p}{1-p}}}, \end{aligned} \quad (5.3.24)$$

From Appendix C.2, the solution to the minimization of $E_1^{\delta_d}(\delta_l, \delta_m, p)$ over \mathcal{D}_1 is

$$\delta_{l_1} = \frac{\delta_d}{2}, \quad \delta_{m_1} = p(1 - \delta_d). \quad (5.3.25)$$

Define $C(E_B^{\delta_d}, p)$ as

$$C(E_B^{\delta_d}, p) \triangleq E_B^{\delta_d} - \left(E_2^{\delta_d}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p) - E_1^{\delta_d}(\delta_{l_1}, \delta_{m_1}, p) \right) \quad (5.3.26)$$

and $M(p)$ as

$$M(p) \triangleq \min_{\mathcal{D}_2 \cap \mathcal{D}'_2} E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p) - E_2^{\delta_d}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p), \quad (5.3.27)$$

where $\mathcal{D}'_2 = \{(\delta_l, \delta_m, \delta_n, \delta_k) \in [0, 1]^4 : \delta_l + \delta_m + \delta_n + \delta_k \leq \delta_{l_1} + \delta_{m_1}\}$. All this said, the following is the main proposition in this chapter.

Proposition 5.4 *For any $0 < C(E_B^{\delta_d}, p) \leq M(p)$, the optimal choice of $E_\eta^{\delta_d}$ is given by*

$$E_\eta^{\delta_d}(\delta_t) = \begin{cases} C(E_B^{\delta_d}, p) & \delta_t > \delta_{l_1} + \delta_{m_1} \\ 0 & \text{else,} \end{cases} \quad (5.3.28)$$

and we have

$$-\frac{1}{N} \log P(\varepsilon) \leq -\delta_d \log \sqrt{4p(1-p)} - E_B^{\delta_d}. \quad (5.3.29)$$

Observe that the requirement $C(E_B^{\delta_d}, p) \leq 0$ is simply the condition on the code (i.e., equation (5.3.10)), with the trivial $E_\eta^{\delta_d}$. Thus, by using a de Caen-based bound, one can only show that the union bound analysis is applicable when $C(E_B^{\delta_d}, p) \leq 0$. However, since it can be easily proved that $M(p) > 0$ for any $0 < p < \frac{1}{2}$, Proposition 5.4 states that by choosing a non trivial $E_\eta^{\delta_d}$, the union bound analysis can be shown to apply in a wider range, $C(E_B^{\delta_d}, p) \leq M(p)$. Furthermore, in Appendix C.4, where Proposition 5.4 is proved, we show that when $0 < C(E_B^{\delta_d}, p) \leq M(p)$, the union bound analysis tightens the bound on the error exponent, with respect to the bound with the trivial $E_\eta^{\delta_d}$, by exactly $C(E_B^{\delta_d}, p)$. When $C(E_B^{\delta_d}, p) > M(p)$, and the new bound does not result in union bound analysis, $E_\eta^{\delta_d}$ as defined in (5.3.28) can still tighten the bound with respect to the trivial $E_\eta^{\delta_d}$, this time by as much as $M(p)$, regardless of $C(E_B^{\delta_d}, p)$.

We give here only an intuitive explanation for Proposition 5.4. The complete proof can be found in Appendix C.4. We wish to prove that the union bound analysis, namely, the bound in (5.3.9) with the trivial $E_\eta^{\delta_d}$, may be applicable even when $C(E_B^{\delta_d}, p) > 0$ (i.e., even when the condition for union bound analysis is not satisfied with the trivial $E_\eta^{\delta_d}$). Observe that the r.h.s. of (5.3.10) is the difference between two minimization problems. Suppose that there exists a function $E_\eta^{\delta_d}$, such that the result of the minimization over \mathcal{D}_2 is increased with respect to the trivial $E_\eta^{\delta_d}$, while the result of the minimization over \mathcal{D}_1 is unchanged. If this is possible, the value of the r.h.s. of (5.3.10) is increased, thus the range in which the union bound analysis apply is widened. The bound in (5.3.9) is the same as it was with the trivial $E_\eta^{\delta_d}$, since the proposed $E_\eta^{\delta_d}$ does not change the result of the minimization over \mathcal{D}_1 . To see that such an $E_\eta^{\delta_d}$ does exist, observe that both $E_2^{\delta_d}$ and $E_1^{\delta_d}$ are convex functions, and their minimization points satisfy $\delta_{l_2} + \delta_{m_2} + \delta_{n_2} + \delta_{k_2} > \delta_{l_1} + \delta_{m_1}$ for every $0 < p < \frac{1}{2}$. Thus, the step function suggested in (5.3.28) can change the result of the minimization over \mathcal{D}_2 without changing the result of the minimization over \mathcal{D}_1 . The threshold value $M(p)$ is due to the fact that the proposed step function cannot unlimitedly increase the result of the minimization over \mathcal{D}_2 . Figure 5.1 includes a basic *one dimensional* example of this concept.

Remark 5.5 Remember that for any received word \mathbf{x} , with $w(\mathbf{x}) = t$, the optimal value of $\eta_i(t)$ is $1/\text{deg}(\mathbf{x})$. Referring to (4.2.27), $\text{deg}(\mathbf{x})$ is a non decreasing function of $w(\mathbf{x})$. Moreover, since the size of any coset is $|\mathcal{C}| = 2^{RN}$, when $R \neq 0$ we expect $\text{deg}(\mathbf{x})$ to grow exponentially with N , at least when $w(\mathbf{x}) = N$ (in this case the exponent is exactly R). Thus, we expect that for a reasonable choice of η_i , there exist $\delta_{t_0} \leq 1$ such that for any $\delta_t \geq \delta_{t_0}$ we have $E_\eta^{\delta_i}(\delta_t) > 0$ and $E_\eta^{\delta_i}(\delta_t) = 0$ otherwise. It is clear that the function $E_\eta^{\delta_d}$ suggested in (5.3.28) answers to this restraint (with $\delta_{t_0} = \delta_{l_1} + \delta_{m_1} < 1$). Note, however, that this choice of $E_\eta^{\delta_d}$ is not necessarily the only optimal choice.

Finally, we return to the general case of the subcode \mathcal{C}_i^* . As explained earlier, the equations required here are cubic, with cumbersome coefficients. Yet, a closed form solution

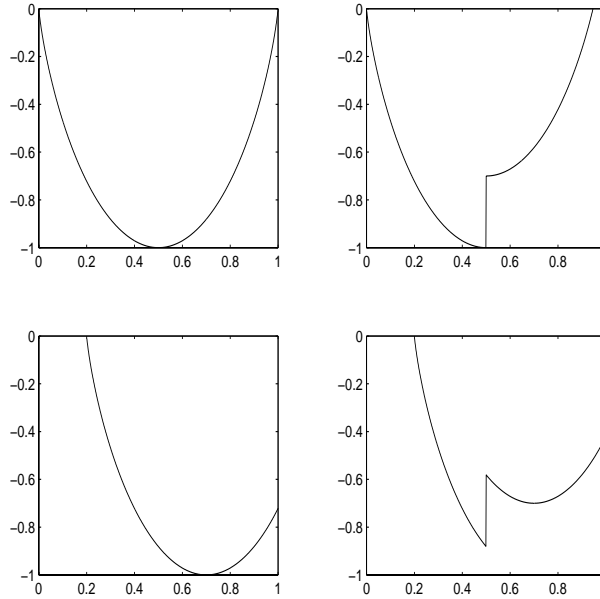


Figure 5.1: A simple one dimensional illustration of the concept behind the proof of Proposition 5.4. The upper left function is $f(x) = -H(x)$, with minimum at $x = 0.5$. The lower left function is $g(x) = -H(x - 0.2)$, with minimum at $x = 0.7$. On the right of each function is the result of adding the step function $s(x)$, $s(x) = 0.3$ for $x > 0.5$ and 0 otherwise. The minimum of $f(x)$ is the same as the minimum of $f(x) + s(x)$. In the lower graphs, however, we have $\min\{g(x) + s(x)\} > \min\{g(x)\}$. Moreover, the minimizing point is changed. In case of a smaller step function, the minimizing point of $g(x) + s(x)$ and $g(x)$ may be the same point, but for a non zero step the value of the minimum will always change.

for these equations exists, and is easily handled using Matlab's symbolic toolbox. We can follow the derivations above (and the proof in Appendix C.4) step by step and find out that the inequality $\delta_{l_2} + \delta_{m_2} + \delta_{n_2} + \delta_{k_2} > \delta_{l_1} + \delta_{m_1}$ is still valid, hence Proposition 5.4 stands solid for any subcode \mathcal{C}_i^* , not necessarily \mathcal{C}_d^* .

5.4 Upper Bounds on the Error Exponent for the Additive White Gaussian Noise Channel

In this section, we calculate the upper bound on the error exponent resulting from the bound in Proposition 3.5. To avoid harassing computations, we restrict ourselves to the *dot product* bound, namely, the bound using the parametric family given in (3.3.9). We follow directly the steps of the previous section, though, it is important to note, we do not seek the optimal optimization function, only compute the bound for the given family of functions and determine the optimal value of the parameter a . The results, in this case, are not as sharp as these of the previous section. Yet, we show that the new bound on the error probability given in Proposition 3.5 results in a tighter bound on the error exponent than the de Caen-based analogue.

