TIGHTENED UPPER BOUNDS ON THE ML DECODING ERROR PROBABILITY OF BINARY LINEAR BLOCK CODES AND APPLICATIONS

MOSHE TWITTO

TIGHTENED UPPER BOUNDS ON THE ML DECODING ERROR PROBABILITY OF BINARY LINEAR BLOCK CODES AND APPLICATIONS

RESEARCH THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS

FOR THE DEGREE OF MASTER OF SCIENCE

IN ELECTRICAL ENGINEERING

MOSHE TWITTO

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

Nisan 5766 HAIFA APRIL 2006

THIS RESEARCH THESIS WAS SUPERVISED BY DR. IGAL SASON UNDER THE AUSPICES OF THE ELECTRICAL ENGINEERING DEPARTMENT

ACKNOWLEDGMENT

I wish to express my sincere and deep gratitude to my supervisor Dr. Igal Sason, for his devoted guidance and invaluable help through all stages of this research.
This whole work was made possible due to the support and encouragement of my dear parents, Rimon and Sara, to whom I dedicate this thesis.

THE GENEROUS FINANCIAL HELP OF TECHNION IS GRATEFULLY ACKNOWLEDGED

To my dear parents, Rimon and Sara.

Contents

List of Figures

Abstract

Since the error performance of coded communication systems rarely admits exact expressions, one resorts to tight analytical upper and lower bounds as useful theoretical and engineering tools for assessing performance and gaining insight into the main system parameters. Since specific good codes are hard to identify, the average performance of ensembles of codes is usually assessed. The reason in this direction was stimulated in the last decade, due to the introduction of capacity-achieving codes, like turbo codes and low-density parity-check codes. Clearly, such bounds should not be subject to the union bounds limitations, as the aforementioned families of codes perform reliably at rates exceeding the cut-off rate of the channel. Furthermore, as the explicit characterization of the Voronoi regions for linear codes is usually unknown, useful bounds should depend solely on the distance spectrum or the input-output weight-enumerating function of the code, which can be found analytically for many codes or ensembles. Although turbo-like codes which closely approach the Shannon capacity limit are decoded using practical and sub-optimal decoding algorithms, the derivation of upper bounds on the maximum-likelihood error probability is of interest. It provides an indication on the ultimate achievable performance of the system under optimal decoding. Tight bounds on the maximum likelihood (ML) decoding error probability also provide an indication on the inherent gap which exists between optimal ML decoding and sub-optimal (e.g., iterative) decoding algorithms.

In addressing some improved versions of the tangential-sphere bound, we focus on the error exponents associated with these bounds. We show that asymptotically, these bounds provide the same error exponent as the tangential-sphere bound of Poltyrev. In the random coding setting, the error exponent of the tangential-sphere bound fails to reproduce the random coding exponent. This motivates us to explore other bounding techniques which may improve the tangential-sphere bound, especially

for high code rates, where the weakness of the tangential-sphere bound is especially pronounced.

In this work, we derive tightened upper bounds on the decoding error probability of binary linear block codes (and ensembles), under maximum-likelihood decoding, where the transmission takes place over an arbitrary binary-input output-symmetric (MBIOS) channel. These bounds are shown to be at least as tight as the Shulman and Feder bound, and are easier to compute than the generalized version of the Duman and Salehi bounds. Hence these bounds reproduce the random coding error exponent and are also suitable for various ensembles of linear codes (e.g., turbo-like codes). For binary linear block codes, we also examine the effect of expurgation of the distance spectrum on the tightness of the new bounds, as well as previously reported bound. The effectiveness of the new bounds is exemplified for various ensembles of turbo-like codes over the AWGN channel; for ensembles of high-rate linear codes, the new bounds appear to be tighter than the tangential-sphere bound.

The good results obtained from the upper bounds which are derived in this thesis, make these bounding techniques applicable to the design and analysis of efficient turbo-like codes. Finally, topics which deserve further research are addressed at the end of this thesis.