Substituting the relevant constants in (3.3.24), the *dot product* bound is given by

$$P(\varepsilon) \geq \frac{\exp\left\{-N \frac{a^2 N_0 E_N}{2}\right\} B_d^N Q^2 \left(\sqrt{\frac{\left(1 + \frac{aN_0}{2}\right)^2 E_N d_N}{N_0/2}} \right)}{Q \left(\sqrt{\frac{(1+aN_0)^2 E_N d_N}{N_0/2}} \right) + (B_d^N - 1) \Psi \left(\frac{1}{2}, \sqrt{\frac{(1+aN_0)^2 E_N d_N}{N_0/2}}, \sqrt{\frac{(1+aN_0)^2 E_N d_N}{N_0/2}} \right)}, \quad (5.4.1)$$

where $a \geq 0$ is a parameter to be optimized. In Appendix C.5, we show that

$$\lim_{x \rightarrow \infty} -\frac{1}{x^2} \ln \Psi \left(\frac{1}{2}, x, x \right) = \frac{2}{3}, \quad (5.4.2)$$

thus, together with (3.3.21), we have

$$P(\varepsilon) \geq \frac{\exp \left\{ -N \left(\frac{a^2 N_0 E_N}{2} - E_B^{\delta_d} \ln 2 + \frac{2 \left(1 + \frac{aN_0}{2}\right)^2 E_N \delta_d}{N_0} \right) \right\}}{\exp \left\{ -N \left(\frac{(1+aN_0)^2 E_N \delta_d}{N_0} \right) \right\} + \exp \left\{ -N \left(\frac{4(1+aN_0)^2 E_N \delta_d}{3N_0} - E_B^{\delta_d} \ln 2 \right) \right\}}. \quad (5.4.3)$$

Analogously to the lower bound for the BSC, the numerator in (5.4.3) is the sum of two exponents. Thus, when computing the bound on the error exponent, two cases should be considered. After some straightforward computations, we have

$$-\frac{1}{N} \ln P(\varepsilon) \leq \frac{E_N \delta_d}{N_0} - E_B^{\delta_d} \ln 2 + \frac{a^2 N_0 E_N (1 - \delta_d)}{2} \quad (5.4.4)$$

if

$$E_B^{\delta_d} \ln 2 \leq \frac{E_N \delta_d}{3N_0} (1 + aN_0)^2 \quad (5.4.5)$$

and

$$-\frac{1}{N} \ln P(\varepsilon) \preceq \frac{a^2 N_0 E_N}{2} + \frac{E_N \delta_d}{N_0} \left(2 \left(1 + \frac{aN_0}{2} \right)^2 - \frac{4}{3} (1 + aN_0)^2 \right) \quad (5.4.6)$$

otherwise. The analysis is similar to the previous section. Condition (5.4.5) can be referred to as a condition on the code. Examine the bound in (5.4.4). $\frac{E_N \delta_d}{N_0}$ is the Bhattacharyya distance for the AWGN channel when the two codewords are at distance $N\delta_d$ apart. Thus, when $a = 0$, the bound in (5.4.4) is simply the union bound analysis discussed earlier, and condition (5.4.5) is the condition for union bound analysis. It is also clear that if this condition is not satisfied with $a = 0$, one can choose $a > 0$ such that the condition is satisfied. However, when $a > 0$, a positive constant is added to the bound in (5.4.4). In other words, when choosing $a > 0$ such that the condition on the code is satisfied, the resulting bound is not as tight as union bound analysis, and the *extension in the range in which the union bound analysis applies*, which was the main result of the previous section, cannot be achieved. Nevertheless, when (5.4.5) is not satisfied with $a = 0$, choosing $a \neq 0$ such that it is satisfied with equality, results in a tighter bound than the bound in (5.4.6) with the trivial $a = 0$. A simple example is given in Section 5.5, Example 5.8.

Finally, we note that the bound in Proposition 3.5, and the resulting bound on the error exponent, can be easily generalized for any subcode \mathcal{C}_i^* , not necessarily \mathcal{C}_d^* . When this is done, analogously to the techniques used for the BSC, one can optimize the bound over all possible subcodes \mathcal{C}_i^* , resulting in a tighter bound. This procedure is mostly useful when there exists a subcode which is not necessarily \mathcal{C}_d^* , yet is large enough and of relatively low weight. In this way one can improve the upper bound on the error exponent.

5.5 Examples

In this section, we give simple numerical examples for the bounds discussed in this chapter. As mentioned earlier, the derivations in this chapter, and thus the results in this section, are valid only for specific codes with known distance distribution. Bounds on the error exponent in general require bounds on the distance distribution of codes.

Example 5.6 (BSC, binomial distance distribution) Consider the binomial distance distribution given by

$$B_i^N = \frac{\binom{N}{i} |\mathcal{C}|}{2^N}. \quad (5.5.1)$$

This distance distribution is the weight enumeration of the average code, and is used as a good approximation for the distance distribution of several linear codes. For example, in [31] Keren and Litsyn derived bounds on the deviation from the binomial distribution for BCH codes. The asymptotic binomial distance distribution is given by

$$E_B^{\delta_i} = H(\delta_i) + R - 1, \quad (5.5.2)$$

where R is the rate of the code. In [28] Barg and Forney showed that the *typical linear code* has a distance distribution which is exactly binomial when $H(\delta_i) + R - 1 \geq 0$ and 0 otherwise. This fact was used to derive the exact error exponent for typical linear codes by the union bound analysis discussed earlier. Barg and Forney used the bound in (5.3.13) and performed an optimization on δ_i over the range $|\delta_i - \frac{1}{2}| \leq \frac{1}{2} - \delta_{GV}(R)$, where the distance distribution is non zero. $\delta_{GV}(R)$ ([27, pp. 557]) is defined by

$$\delta_{GV}(R) \triangleq H^{-1}(1 - R), \quad \delta_{GV}(R) \leq \frac{1}{2}. \quad (5.5.3)$$

We will not perform this minimization here, only plot as an example, where the considered subcode is always $\mathcal{C}_{\frac{N}{2}}^*$, i.e., $\delta_i = \frac{1}{2}$ and $E_B^{\delta_i} = R$. The results for the BSC are in Figure 5.2. The bounds, top to bottom at $R = 0.02$ are: The sphere packing bound (see, for example, [11]). Beneath it is the bound in (5.3.13), when $\delta_i = \delta_d = \delta_{LP}$. δ_{LP} is an upper bound on

the minimum distance of codes derived by McEliece-Rodemich-Rumsey-Welch's ([27, pp. 559]). For low rates ($R \leq R_0 \approx 0.305$) the bound is given by

$$\delta_d(R) \leq \delta_{LP}(R) = \frac{1}{2} \sqrt{H^{-1}(R)(1 - H^{-1}(R))}. \quad (5.5.4)$$

Note that in this case $E_B^{\delta_d} = 0$, i.e., this is exactly the two codewords bound. Below are the new bounds, which diverge at higher rates, followed by the straight line bound ([9]) and a lower bound on the error exponent (as summarized in [11]). It is clear that using a non trivial $E_\eta^{\delta_i}$ tightens the bound. As mentioned earlier, non trivial $E_\eta^{\delta_i}$ allows us to use the union bound analysis in a wider range than this of a de Caen-based bound. This fact is clearly seen in the graphs.

In the following two examples we consider the bounds derived in this chapter when $\delta_i = \delta_d$. As noted earlier, these are not the tightest bounds achievable, since no optimization on δ_i is performed. Still, since in this case it is easy to refer to specific codes, and not just *typical codes*, we use this subcode to show that there exist codes for which the new bounds presented in this work are tighter than the de Caen-based bounds. When $\delta_i = \delta_d$, the code must have an exponential number of minimum distance codewords for the union bound analysis to be effective. To the author's knowledge, the only known sequences of linear codes satisfying this condition were recently discovered by Ashikhmin, Barg and Vlăduț (ABV) in [29]. In these codes, however, the value of $E_B^{\delta_d}$, as well as the rate R , is very small. Thus, for the BSC for example, we give the results for very high values of p . Yet, it is important to note, since the code's rate is very low it is still below the channel capacity for most values of p presented.

Example 5.7 (BSC, ABV codes) We use the bound on the error exponent derived in Section 5.3 with ABV codes. Figure 5.3 includes the results. The two upmost curves are the discussed bound, with trivial $E_\eta^{\delta_d}$ above and non trivial $E_\eta^{\delta_d}$ below. The horizontal line is the value of $E_B^{\delta_d}$. The lowermost curve is the condition on the code for this case. It is clear that for values of p for which the condition is not satisfied, non trivial $E_B^{\delta_d}$ tightens

the bound. Again, it is clear that the improvement is achieved by continuing the usage of union bound analysis, until a certain threshold is exceeded. From this point on, the union bound analysis does not apply, yet the bound with non trivial $E_\eta^{\delta a}$ is still tighter.

Example 5.8 (AWGN, ABV codes) We use the bound on the error exponent derived in Section 5.4, again, with ABV codes. Figure 5.4 includes the results. The two upmost curves are the bound with $a = 0$ above and non trivial a below. The lower curve is the sphere packing bound. It is clear that the bound on the error exponent can be improved with non trivial a , though only in a range of rates where the bound is trivial, i.e., where it is looser than the sphere packing bound. It is also clear that the improvement is where the rate of the code is relatively high, as expected in Section 3.3, where the rational behind the approximation (3.3.9) was discussed.

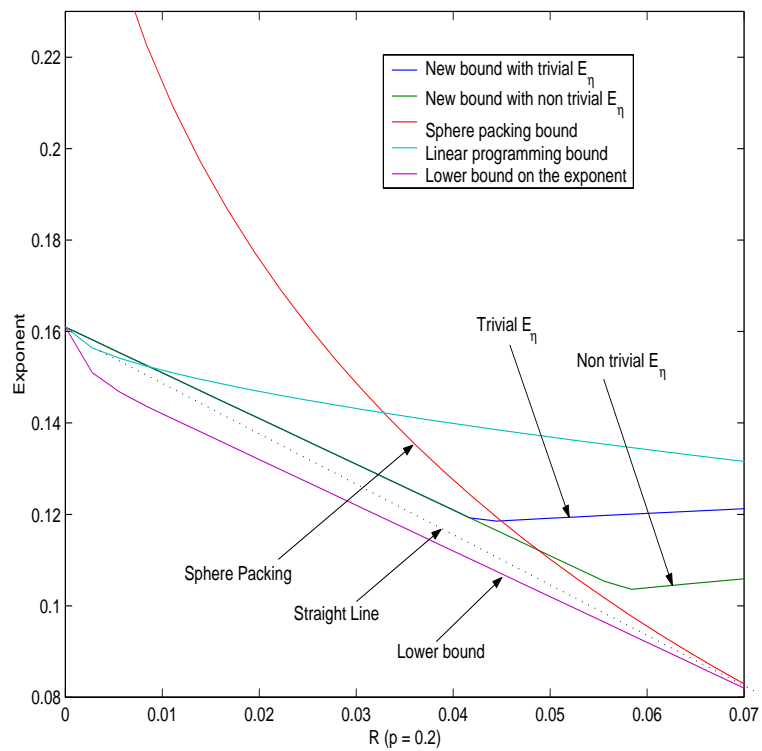


Figure 5.2: Bounds on the error exponent, Example 5.6. Top to bottom, at $R = 0.02$, the bounds are: the sphere packing bound, linear programming bound, the new bounds (which diverge for higher rates), the straight line bound and a lower bound on the error exponent.

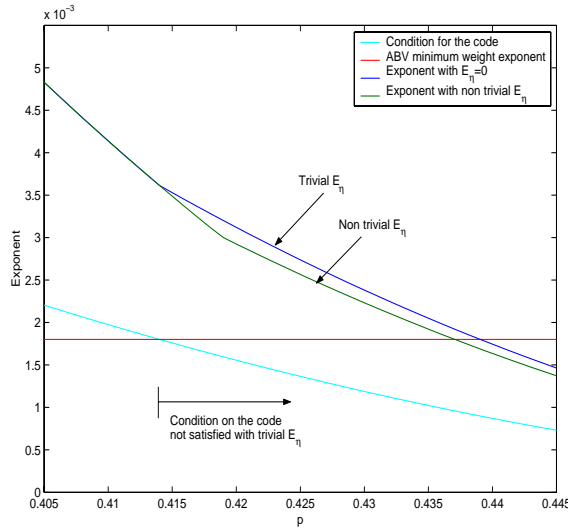


Figure 5.3: Bounds on the error exponent, BSC, Example 5.7 - ABV codes and the subcode \mathcal{C}_d^* . The two upmost curves are the discussed bound, with trivial $E_{\eta}^{\delta_d}$ above and non trivial $E_{\eta}^{\delta_d}$ below. The horizontal line is the value of $E_B^{\delta_d}$. The lowermost curve is the condition on the code. It is clear that for values of p for which the condition is not satisfied with the trivial $E_{\eta}^{\delta_d}$, non trivial $E_B^{\delta_d}$ tightens the bound.

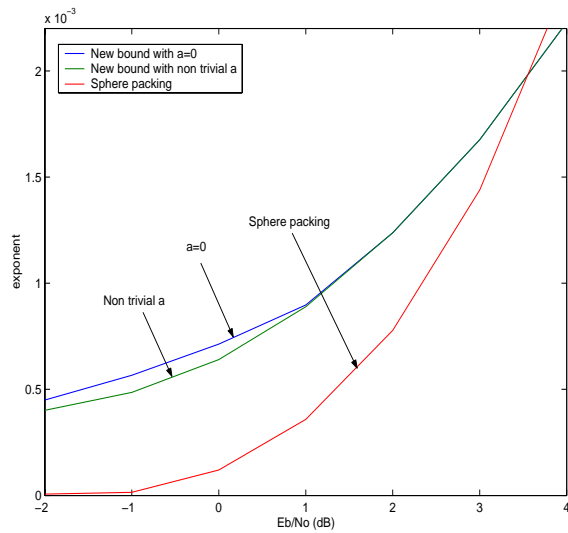


Figure 5.4: Bounds on the error exponent, AWGN, Example 5.8 - ABV codes and the subcode \mathcal{C}_d^* . The two upmost curves are the bound with $a = 0$ above and non trivial a below. The lowermost curve is the sphere packing bound.

Chapter 6

Discussion and Future Work

In this work, lower bounds on the error probability in coded communication were derived. The bounds were derived for both block codes and linear codes over the AWGN channel and the BSC. The first part of the work included a new lower bound on the probability of a union of events. This lower bound was used as a framework for deriving the bounds in the proceeding chapters. The derived framework was easy to use, since in each chapter only channel-specific derivations were required. As explained therein, the core of the bound is the ability optimize the result over a wide family of functions. Moreover, the function which is the optimal choice is known, though not always mathematically endurable, thus may act as a guiding light in the optimization process. Additionally, a tradeoff between computational complexity and performance can be achieved by restricting the set of possible functions for optimization.

In Chapter 3, the bound was used to derive lower bounds on the error probability for the AWGN channel. Several families of functions were suggested for optimization, all including Seguin's bound as private case. The resulting bounds were shown, by means of numerical analysis, to yield the tightest results currently known for a range of E_b/N_0 . In addition, a new bound based on Kounias' lower bound was also derived. This bound was shown to have a very simple form, and, nevertheless, results superior to Seguin's bound for every value of

E_b/N_0 . In Chapter 4, analogous derivations for the BSC were carried out. The resulting bound was shown to perform at least as good as Keren and Litsyn's bound for every value of p , the channel's crossover probability, yet, in this case, the major improvement was achieved in the range where both bounds are inferior to the sphere packing bound.

In Chapter 5, the upper bounds on the error exponent resulting from the bounds in Chapters 3 and 4 were discussed. For the BSC, it was shown that under certain conditions a union bound analysis can be used to achieve an exponentially valid lower bound on the error probability. Moreover, it was shown that using a non trivial choice of the optimization function may weaken these conditions, thus yielding a tighter bound. For the AWGN channel, it was shown that although a non trivial value of the parameter for optimization cannot widen the range of the union bound analysis, it can tighten the bound on the error exponent. It is important to remember, however, that the results for the AWGN channel refer to a bound derived using a specific approximation, and do not claim to propose the best achievable results of the new bounds. The analysis of Chapter 5 also gave a strong clue about how to choose a good family of functions for optimization when long codes are discussed.

To conclude, it was shown that the new bound on the probability of a union gives a powerful framework for deriving lower bounds on the error probability.

As for future work, several suggestions can be examined. The bounds on the error exponent, derived in Chapter 5, are applicable only for specific codes, with known distance distribution. The usage of known and new bounds on the distance distribution of *any* binary (or binary linear) code may be interesting as well. In this case, future work should refer to the bounds and techniques appearing in the works of Litsyn [11] and Burnashev [12].

In both Chapter 3 and Chapter 4, the subcode \mathcal{C}_d^* was used to derive the bounds. However, to specialize the bounds for linear codes and achieve bounds which depend only on the weight enumeration of the code, we assumed that each of the minimum distance codewords are at distance d apart. It is highly reasonable to assume that this is, in general,

not true. Deriving a bound on the number of pairs of codewords at distance d apart in \mathcal{C}_d^* may improve the bound significantly.

Finally, we note that since the new bound on the probability of a union suggests a framework for deriving bounds on the error probability, new bounds can be derived for different channel models. Moreover, the proposed bounds may be improved by seeking new families of functions for optimization.

Appendix A

Computation of the Integrals Required for Proposition 3.2

In this appendix, we compute the generalized pairwise error probability integral (3.3.2) and the generalized triplets error probability integral (3.3.3), required for Proposition 3.2.

A.1 Generalized Pairwise Error Probability Integral

We first compute the integral

$$\int_{\varepsilon_{0i}} p(\mathbf{r}|\mathbf{s}_0)m(\mathbf{r}|\mathbf{s}_0)d\mathbf{r}, \quad (\text{A.1.1})$$

where $\varepsilon_{0i} = \{\mathbf{r} \in \mathbb{R}^K : \|\mathbf{r} - \mathbf{s}_i\| < \|\mathbf{r} - \mathbf{s}_0\|\}$ and $m(\mathbf{r}|\mathbf{s}_0)$ is given by

$$m(\mathbf{r}|\mathbf{s}_0) = \exp \left\{ -(a\|\mathbf{r}\|^2 + b\langle \mathbf{r}, \mathbf{s}_0 \rangle + c\|\mathbf{s}_0\|^2) \right\}. \quad (\text{A.1.2})$$

Remembering that

$$p(\mathbf{r}|\mathbf{s}_0) = (\pi N_0)^{-\frac{K}{2}} \exp \left\{ -\frac{1}{N_0} \|\mathbf{r} - \mathbf{s}_0\|^2 \right\}, \quad (\text{A.1.3})$$

we have

$$\begin{aligned}
& \int_{\varepsilon_{0i}} p(\mathbf{r}|\mathbf{s}_0)m(\mathbf{r}|\mathbf{s}_0)d\mathbf{r} \\
&= \int_{\varepsilon_{0i}} (\pi N_0)^{-\frac{K}{2}} \exp\left\{-\frac{1}{N_0}\|\mathbf{r}-\mathbf{s}_0\|^2\right\} \exp\left\{-(a\|\mathbf{r}\|^2+b\langle\mathbf{r},\mathbf{s}_0\rangle+c\|\mathbf{s}_0\|^2)\right\}d\mathbf{r} \\
&= \int_{\varepsilon_{0i}} (\pi N_0)^{-\frac{K}{2}} \exp\left\{-\left(\left(\frac{1}{N_0}+a\right)\|\mathbf{r}\|^2+\left(b-\frac{2}{N_0}\right)\langle\mathbf{r},\mathbf{s}_0\rangle+\left(\frac{1}{N_0}+c\right)\|\mathbf{s}_0\|^2\right)\right\}d\mathbf{r} \\
&= \int_{\varepsilon_{0i}} (\pi N_0)^{-\frac{K}{2}} \exp\left\{-\frac{1}{N'_0}\|\mathbf{r}-\alpha\mathbf{s}_0\|^2-\beta\|\mathbf{s}_0\|^2\right\}d\mathbf{r},
\end{aligned} \tag{A.1.4}$$