List of notation and abbreviations

AHP	:Added-Hyper-Plane
AWGN	:Additive White Gaussian Noise
BER	:Bit error rate
BPSK	:Binary phase shift keying
CPM	:Continuous phase modulation
DS2	:The second version of Duman and Salehi bound
i.i.d.	:independent identically distributed
ITSB	:Improved tangential-sphere bound
IOWEF	:Input-Output Weight Enumeration Function
LDPC	:Low-Density Parity-Check
MBIOS	:Memoryless, Binary-Input and Output-Symmetric
ML	:Maximum-Likelihood
PHN	:Phase noise
RA	:Repeat and Accumulate
SFB	:Shulman and Feder bound
TSB	:Tangential-sphere bound

$\frac{E_\mathrm{b}}{N_0}$	:Energy per bit to spectral noise density
E_s	:Energy per symbol
d_min	:minimum distance of a block code
$h_2(\cdot)$	:The binary entropy function
K	:The dimension of a linear block code
N	:The length of a block code
N_0	:The one-sided spectral power density of the additive white Gaussian noise
$\Pr(A)$	:The probability of event A
$P_\mathrm{b}(\mathcal{C})$	:The bit error probability of the code $\mathcal{C}$
$P_\mathrm{e}(\mathcal{C})$	:The block error probability of the code $\mathcal{C}$
$Q(\cdot)$	:The Q-function
R	:Code rate
R_0	:The cuttoff rate of a channel
$\Gamma(\cdot)$	:The complete Gamma function
$\gamma(\cdot,\cdot)$	:The incomplete Gamma function

Chapter 1

Introduction

Since the advent of information theory, the search for efficient coding systems has motivated the introduction of efficient bounding techniques tailored to specific codes or some carefully chosen ensembles of codes. The incentive for introducing and applying such bounds has strengthened with the introduction of various families of codes defined on graphs which closely approach the channel capacity with feasible complexity (e.g., turbo codes, repeat-accumulate codes [13], and low-density parity-check (LDPC) [19, 27] codes). Moreover, a lot of applications for turbo-like codes were suggested for a variety of digital communication systems, such as Digital Video Broadcasting (DVB-S2), deep space communications and the third generation of CDMA (WCDMA). Their broad applications and the existence of efficient algorithms implemented in custom VLSI circuits (e.g., [26], [52], [53]) enable to apply these iterative decoding schemes to a variety of digital communication systems. Clearly, the desired bounds must not be subject to the union bound limitation, since for long blocks these ensembles of turbo-like codes perform reliably at rates which considerably exceeds the cutoff rate (R_0) of the channel (recalling that for long codes, union bounds are not informative at the portion of the rate region exceeding the cut-off rate of the channel, where the performance of these capacity-approaching codes is most appealing). Although maximum-likelihood (ML) decoding is in general prohibitively complex for long codes, the derivation of bounds on the ML decoding error probability is of interest, providing an ultimate indication of the system performance. Further, the structure of efficient codes is usually not available, necessitating efficient bounds on performance to solely rely on basic features, such as the distance spectrum and

input-output weight enumeration function (IOWEF) of the examined code (for the evaluation of the block and bit error probabilities, respectively, of a specific code or ensemble).

A basic concept which lies in the base of many previously reported upper bounds was introduced by Fano [16] in 1960. It relies on the following inequality:

$$\Pr(\text{word error} \mid \mathbf{c}) \leq \Pr(\text{word error}, \mathbf{y} \in \mathcal{R} \mid \mathbf{c}) + \Pr(\mathbf{y} \notin \mathcal{R} \mid \mathbf{c}) \tag{1.1}$$