where

$$N'_0 = \frac{N_0}{1+aN_0}, \quad a \neq -\frac{1}{N_0}, \tag{A.1.5}$$

$$\alpha = \frac{\left(\frac{1}{N_0}-\frac{b}{2}\right)}{\left(a+\frac{1}{N_0}\right)}, \tag{A.1.6}$$

$$\beta = \frac{\left(\frac{1}{N_0}+a\right)\left(\frac{1}{N_0}+c\right)-\left(\frac{1}{N_0}-\frac{b}{2}\right)^2}{\frac{1}{N_0}+a}, \tag{A.1.7}$$

and the last equality in (A.1.4) results from a simple completion of squares. Hence,

$$\begin{aligned}
& \int_{\varepsilon_{0i}} p(\mathbf{r}|\mathbf{s}_0)m(\mathbf{r}|\mathbf{s}_0)d\mathbf{r} \\
&= \exp\{-\beta\|\mathbf{s}_0\|^2\} \left(\frac{N'_0}{N_0}\right)^{\frac{K}{2}} \int_{\varepsilon_{0i}} (\pi N'_0)^{-\frac{K}{2}} \exp\left\{-\frac{1}{N'_0}\|\mathbf{r}-\alpha\mathbf{s}_0\|^2\right\}d\mathbf{r}.
\end{aligned} \tag{A.1.8}$$

Assume $N'_0 > 0$, i.e., $a > -\frac{1}{N_0}$. To compute the integral in (A.1.8) note that

$$\begin{aligned}
& \int_{\varepsilon_{0i}} (\pi N'_0)^{-\frac{K}{2}} \exp\left\{-\frac{1}{N'_0}\|\mathbf{r}-\alpha\mathbf{s}_0\|^2\right\}d\mathbf{r} \\
&= P\left(\{\mathbf{r} \in \mathbb{R}^K : \|\mathbf{r}-\mathbf{s}_i\| < \|\mathbf{r}-\mathbf{s}_0\|\} | \mathbf{r} = \tilde{\mathbf{n}} + \alpha\mathbf{s}_0\right) \\
&= P\left(\frac{\langle\tilde{\mathbf{n}},\mathbf{s}_0-\mathbf{s}_i\rangle}{\sqrt{\frac{N'_0}{2}}\|\mathbf{s}_0-\mathbf{s}_i\|} < \frac{(\alpha-1)^2\|\mathbf{s}_0\|^2-\|\alpha\mathbf{s}_0-\mathbf{s}_i\|^2}{\sqrt{2N'_0}\|\mathbf{s}_0-\mathbf{s}_i\|}\right)
\end{aligned} \tag{A.1.9}$$

where $\tilde{\mathbf{n}}$ is a K -dimensional vector of i.i.d. $\mathcal{N}\left(0,\frac{N'_0}{2}\right)$ random variables. Finally, since

$$X'_i \triangleq \frac{\langle\tilde{\mathbf{n}},\mathbf{s}_0-\mathbf{s}_i\rangle}{\sqrt{\frac{N'_0}{2}}\|\mathbf{s}_0-\mathbf{s}_i\|} \tag{A.1.10}$$

is an $\mathcal{N}(0, 1)$ random variable, we have

$$\begin{aligned} & \int_{\varepsilon_{0i}} p(\mathbf{r}|\mathbf{s}_0)m(\mathbf{r}|\mathbf{s}_0)d\mathbf{r} \\ &= \exp\{-\beta\|\mathbf{s}_0\|^2\} \left(\frac{N'_0}{N_0}\right)^{\frac{K}{2}} Q\left(\frac{\|\alpha\mathbf{s}_0 - \mathbf{s}_i\|^2 - (\alpha - 1)^2\|\mathbf{s}_0\|^2}{\sqrt{2N'_0}\|\mathbf{s}_0 - \mathbf{s}_i\|}\right), \end{aligned} \quad (\text{A.1.11})$$

where $Q(\cdot)$ is the error function defined in (3.2.4).

A.2 Generalized Triplets Error Probability Integral

We are interested in computing the integral

$$\int_{\varepsilon_{0i} \cap \varepsilon_{0j}} p(\mathbf{r}|\mathbf{s}_0)m^2(\mathbf{r}|\mathbf{s}_0)d\mathbf{r}, \quad (\text{A.2.1})$$

where $m(\mathbf{r}|\mathbf{s}_0)$ is as given in (A.1.2). Repeating the derivations of Section A.1, we have

$$\begin{aligned} & \int_{\varepsilon_{0i} \cap \varepsilon_{0j}} p(\mathbf{r}|\mathbf{s}_0)m^2(\mathbf{r}|\mathbf{s}_0)d\mathbf{r} \\ &= \exp\{-\beta'\|\mathbf{s}_0\|^2\} \left(\frac{N''_0}{N_0}\right)^{\frac{K}{2}} \int_{\varepsilon_{0i} \cap \varepsilon_{0j}} (\pi N''_0)^{-\frac{K}{2}} \exp\left\{-\frac{1}{N''_0}\|\mathbf{r} - \alpha'\mathbf{s}_0\|^2\right\} d\mathbf{r}, \end{aligned} \quad (\text{A.2.2})$$

where

$$N''_0 = \frac{N_0}{1 + 2aN_0}, \quad a \neq -\frac{1}{2N_0}, \quad (\text{A.2.3})$$

$$\alpha' = \left(\frac{\frac{1}{N_0} - b}{2a + \frac{1}{N_0}}\right), \quad (\text{A.2.4})$$

$$\beta' = \frac{\left(\frac{1}{N_0} + 2a\right)\left(\frac{1}{N_0} + 2c\right) - \left(\frac{1}{N_0} - b\right)^2}{\frac{1}{N_0} + 2a}. \quad (\text{A.2.5})$$

Assuming $N''_0 > 0$, i.e., $a > -\frac{1}{2N_0}$, the integral in (A.2.2) yields

$$\begin{aligned} & \int_{\varepsilon_{0i} \cap \varepsilon_{0j}} (\pi N''_0)^{-\frac{K}{2}} \exp\left\{-\frac{1}{N''_0}\|\mathbf{r} - \alpha'\mathbf{s}_0\|^2\right\} d\mathbf{r} \\ &= P\left(\{\mathbf{r} \in \mathbb{R}^K : \|\mathbf{r} - \mathbf{s}_i\| < \|\mathbf{r} - \mathbf{s}_0\|, \|\mathbf{r} - \mathbf{s}_j\| < \|\mathbf{r} - \mathbf{s}_0\|\} | \mathbf{r} = \hat{\mathbf{n}} + \alpha'\mathbf{s}_0\right) \\ &= P\left(X''_i < \frac{(\alpha' - 1)^2\|\mathbf{s}_0\|^2 - \|\alpha'\mathbf{s}_0 - \mathbf{s}_i\|^2}{\sqrt{2N''_0}\|\mathbf{s}_0 - \mathbf{s}_i\|}, X''_j < \frac{(\alpha' - 1)^2\|\mathbf{s}_0\|^2 - \|\alpha'\mathbf{s}_0 - \mathbf{s}_j\|^2}{\sqrt{2N''_0}\|\mathbf{s}_0 - \mathbf{s}_j\|}\right), \end{aligned} \quad (\text{A.2.6})$$

where $\hat{\mathbf{n}}$ is a K -dimensional vector of i.i.d. $\mathcal{N}\left(0, \frac{N_0''}{2}\right)$ random variables and

$$X_i'' \triangleq \frac{\langle \hat{\mathbf{n}}, \mathbf{s}_0 - \mathbf{s}_i \rangle}{\sqrt{\frac{N_0''}{2}} \|\mathbf{s}_0 - \mathbf{s}_i\|} \quad (\text{A.2.7})$$

is an $\mathcal{N}(0, 1)$ random variable. To conclude, it is easy to verify that

$$\mathbb{E}\{X_i'' X_j''\} = \frac{\langle \mathbf{s}_i - \mathbf{s}_0, \mathbf{s}_j - \mathbf{s}_0 \rangle}{\|\mathbf{s}_i - \mathbf{s}_0\| \|\mathbf{s}_j - \mathbf{s}_0\|} = \rho_{ij}, \quad (\text{A.2.8})$$

where ρ_{ij} was defined in (3.2.6). Consequently,

$$\int_{\varepsilon_{0i} \cap \varepsilon_{0j}} p(\mathbf{r}|\mathbf{s}_0) m^2(\mathbf{r}|\mathbf{s}_0) d\mathbf{r} = \exp\{-\beta' \|\mathbf{s}_0\|^2\} \left(\frac{N_0''}{N_0}\right)^{\frac{K}{2}} \cdot \Psi\left(\rho_{ij}, \frac{\|\alpha' \mathbf{s}_0 - \mathbf{s}_i\|^2 - (\alpha' - 1)^2 \|\mathbf{s}_0\|^2}{\sqrt{2N_0''} \|\mathbf{s}_0 - \mathbf{s}_i\|}, \frac{\|\alpha' \mathbf{s}_0 - \mathbf{s}_j\|^2 - (\alpha' - 1)^2 \|\mathbf{s}_0\|^2}{\sqrt{2N_0''} \|\mathbf{s}_0 - \mathbf{s}_j\|}\right), \quad (\text{A.2.9})$$

where $\Psi(\cdot, \cdot, \cdot)$ is the bivariate normal distribution defined in (3.2.7).

Appendix B

Proofs of Propositions 4.1 and 4.5

B.1 Proof of Proposition 4.1

The proof is as follows. First, we examine a simpler expression than $\tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j)$, in which the sum is only over words $\mathbf{x} \in \varepsilon_{0i} \cap \varepsilon_{0j}$, $i \neq j$ with constant weight $w(\mathbf{x}) = u$, i.e.,

$$\begin{aligned} \tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j, u) \triangleq & \sum_{l=0}^{w(\mathbf{c}_i\mathbf{c}_j)} \sum_{m=\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor + 1 - l}^{w(\mathbf{c}_i) - w(\mathbf{c}_i\mathbf{c}_j)} \sum_{n=\lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor + 1 - l}^{w(\mathbf{c}_j) - w(\mathbf{c}_i\mathbf{c}_j)} \binom{w(\mathbf{c}_i\mathbf{c}_j)}{l} \binom{w(\mathbf{c}_i) - w(\mathbf{c}_i\mathbf{c}_j)}{m} \\ & \cdot \binom{w(\mathbf{c}_j) - w(\mathbf{c}_i\mathbf{c}_j)}{n} \binom{N - w(\mathbf{c}_i) - w(\mathbf{c}_j) + w(\mathbf{c}_i\mathbf{c}_j)}{u - l - m - n} p^u (1-p)^{N-u} \eta_i(u). \end{aligned} \quad (\text{B.1.1})$$

Since $\tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j) = \sum_u \tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j, u)$, if $\tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j, u)$ is monotonically increasing in $w(\mathbf{c}_i\mathbf{c}_j)$ the proposition is proved. Next, observe that $p^u(1-p)^{N-u}\eta_i(u) \geq 0$ does not affect the behavior of $\tilde{P}_{den}(\mathbf{c}_i, \mathbf{c}_j, u)$ as a function of $w(\mathbf{c}_i\mathbf{c}_j)$. Hence, it is enough to prove that the expression

$$\begin{aligned} & \sum_{l=0}^{w(\mathbf{c}_i\mathbf{c}_j)} \sum_{m=\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor + 1 - l}^{w(\mathbf{c}_i) - w(\mathbf{c}_i\mathbf{c}_j)} \sum_{n=\lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor + 1 - l}^{w(\mathbf{c}_j) - w(\mathbf{c}_i\mathbf{c}_j)} \binom{w(\mathbf{c}_i\mathbf{c}_j)}{l} \binom{w(\mathbf{c}_i) - w(\mathbf{c}_i\mathbf{c}_j)}{m} \\ & \cdot \binom{w(\mathbf{c}_j) - w(\mathbf{c}_i\mathbf{c}_j)}{n} \binom{N - w(\mathbf{c}_i) - w(\mathbf{c}_j) + w(\mathbf{c}_i\mathbf{c}_j)}{u - l - m - n} \end{aligned} \quad (\text{B.1.2})$$

is monotonic in $w(\mathbf{c}_i\mathbf{c}_j)$. The expression in (B.1.2) is the number of words in a sub set of $GF(2)^N$, which we denote by $V_{ij}(u, w)$, where $w = w(\mathbf{c}_i\mathbf{c}_j)$. The value of (B.1.2) is

Therefore, $\mathbf{x}_{S_j \setminus S_{j'}} = 0$, $\mathbf{x}_{S_{j'} \setminus S_j} = 1$ and $w(\mathbf{x}_{S_j \cap S_{j'}}) = \lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor$. Accordingly, we have

$$\begin{aligned} & |V_{ij'}(u, w+1) \cap V_{ij}^c(u, w)| \\ &= \sum_{l=0}^w \binom{w}{l} \sum_{m=\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor - l}^{w(\mathbf{c}_i) - w - 1} \binom{w(\mathbf{c}_i) - w - 1}{m} \binom{w(\mathbf{c}_j) - w - 1}{\lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor - l} \binom{N - w(\mathbf{c}_i) - w(\mathbf{c}_j) + w}{u - m - 1 - \lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor}. \end{aligned} \quad (\text{B.1.5})$$

To evaluate $V_{ij}(u, w) \cap V_{ij'}^c(u, w+1)$, note that now $\mathbf{x}_{S_j \setminus S_{j'}} = 1$ and $\mathbf{x}_{S_{j'} \setminus S_j} = 0$. Thus,

$$\begin{aligned} & |V_{ij}(u, w) \cap V_{ij'}^c(u, w+1)| \\ &= \sum_{l=0}^w \binom{w}{l} \sum_{m=\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor - l + 1}^{w(\mathbf{c}_i) - w - 1} \binom{w(\mathbf{c}_i) - w - 1}{m} \binom{w(\mathbf{c}_j) - w - 1}{\lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor - l} \binom{N - w(\mathbf{c}_i) - w(\mathbf{c}_j) + w}{u - m - 1 - \lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor}. \end{aligned} \quad (\text{B.1.6})$$

Hence,

$$\begin{aligned} & |V_{ij'}(u, w+1) \cap V_{ij}^c(u, w)| - |V_{ij}(u, w) \cap V_{ij'}^c(u, w+1)| \\ &= \sum_{l=0}^w \binom{w}{l} \cdot \left[\binom{w(\mathbf{c}_i) - w - 1}{\lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor - l} \binom{w(\mathbf{c}_j) - w - 1}{\lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor - l} \binom{N - w(\mathbf{c}_i) - w(\mathbf{c}_j) + w}{u - \lfloor \frac{w(\mathbf{c}_i)}{2} \rfloor - \lfloor \frac{w(\mathbf{c}_j)}{2} \rfloor + l - 1} \right], \end{aligned} \quad (\text{B.1.7})$$

an expression which is clearly non negative for any $0 \leq w \leq \min \{w(\mathbf{c}_i) - 1, w(\mathbf{c}_j) - 1\}$, $w(\mathbf{c}_i)$, $w(\mathbf{c}_j)$, and u .

B.2 Proof of Proposition 4.5

The proof is as follows. First, taking $M = N$ only weakens the bound since (4.2.32) is a monotonic decreasing function of M . We have

$$P(\varepsilon) \geq LB(p) = \frac{B_d P_{num}^2(d)}{P_{den}(d) + (B_d - 1)P_{den}(d, d)}, \quad (\text{B.2.1})$$

where p is the crossover probability of the channel appearing in the definitions of $P_{num}(d)$, $P_{den}(d)$ and $P_{den}(d, d)$ (equations (4.2.8), (4.2.11) and (4.2.16) respectively). The proof is

analogous to Seguin's proof, as it appears in [4]. As an upper bound we use the union bound

$$P(\varepsilon) \leq UB(p) = \sum_{i=1}^N B_i P(i), \quad (\text{B.2.2})$$

where $P(i)$ is given by

$$P(i) = \sum_{l=\lfloor \frac{i}{2} \rfloor + 1}^i \binom{i}{l} p^l (1-p)^{i-l}. \quad (\text{B.2.3})$$

Thus, we have

$$\begin{aligned} \frac{LB(p)}{UB(p)} &= \frac{B_d P_{num}^2(d)}{\left(\sum_{i=1}^N B_i P(i) \right) (P_{den}(d) + (B_d - 1) P_{den}(d, d))} \\ &= \frac{B_d}{\left(\sum_{i=1}^N B_i \frac{P(i)}{P_{num}(d)} \right) \left(\frac{P_{den}(d)}{P_{num}(d)} + (B_d - 1) \frac{P_{den}(d, d)}{P_{num}(d)} \right)}. \end{aligned} \quad (\text{B.2.4})$$

Observe that both $P(i)$ and $P_{num}(d)$ go to 0 as p goes to 0, hence we may apply l'Hopital's rule until one of them is a non zero constant. Since the expression with the lowest power of p is the first to yield a non zero constant after successive differentiations, we have

$$\lim_{p \rightarrow 0} \frac{P(i)}{P_{num}(d)} = \begin{cases} 0 & i > d \\ \frac{1}{\eta_i(\lfloor \frac{i}{2} \rfloor + 1)} & i = d \\ \infty & i < d, \end{cases} \quad (\text{B.2.5})$$

thus,

$$\lim_{p \rightarrow 0} \sum_{i=1}^N B_i \frac{P(i)}{P_{num}(d)} = \frac{B_d}{\eta_i(\lfloor \frac{d}{2} \rfloor + 1)}. \quad (\text{B.2.6})$$

Using the same method we have

$$\lim_{p \rightarrow 0} \frac{P_{den}(d)}{P_{num}(d)} = \eta_i \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \quad (\text{B.2.7})$$

and

$$\lim_{p \rightarrow 0} \frac{P_{den}(d, d)}{P_{num}(d)} = 0, \quad (\text{B.2.8})$$

therefore,

$$\lim_{p \rightarrow 0} \frac{LB(p)}{UB(p)} = 1. \quad (\text{B.2.9})$$

Appendix C

Proofs and computations for Chapter 5

C.1 Proof of Proposition 5.2

The proof is as follows. We wish to determine the exponential rate (as $N \rightarrow \infty$) of the r.h.s. of (5.3.1). First, consider $P_{num}(i)$. Since the number of summands in (4.2.8) is polynomial in N , the exponential rate is determined by the summand with the maximal exponent. Remembering that

$$\binom{N}{k} \doteq 2^{NH\left(\frac{k}{N}\right)}, \quad (\text{C.1.1})$$

we have

$$\begin{aligned} \lim_{N \rightarrow \infty} -\frac{1}{N} \log \left(\binom{i}{l} \binom{N-i}{m} p^{l+m} (1-p)^{N-l-m} \eta_i(l+m) \right) &= \\ -\delta_i H\left(\frac{\delta_l}{\delta_i}\right) - (1-\delta_i) H\left(\frac{\delta_m}{1-\delta_i}\right) - (\delta_l + \delta_m) \log(p) - (1-\delta_l - \delta_m) \log(1-p) + E_\eta^{\delta_i}(\delta_l + \delta_m) & \\ = E_1^{\delta_i}(\delta_l, \delta_m, p) + E_\eta^{\delta_i}(\delta_l + \delta_m). & \quad (\text{C.1.2}) \end{aligned}$$