where $\mathbf{y}$ denotes the received vector at the output of the channel, $\mathcal{R}$ is an arbitrary geometrical region which can be interpreted as a subset of the observation space, and $\mathbf{c}$ is an arbitrary transmitted codeword. The idea of this bounding technique is to use the union bound only for the joint event where the decoder fails to decode correctly, and in addition, the received signal vector falls inside the region $\mathcal{R}$ (i.e., the union bound is used to upper bound the first term in the RHS of (1.1)). On the other hand, the second term in the RHS of (1.1) represents the probability of the event where the received vector $\mathbf{y}$ falls outside the region $\mathcal{R}$; this event, which is likely to happen in the low SNR regime, is calculated only one time. If we choose the region $\mathcal{R}$ to be the whole observation space, then (1.1) provides the union bound. However, since the upper bound (1.1) is valid for an arbitrary choice of $\mathcal{R}$, many improved upper bounds can be derived by an appropriate selection of this region. Upper bounds from this category differ in the chosen region. For instance, the tangential bound of Berlekamp [6] used the basic inequality in (1.1), where the volume $\mathcal{R}$ is a the n-dimensional Euclidian space separated by a plane. For the derivation of the sphere bound [21], Herzberg and Poltyrev have chosen the region $\mathcal{R}$ to be a sphere around the transmitted signal vector, and then optimized the radius of the sphere in order to get the tightest upper bound within this form. The region $\mathcal{R}$ in Divsalar's bound [11] was chosen to be a hyper-sphere with an additional degree of freedom with respect to the location of its center. It should be noted, however, that the final form of Divsalar's bound was obtained by applying the Chernoff bounds on the encountered probabilities, which results in a simple bound where nor integration or numerical optimizations are needed. Finally, the tangential-sphere bound (TSB) which was proposed for binary linear block codes by Poltyrev, and for M-ary phase-shift keying (PSK) block coded-modulation schemes by Herzberg and Poltyrev [22] incorporate $\mathcal{R}$ as a circular cone of half-angle θ, whose central axis passes through the transmitted

signal vector and the origin. The TSB is one of the tightest upper bounds known to-date for linear block codes whose transmission takes place over the AWGN channel (see [32, 33]). In [45] , Yousefi and Khandani show that the conical region of the TSB is the optimal region among all regions which have azimuthal symmetry w.r.t. the line which passes through the transmitted signal and the origin. This justifies the tightness of the TSB, based on its geometrical interpretation.

All the aforementioned upper bounds are obtained by applying the union bound on the first term in the RHS of (1.1). In [46], Yousefi and Khandani suggest to use the Hunter bound [23] (an upper bound which belongs to the family of second-order Bonferroni-type inequalities [17]) instead of the union bound, in order to get an upper bound on the joint probability in the RHS of (1.1). This modification should result in a tighter upper bound. They refer to the resulting upper bound as the added-hyper-plane (AHP) bound. Yousefi and Mehrabian also apply the Hunter bound, but implement it in a quite different way to obtain an upper bound which is called the improved tangential-sphere bound (ITSB). The tightness of the AHP is demonstrated for some short BCH codes [46] where it is shown to slightly outperform the TSB at the low SNR range. In [47], a comparison between the ITSB and the TSB for few short codes also slightly outperform the TSB, but in parallel, the computational complexity of these bounds as compared to the TSB is significantly larger. An important question which is not addressed analytically in [46, 47] is whether the new upper bounds (namely, the AHP bound and the ITSB) provide an improved lower bound on the error exponent as compared to the error exponent of the TSB. In this thesis, we address this question, and prove that the error exponents of these improved tangential-sphere bounds coincide with the error exponent of the TSB [44]. We note however that the TSB fails to reproduce the random coding error exponent, especially for high-rate linear block codes [21].