Thus,

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log P_{num}(i) = \min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(\delta_l, \delta_m, p) + E_\eta^{\delta_i}(\delta_l + \delta_m) \right\}. \quad (\text{C.1.3})$$

Note that since $E_1^{\delta_i}(\delta_l, \delta_m, p)$ is continuous and $E_\eta^{\delta_i}(\delta_l + \delta_m)$ is piecewise continuous, for large enough N the minimum can be taken over \mathcal{D}_1 , a continuous interval, ignoring the requirements for rational values of δ_l and δ_m . The requirements for integer values in the summation bounds of (4.2.8) were also relaxed for the same reason. The resulting optimization problem is much simpler to solve. The same applies for $P_{den}(i)$ as well, thus, we have

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log P_{den}(i) = \min_{\mathcal{D}_1} \left\{ E_1^{\delta_i}(\delta_l, \delta_m, p) + 2E_\eta^{\delta_i}(\delta_l + \delta_m) \right\}. \quad (\text{C.1.4})$$

As for $P_{den}(i, i)$, except for considering the set of feasible points (for which the preceding argument apply, and we may ignore the requirements for rational values), note that

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \log \binom{i - \lceil \frac{d}{2} \rceil}{l} \binom{\lceil \frac{d}{2} \rceil}{m} \binom{\lceil \frac{d}{2} \rceil}{n} \binom{N - i - \lceil \frac{d}{2} \rceil}{k} \\ = \lim_{N \rightarrow \infty} \frac{1}{N} \log \binom{i - \frac{d}{2}}{l} \binom{\frac{d}{2}}{m} \binom{\frac{d}{2}}{n} \binom{N - i - \frac{d}{2}}{k}, \end{aligned} \quad (\text{C.1.5})$$

thus, we have

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log P_{den}(i, i) = \min_{\mathcal{D}_2} \left\{ E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p) + 2E_\eta^{\delta_i}(\delta_l + \delta_m + \delta_n + \delta_k) \right\}. \quad (\text{C.1.6})$$

To conclude, observe that when considering the exponent of the sum $P_{den}(i) + (B_i - 1)P_{den}(i, i)$, we distinguish between two cases. The first is when

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log P_{den}(i) \leq \lim_{N \rightarrow \infty} -\frac{1}{N} \log ((B_i - 1)P_{den}(i, i)), \quad (\text{C.1.7})$$

namely, when condition (5.3.10) is satisfied, $P_{den}(i)$ dominates $(B_i - 1)P_{den}(i, i)$ and we have (5.3.9). The second is when condition (5.3.10) is not satisfied, $(B_i - 1)P_{den}(i, i)$ dominates $P_{den}(i)$ and we have (5.3.11).

C.2 The minimization of $E_1^{\delta_i}(\delta_l, \delta_m, p)$ over \mathcal{D}_1

To minimize $E_1^{\delta_i}(\delta_l, \delta_m, p)$, as defined in (5.3.7), over \mathcal{D}_1 , as defined in (5.3.4), observe that the minimization may be carried out separately, on δ_l and on δ_m . By the strict convexity of

$-H(\cdot)$, $E_1^{\delta_i}(\delta_l, \delta_m, p)$ is strictly convex with respect to δ_l and δ_m . The boundaries for each variable are intersection of linear inequalities, hence form convex sets. We require

$$\frac{\partial E_1^{\delta_i}}{\partial \delta_l} = -\log\left(\frac{\delta_i}{\delta_l} - 1\right) + \log\left(\frac{1-p}{p}\right) = 0, \quad (\text{C.2.1})$$

yielding $\delta_l = \delta_i p$. However, for $p < \frac{1}{2}$, $\delta_l = \delta_i p < \frac{\delta_i}{2}$, which does not satisfy $\frac{\delta_i}{2} \leq \delta_l \leq \delta_i$, hence the optimal value for δ_l is $\delta_i/2$, which we denote by δ_{l_1} . For δ_m we require

$$\frac{\partial E_1^{\delta_i}}{\partial \delta_m} = -\log\left(\frac{1-\delta_i}{\delta_m} - 1\right) + \log\left(\frac{1-p}{p}\right) = 0, \quad (\text{C.2.2})$$

yielding $\delta_m = p(1-\delta_i)$, which satisfies the condition on δ_m for $0 < p < \frac{1}{2}$, hence is the optimal value for δ_m , denoted by δ_{m_1} . Consequently,

$$\begin{aligned} \min_{\mathcal{D}_1} E_1^{\delta_i}(\delta_l, \delta_m, p) &= E_1^{\delta_i}(\delta_{l_1}, \delta_{m_1}, p) \\ &= -\delta_i H\left(\frac{\delta_i/2}{\delta_i}\right) - (1-\delta_i) H\left(\frac{p(1-\delta_i)}{1-\delta_i}\right) \\ &\quad + (\delta_i/2 + p(1-\delta_i)) \log\left(\frac{1-p}{p}\right) - \log(1-p) \\ &= -\delta_i \log \sqrt{4p(1-p)}. \end{aligned} \quad (\text{C.2.3})$$

C.3 The minimization of $E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p)$ over \mathcal{D}_2

The minimization of $E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p)$, as defined in (5.3.8), over \mathcal{D}_2 , as defined in (5.3.5), is a bit more complex than that of $E_1^{\delta_i}(\delta_l, \delta_m, p)$, since it cannot be performed over each variable separately. Hereafter, we follow the definitions of [32].

Definition C.1 ([32]) *A feasible vector for a given minimization problem is a vector which satisfies all the constraints for that problem.*

For a minimization problem with inequality constraints, we denote the i 'th constraint as $g_i(\mathbf{v}) \leq 0$, e.g., $\frac{1}{2}(\delta_i - \delta_d) \leq \delta_l$ will be denoted as $g_1(\delta_l, \delta_m, \delta_n, \delta_k) = \frac{1}{2}(\delta_i - \delta_d) - \delta_l \leq 0$, and say that the constraint is *active at* \mathbf{v} if it is satisfied with equality for that \mathbf{v} .

Definition C.2 ([32]) A feasible vector \mathbf{v} is said to be regular if the active inequality constraint gradients $\{\nabla g_i(\mathbf{v}) : g_i(\mathbf{v}) = 0\}$ are linearly independent.

When minimizing the function $f(\mathbf{v})$ under the constraints $g_i(\mathbf{v}) \leq 0, i = 1, \dots, r$, the Lagrangian function is defined as

$$L(\mathbf{v}, \mu) = f(\mathbf{v}) + \sum_{j=1}^r \mu_j g_j(\mathbf{v}). \quad (\text{C.3.1})$$

Under the preceding definitions we have the following theorem, which we state in a slightly weaker form, nevertheless sufficient for our purpose.

Theorem C.3 (Karush-Kuhn-Tucker Necessary Conditions, [32]) Let \mathbf{v}^* be a regular local minimum when minimizing $f(\mathbf{v})$ under the constraints $g_i(\mathbf{v}) \leq 0, i = 1, \dots, r$, when f and g_i are continuously differentiable. Then there exists a unique Lagrange Multiplier vector $\mu^* = (\mu_1^*, \dots, \mu_r^*)$ such that $\nabla_{\mathbf{v}} L(\mathbf{v}^*, \mu^*) = 0$. The multipliers μ_j^* satisfy $\mu_j^* \geq 0, j = 1, \dots, r$, and $\mu_j^* = 0$ for each constraint g_j which is inactive at \mathbf{v}^* .

To use Theorem C.3 for our purposes, first observe that the Hessian matrix of $E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p)$ is positive definite (it is a diagonal matrix with positive diagonal elements). Hence $E_2^{\delta_i}(\delta_l, \delta_m, \delta_n, \delta_k, p)$ is strictly convex, and the conditions in Theorem C.3 are necessary and sufficient. Since \mathcal{D}_2 is defined by eight inequality constraints (i.e., $r = 8$), we choose to minimize over a larger set, $\tilde{\mathcal{D}}_2$, defined by the inequalities

$$g_1(\delta_l, \delta_m, \delta_n, \delta_k) = \frac{\delta_i}{2} - \delta_l - \delta_m \leq 0, \quad (\text{C.3.2})$$

$$g_2(\delta_l, \delta_m, \delta_n, \delta_k) = \frac{\delta_i}{2} - \delta_l - \delta_n \leq 0. \quad (\text{C.3.3})$$

Fortunately, we will see that the solution \mathbf{v}^* satisfies $\mathbf{v}^* \in \mathcal{D}_2$. Since our objective function is convex, and the set $\tilde{\mathcal{D}}_2$ is convex, it is clear that this solution is also a global minimum. Moreover, for this set of inequalities we have

$$\begin{aligned} \nabla g_1(\delta_l, \delta_m, \delta_n, \delta_k) &= (-1, -1, 0, 0), \\ \nabla g_2(\delta_l, \delta_m, \delta_n, \delta_k) &= (-1, 0, -1, 0), \end{aligned} \quad (\text{C.3.4})$$

a set of gradients which are clearly linearly independent, hence any feasible point is regular.

When any feasible point is regular all the local minimums can be found using Theorem C.3.

To conclude, the equations $\nabla_{\mathbf{v}}L(\mathbf{v}^*, \mu^*) = 0$ translate to

$$-\log\left(\frac{\delta_i - \delta_d/2}{\delta_l} - 1\right) + \log\left(\frac{1-p}{p}\right) = \mu_1 + \mu_2, \quad (\text{C.3.5})$$

$$-\log\left(\frac{\delta_d/2}{\delta_m} - 1\right) + \log\left(\frac{1-p}{p}\right) = \mu_1, \quad (\text{C.3.6})$$

$$-\log\left(\frac{\delta_d/2}{\delta_n} - 1\right) + \log\left(\frac{1-p}{p}\right) = \mu_2, \quad (\text{C.3.7})$$

$$-\log\left(\frac{1 - \delta_i - \delta_d/2}{\delta_k} - 1\right) + \log\left(\frac{1-p}{p}\right) = 0, \quad (\text{C.3.8})$$

while the inequality constraints together with the constraints on μ_i translate to

$$\mu_1 \left(\frac{\delta_i}{2} - \delta_m - \delta_l\right) = 0, \quad (\text{C.3.9})$$

$$\mu_2 \left(\frac{\delta_i}{2} - \delta_n - \delta_l\right) = 0. \quad (\text{C.3.10})$$

For example, equation (C.3.9) means that either $\mu_1 = 0$ or $g_1(\delta_l, \delta_m, \delta_n, \delta_k)$ is active. The solution of equations (C.3.5) to (C.3.10) is as follows: First, we see that equation (C.3.8) can be solved separately, yielding

$$\delta_k = p \left(1 - \delta_i - \frac{\delta_d}{2}\right). \quad (\text{C.3.11})$$

For $0 < p < \frac{1}{2}$ this value of δ_k satisfies the constraints in the definition of \mathcal{D}_2 , and will be denoted as δ_{k_2} . For δ_l, δ_m and δ_n we should distinguish between four cases:

case 1, $\mu_1 = 0, \mu_2 = 0$: In this case equations (C.3.5) to (C.3.7) can be solved independently for δ_l, δ_m and δ_n , yielding $(\delta_l, \delta_m, \delta_n, \delta_k) = (p(\delta_i - \frac{\delta_d}{2}), \frac{p\delta_d}{2}, \frac{p\delta_d}{2}, \delta_{k_2})$. However, for $p < \frac{1}{2}$ this solution does not satisfy (C.3.2) and (C.3.3) hence is not feasible.

case 2, $\mu_1 \neq 0, \mu_2 = 0$: Here we have $(\delta_l, \delta_m, \delta_n, \delta_k) = (\frac{\delta_i}{2} - \frac{\delta_d}{4}, \frac{\delta_d}{4}, \frac{p\delta_d}{2}, \delta_{k_2})$. For $p < \frac{1}{2}$ this solution does not satisfy (C.3.3).

case 3, $\mu_1 = 0, \mu_2 \neq 0$: Here we have $(\delta_l, \delta_m, \delta_n, \delta_k) = (\frac{\delta_i}{2} - \frac{\delta_d}{4}, \frac{p\delta_d}{2}, \frac{\delta_d}{4}, \delta_{k_2})$. In this case, for $p < \frac{1}{2}$ this solution does not satisfy (C.3.2).

case 4, $\mu_1 \neq 0, \mu_2 \neq 0$: In this case both inequalities are active. We can factor out μ_1 and μ_2 from (C.3.9) and (C.3.10) respectively, and solve the equations for δ_l, δ_m and δ_n . The result is the following cubic equation for δ_l

$$\frac{\delta_l \left(\frac{\delta_d}{2} - \frac{\delta_i}{2} + \delta_l \right)^2}{\left(\frac{\delta_i}{2} - \delta_l \right)^2 \left(\delta_i - \frac{\delta_d}{2} - \delta_l \right)} = \frac{1-p}{p}. \quad (\text{C.3.12})$$

Clearly, the above equation must have one and only one solution, which together with $\delta_{m_2} = \delta_i/2 - \delta_{l_2}, \delta_{n_2} = \delta_i/2 - \delta_{l_2}$ and δ_{k_2} , yields a feasible solution in terms of $\tilde{\mathcal{D}}_2$. Since the solution is rather cumbersome to analyze, we do not include it here. Nevertheless, using Matlab's symbolic toolbox, it is easy to verify that this solution is also feasible in terms of \mathcal{D}_2 , and thus is the required solution for the minimization problem discussed. As noted in Section 5.3, in the special case of $\delta_i = \delta_d$, however, the expressions are simpler and easier to analyze.

C.4 Proof of Proposition 5.4

The proof is as follows. First, we show that indeed $M(p) > 0$ for every $0 < p < \frac{1}{2}$. The proof is based on the fact that for every $0 < p < \frac{1}{2}$ we have

$$\delta_{l_2} + \delta_{m_2} + \delta_{n_2} + \delta_{k_2} > \delta_{l_1} + \delta_{m_1}, \quad (\text{C.4.1})$$

with equality in (C.4.1) only for $p = 0, \frac{1}{2}$. Thus,

$$\min_{\mathcal{D}_2 \cap \mathcal{D}'_2} E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p) > \min_{\mathcal{D}_2} E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p) = E_2^{\delta_d}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p). \quad (\text{C.4.2})$$

To prove the inequality in (C.4.2), assume that there exists $(\delta'_l, \delta'_m, \delta'_n, \delta'_k) \in \mathcal{D}_2 \cap \mathcal{D}'_2$ such that $E_2^{\delta_d}(\delta'_l, \delta'_m, \delta'_n, \delta'_k, p) = E_2^{\delta_d}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p)$. Since $(\delta'_l, \delta'_m, \delta'_n, \delta'_k) \in \mathcal{D}_2$, $(\delta'_l, \delta'_m, \delta'_n, \delta'_k) \neq (\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2})$, by the strict convexity of $E_2^{\delta_d}$ and the convexity of \mathcal{D}_2 there exists $(\delta''_l, \delta''_m, \delta''_n, \delta''_k) \in \mathcal{D}_2$ such that $E_2^{\delta_d}(\delta''_l, \delta''_m, \delta''_n, \delta''_k, p) < E_2^{\delta_d}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p)$, which contradicts the minimality of $E_2^{\delta_d}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p)$. Note that the minimum over $\mathcal{D}_2 \cap \mathcal{D}'_2$

can be calculated¹ using Theorem C.3.

We may now consider the minimization problems in the r.h.s. of (5.3.10), where $\delta_i = \delta_d$, $E_\eta^{\delta_d}$ is as defined in (5.3.28) and $C(E_B^{\delta_d}, p) > 0$. Since $E_\eta^{\delta_d} \geq 0$, and $E_\eta^{\delta_d}(\delta_{l_1} + \delta_{m_1}) = 0$, where $(\delta_{l_1}, \delta_{m_1})$ is the minimizing point of $E_1^{\delta_d}$, it is clear that

$$\min_{\mathcal{D}_1} \left\{ E_1^{\delta_d}(\delta_l, \delta_m, p) + 2E_\eta^{\delta_d}(\delta_l + \delta_m) \right\} = \min_{\mathcal{D}_1} E_1^{\delta_d}(\delta_l, \delta_m, p), \quad (\text{C.4.3})$$

for every $C(E_B^{\delta_d}, p) > 0$. Namely, the result of the minimization over \mathcal{D}_1 is unchanged. However, when considering the minimization of $E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p) + 2E_\eta^{\delta_d}(\delta_l + \delta_m + \delta_n + \delta_k)$ over \mathcal{D}_2 , the value of $C(E_B^{\delta_d}, p)$ is important. Since $\delta_{l_2} + \delta_{m_2} + \delta_{n_2} + \delta_{k_2} > \delta_{l_1} + \delta_{m_1}$, the step function $E_\eta^{\delta_d}$, defined in (5.3.28), “lifts” $E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p)$ in a range which includes its minimizing point. For large enough $C(E_B^{\delta_d}, p)$, i.e., $C(E_B^{\delta_d}, p) > M(p)$, the new minimum must be at a new point, $(\delta'_l, \delta'_m, \delta'_n, \delta'_k) \in \mathcal{D}_2 \cap \mathcal{D}'_2$, yielding

$$\begin{aligned} \min_{\mathcal{D}_2} \left\{ E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p) + 2E_\eta^{\delta_d}(\delta_l + \delta_m + \delta_n + \delta_k) \right\} &= \min_{\mathcal{D}_2 \cap \mathcal{D}'_2} E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p) \\ &= E_2^{\delta_d}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p) + M(p). \end{aligned} \quad (\text{C.4.4})$$

However, for smaller values of $C(E_B^{\delta_d}, p)$, $(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2})$ remains the minimizing point, yielding

$$\min_{\mathcal{D}_2} \left\{ E_2^{\delta_d}(\delta_l, \delta_m, \delta_n, \delta_k, p) + 2E_\eta^{\delta_d}(\delta_l + \delta_m + \delta_n + \delta_k) \right\} = E_2^{\delta_d}(\delta_{l_2}, \delta_{m_2}, \delta_{n_2}, \delta_{k_2}, p) + C(E_B^{\delta_d}, p). \quad (\text{C.4.5})$$

This far, we have proved that when $0 < C(E_B^{\delta_d}, p) \leq M(p)$, $E_\eta^{\delta_d}$ as defined in (5.3.28) alters the condition on the code in such a way that it is not repealed, and the union bound analysis apply. It remains to prove that this is the optimal choice when $0 < C(E_B^{\delta_d}, p) \leq M(p)$, and to quantify the improvement over the bound with the trivial $E_\eta^{\delta_d}$ for every $C(E_B^{\delta_d}, p) > 0$. To see this, Subtract the r.h.s. of (5.3.9) from the r.h.s. of (5.3.11).

¹The solution for this minimization is not required for this proof, only for presenting numeric results. Furthermore, this solution requires solving a cubic equation. Therefore, we solve the cubic equation using Matlab’s symbolic toolbox and present only the relevant results in Section 5.5.