Another approach for the derivation of improved upper bounds is based on the Gallager bounding technique which provides a conditional upper bound on the ML decoding error probability given an arbitrary transmitted (length-N) codeword $\mathbf{c}_m$ ($P_{\mathrm{e}|m}$). The conditional decoding error probability is upper bounded by

$$P_{\mathrm{e}|m} \leq \left(\sum_{m' \neq m} \sum_{\mathbf{y}} p_N(\mathbf{y}|\mathbf{c}_m)^{\frac{1}{\rho}} \ \psi_N^m(\mathbf{y})^{1-\frac{1}{\rho}} \ \left(\frac{p_N(\mathbf{y}|\mathbf{c}_{m'})}{p_N(\mathbf{y}|\mathbf{c}_m)} \right)^{\lambda} \right)^{\rho} \tag{1.2}$$

where $0 \leq \rho \leq 1$ and $\lambda \geq 0$ (see [15, 35]). Here, $\psi_N^m(\mathbf{y})$ is an arbitrary probability tilting measure (which may depend on the transmitted codeword $\mathbf{c}_m$), and $p_N(\mathbf{y}|\mathbf{c})$ designates the transition probability measure of the channel. Connections between these two seemingly different bounding techniques in (1.1) and (1.2) were demonstrated in [38], showing that many previously reported bounds (or their Chernoff versions) whose original derivation relies on the concept shown in inequality (1.1) can in fact be re-produced as particular cases of the bounding technique used in (1.2). To this end, one simply needs to choose the suitable probability tilting measure ψ_N^m which serves as the "kernel" for reproducing various previously reported bounds. The observations in [38] relied on some fundamental results which were reported by Divsalar [11].

The tangential-sphere bound (TSB) of Poltyrev often happens to be the tightest upper bound on the ML decoding error probability of block codes whose transmission takes place over a binary-input AWGN channel. However, in the random coding setting, it fails to reproduce the random coding exponent [20] while the second version of the Duman and Salehi (DS2) bound does (see [38]). In fact, also the Shulman-Feder bound (SFB) [37] which is a special case of the latter bound achieves capacity for the ensemble of fully random block codes. Though the SFB is informative for some structured linear block codes with good Hamming properties, it appears to be rather loose when considering sequences of linear block codes whose minimum distance grows sub-linearly with the block length, as is the case with most capacity-approaching ensembles of LDPC and turbo codes. However, the tightness of this bounding technique is significantly improved by combining the SFB with the union bound; this approach was exemplified for some structured ensembles of LDPC codes (see e.g., [28] and the proof of [36, Theorem 2.2]).

In this thesis, we introduce improved upper bounds on the ML decoding error probability of binary linear block codes, which are tightened versions of the SFB. These bounds on the block and bit error probabilities depend on the distance spectrum of the code (or the average distance spectrum of the ensemble) and the input-output weight enumeration function, respectively. The effect of an expurgation of the distance spectrum on the tightness of the resulting bounds is also considered. By applying the new bounds to ensembles of turbo-like codes over the binary-input AWGN channel, we demonstrate the usefulness of these new bounds, especially for

some coding structures of high rates.

The thesis is organized as follows: we present some improved versions of the TSB in Chapter 2, and derive their error exponents. In Chapter 3, we introduce an upper bound on the block error probability which is in general tighter than the SFB, and combine the resulting bound with the union bound. Similarly, appropriate upper bounds on the bit error probability are introduced. Finally, we conclude our work and propose some future research directions in Chapter 4.

For an extensive tutorial paper on performance bounds of linear codes, the reader is referred to [35].

Chapter 2

The Error Exponents of Some Improved Tangential-Sphere Bound

Chapter overview: The performance of maximum-likelihood (ML) decoded binary linear block codes over the AWGN channel is addressed via the tangential-sphere bound (TSB) and some of its improved variations. This chapter is focused on the derivation of the error exponents of these bounds. Although it was previously exemplified that some variations of the TSB suggest an improvement over the TSB for finite-length codes, it is demonstrated in this chapter that all of these bounds possess the same error exponent. Their common value is equal to the error exponent of the TSB, where the latter error exponent was previously derived by Poltyrev and later its expression was simplified by Divsalar.

This chapter is based on the following papers:

- M. Twitto and I. Sason, "On the Error Exponents of Some Improved Tangential-Sphere Bounds," accepted to the *IEEE Trans. on Information Theory*, August 2006.

- M. Twitto and I. Sason, "On the Error Exponents of Some Improved Tangential-Sphere Bounds," submitted to the *24th IEEE Convention of Electrical and Electronics Engineers in Israel*, Eilat, Israel, Nov. 15–17, 2006.