Requiring the result to be negative is no other than the condition on the code (5.3.10). Namely, when the condition on the code is not satisfied, and (5.3.11) is valid, the bound in (5.3.9) is tighter. Thus, in this case, the best choice of $E_\eta^{\delta_d}$ can improve the error exponent by no more than equalizing it to (5.3.9). Since this can be done by the $E_\eta^{\delta_d}$ proposed in (5.3.28), we draw the conclusion that it is the optimal choice. Another, more intuitive, explanation for this result is obtained by noticing that no tighter lower bound on the error probability, calculated on a subcode \mathcal{C}_d^* , can be achieved, than the one which coincides with the union bound. The improvement over the bound with the trivial $E_\eta^{\delta_d}$ is simply the change in the value of the r.h.s. of (5.3.11) caused by our choice of $E_\eta^{\delta_d}$, which is $C(E_B^{\delta_d}, p)$ when $0 < C(E_B^{\delta_d}, p) \leq M(p)$, and $M(p)$ when $C(E_B^{\delta_d}, p) > M(p)$.

C.5 Computation of the exponential rate of $\Psi(\frac{1}{2}, x, x)$

We wish to prove that

$$\lim_{x \rightarrow \infty} -\frac{1}{x^2} \ln \Psi \left(\frac{1}{2}, x, x \right) = \frac{2}{3}, \quad (\text{C.5.1})$$

where Ψ is as defined in (3.2.7). From (3.4.2), we have

$$\Psi \left(\frac{1}{2}, x, x \right) = \frac{\sqrt{3}}{2\pi} \int_0^{\frac{\pi}{4}} \frac{1}{1 - \frac{1}{2} \sin(2\theta)} \exp \left\{ -\frac{2}{3} x^2 \frac{1 - \frac{1}{2} \sin(2\theta)}{\sin^2(\theta)} \right\} d\theta. \quad (\text{C.5.2})$$

For every $x > 1.5$, divide the integral in (C.5.2) to $[x]$ parts, where $[x]$ is the closest integer to x . We have,

$$\begin{aligned} \Psi \left(\frac{1}{2}, x, x \right) &= \frac{\sqrt{3}}{2\pi} \left(\int_0^{\frac{\pi}{4[x]}} \frac{1}{1 - \frac{1}{2} \sin(2\theta)} \exp \left\{ -\frac{2}{3} x^2 \frac{1 - \frac{1}{2} \sin(2\theta)}{\sin^2(\theta)} \right\} d\theta + \right. \\ &\quad \left. \cdots + \int_{\frac{\pi([x]-1)}{4[x]}}^{\frac{\pi}{4}} \frac{1}{1 - \frac{1}{2} \sin(2\theta)} \exp \left\{ -\frac{2}{3} x^2 \frac{1 - \frac{1}{2} \sin(2\theta)}{\sin^2(\theta)} \right\} d\theta \right). \end{aligned} \quad (\text{C.5.3})$$

Observe that, for any $0 \leq \theta \leq \frac{\pi}{4}$, we have

$$1 \leq \frac{1}{1 - \frac{1}{2} \sin(2\theta)} \leq 2 \quad (\text{C.5.4})$$

and

$$1 \leq \frac{1 - \frac{1}{2} \sin(2\theta)}{\sin^2(\theta)} < \infty, \quad (\text{C.5.5})$$

where the expression in (C.5.4) is monotonically increasing in θ , and the expression in (C.5.5) is monotonically decreasing in θ . Thus,

$$\begin{aligned}\Psi\left(\frac{1}{2}, x, x\right) &\leq \frac{\sqrt{3}[x]}{\pi} \int_{\frac{\pi([x]-1)}{4[x]}}^{\frac{\pi}{4}} \exp\left\{-\frac{2}{3}x^2 \frac{1 - \frac{1}{2}\sin(2\theta)}{\sin^2(\theta)}\right\} d\theta \\ &\leq \frac{\sqrt{3}[x]}{\pi} \frac{\pi}{4[x]} \exp\left\{-\frac{2}{3}x^2\right\} \\ &= \frac{\sqrt{3}}{4} \exp\left\{-\frac{2}{3}x^2\right\},\end{aligned}\tag{C.5.6}$$

and

$$\begin{aligned}\Psi\left(\frac{1}{2}, x, x\right) &\geq \frac{\sqrt{3}}{2\pi} \int_{\frac{\pi([x]-1)}{4[x]}}^{\frac{\pi}{4}} \exp\left\{-\frac{2}{3}x^2 \frac{1 - \frac{1}{2}\sin(2\theta)}{\sin^2(\theta)}\right\} d\theta \\ &\geq \frac{\sqrt{3}}{2\pi} \frac{\pi}{4[x]} \exp\left\{-\frac{2}{3}x^2 \left(\frac{1 - \frac{1}{2}\sin\left(2\frac{\pi([x]-1)}{4[x]}\right)}{\sin^2\left(\frac{\pi([x]-1)}{4[x]}\right)}\right)\right\}.\end{aligned}\tag{C.5.7}$$

Taking the natural logarithm of both sides of (C.5.6) and (C.5.7), and dividing by $-x^2$, we have

$$-\frac{1}{x^2} \ln \frac{\sqrt{3}}{4} + \frac{2}{3} \leq -\frac{1}{x^2} \ln \Psi\left(\frac{1}{2}, x, x\right) \leq -\frac{1}{x^2} \ln \frac{\sqrt{3}}{8} - \frac{1}{x^2} \ln \frac{1}{[x]} + \frac{2}{3} \left(\frac{1 - \frac{1}{2}\sin\left(2\frac{\pi([x]-1)}{4[x]}\right)}{\sin^2\left(\frac{\pi([x]-1)}{4[x]}\right)}\right).\tag{C.5.8}$$

For $x \rightarrow \infty$, equation (C.5.1) immediately follows after applying the sandwich rule.

References

- [1] D. de Caen, “A lower bound on the probability of a union,” *Discr. Math.*, vol. 169, pp. 217–220, 1997.
- [2] H. Kuai, F. Alajaji, and G. Takahara, “A lower bound on the probability of a finite union of events,” *Discr. Math.*, vol. 215, pp. 147–158, March 2000.
- [3] E. G. Kounias, “Bounds on the probability of a union, with applications,” *Ann. Math. Statist.*, vol. 39, no. 6, pp. 2154–2158, 1968.
- [4] G. E. Seguin, “A lower bound on the error probability for signals in white Gaussian noise,” *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 3168–3175, November 1998.
- [5] O. Keren and S. Litsyn, “A lower bound on the probability of decoding error over a BSC channel,” *The 21st IEEE Electrical and Electronic Engineers in Israel*, pp. 271–273, 2000.
- [6] G. Poltyrev, “Bounds on decoding error probability of binary linear codes via their spectra,” *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1284–1292, July 1994.
- [7] C. E. Shannon, “Probability of error for optimal codes in a Gaussian channel,” *Bell Syst. Tech. J.*, vol. 38, no. 3, pp. 611–656, May 1959.

- [8] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inform. Theory*, vol. 11, pp. 3–18, January 1965.
- [9] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels," *Inform. Contr.*, vol. 10, pp. 65–103 (Part 1); 522–552 (Part 2), 1967.
- [10] R. J. McEliece and J. K. Omura, "An improved upper bound on the block coding error exponent for binary-input discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. 23, pp. 611–613, September 1977.
- [11] S. Litsyn, "New upper bounds on error exponents," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 385–398, March 1999.
- [12] M. V. Burnashev, "On the relation between the code spectrum and the decoding error probability," *Probl. Inform. Transm.*, vol. 36, no. 4, pp. 285–304, 2000.
- [13] D. A. Dawson and D. Sankoff, "An inequality for probabilities," *Proc. Amer. Math. Soc.*, vol. 18, pp. 504–507, 1976.
- [14] D. Hunter, "An upper bound for the probability of a union," *J. Appl. Probab.*, vol. 13, pp. 597–603, 1976.
- [15] A. Dembo, "Unpublished notes," Communicated by I. Sason, 2000.
- [16] P. Swaszek, "A lower bound on the error probability for signals in white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 837–841, May 1995.
- [17] B. Hughes, "On the error probability of signals in additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 37, no. 1, pp. 151–155, January 1991.
- [18] E. R. Berlekamp, "The technology of error correcting codes," *Proc. IEEE*, vol. 68, no. 8, pp. 564–593, May 1980.

- [19] H. Kuai and G. Takahara F. Alajaji, "Tight error bounds for nonuniform signaling over AWGN channels," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2712–2718, November 2000.
- [20] A.J. Viterbi and J.K. Omura, *Principles of Digital Communication and Coding*, McGraw-Hill, Singapore, 1979.
- [21] J. Galambos and I. Simonelli, *Bonferroni-type Inequalities with Applications*, Springer, 1996.
- [22] M.-S. Alouini and A. J. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1324–1334, September 1999.
- [23] J. W. Craig, "A new, simple and exact result for calculating the probability of error for two dimensional signal constellations," *IEEE MILCOM91 Conf. Rec.*, Boston, MA, pp. 25.5.1–25.5.5, 1991.
- [24] N. C. Beaulieu, "A simple series for personal computer computation of the error function $Q(\cdot)$," *IEEE Trans. Commun.*, vol. 37, no. 9, pp. 989–991, September 1989.
- [25] M. K. Simon and D. Divsalar, "Some new twists to problems involving the Gaussian probability integral," *IEEE Trans. commun.*, vol. 46, no. 2, pp. 200–210, February 1998.
- [26] O. Keren and S. Litsyn, "A simple lower bound on the probability of decoding error over a BSC," Unpublished notes, 2001.
- [27] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland Publishing Company, New York, 1977.

- [28] A. Barg and G. D. Forney, “Random codes: minimum distances and error exponents,” *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2568–2573, September 2002.
- [29] A. Ashikhmin, A. Barg, and S. Vlăduț, “Linear codes with exponentially many light vectors,” *Journal of Combinatorial Theory*, vol. A 96, no. 2, pp. 396–399, November 2001.
- [30] R. G. Gallager, “The random coding bound is tight for the average code,” *IEEE Trans. Inform. Theory*, vol. 19, pp. 244–246, March 1973.
- [31] O. Keren and S. Litsyn, “More on the distance distribution of BCH codes,” *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 251–255, January 1999.
- [32] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, second edition, 1999.