2.1 Introduction

In recent years, much effort has been put into the derivation of tight performance bounds on the error probability of linear block codes under soft-decision maximum-likelihood (ML) decoding. During the last decade, this research was stimulated by the introduction of various codes defined on graphs and iterative decoding algorithms, achieving reliable communication at rates close to capacity with feasible complexity. The remarkable performance of these codes at a portion of the rate region between the channel capacity and cut-off rate, makes the union bound useless for their performance evaluation. Hence, tighter performance bounds are required to gain some insight on the performance of these efficient codes at rates remarkably above the cut-off rate. Duman and Salehi pioneered this research work by adapting the Gallager bounding technique in [19] and making it suitable for the performance analysis of ensembles, based on their average distance spectrum. They have also applied their bound to ensembles of turbo codes and exemplified its superiority over the union bound [14, 15]. Other performance bounds under ML decoding or 'typical pairs decoding' are derived and applied to ensembles of turbo-like codes by Divsalar [11], Divsalar and Biglieri [12], Jin and McEliece [24, 25], Miller and Burshtein [28], Sason and Shamai [32, 33, 35] and Viterbi [49, 50].

The tangential-sphere bound of Poltyrev [31] forms one of the tightest performance bounds for ML decoded linear block codes transmitted over the binary-input additive white Gaussian noise (BIAWGN) channel. The TSB was modified by Sason and Shamai [32] for the analysis of the bit error probability of linear block codes, and was slightly refined by Zangl and Herzog [51]. This bound only depends on the distance spectrum of the code (or the input-output weight enumerating function (IOWEF) of the code for the bit-error analysis [32]), and hence, it can be applied to various codes or ensembles. The TSB falls within the class of upper bounds whose derivation relies on the basic inequality

$$\Pr(\text{word error} \mid \mathbf{c}_0) \leq \Pr(\text{word error},\ \mathbf{y} \in \mathcal{R} \mid \mathbf{c}_0) + \Pr(\mathbf{y} \notin \mathcal{R} \mid \mathbf{c}_0) \tag{2.1}$$

where $\mathbf{c}_0$ is the transmitted codeword, $\mathbf{y}$ denotes the received vector at the output of the channel, and $\mathcal{R}$ designates an arbitrary geometrical region which can be interpreted as a subset of the observation space. The basic idea of this bounding technique is to reduce the number of overlaps between the decision regions associated with the

pairwise error probabilities used for the calculation of union bounds. This is done by separately bounding the error events for which the noise resides in a region $\mathcal{R}$. The TSB, for example, uses a circular hyper-cone as the region $\mathcal{R}$. Other important upper bounds from this family include the simple bound of Divsalar [11], the tangential bound of Berlekamp [6], and the sphere bound of Herzberg and Poltyrev [21]. In [45], Yousefi and Khandani prove that among all the volumes $\mathcal{R}$ which posses some symmetry properties, the circular hyper-cone yields the tightest bound. This finding demonstrates the optimality of the TSB among a family of bounds associated with geometrical regions which possess some symmetry properties, and which are obtained by applying the *union bound* on the first term in the RHS of (2.1). In [46], Yousefi and Khandani suggest to use the Hunter bound [23] (an upper bound which belongs to the family of second-order Bonferroni-type inequalities) instead of the union bound. This modification should result in a tighter upper bound, and they refer to the resulting upper bound as the added hyper plane (AHP) bound. Yousefi and Mehrabian also apply the Hunter bound, but implement it in a quite different way in order to obtain an improved tangential-sphere bound (ITSB) which solely depends on the distance spectrum of the code. The tightness of the ITSB and the AHP bound is exemplified in [46, 47] for some short linear block codes, where these bounds slightly outperform the TSB at the low SNR range.

An issue which is not addressed analytically in [46, 47] is whether the new upper bounds (namely, the AHP bound and the ITSB) provide an improved lower bound on the error exponent as compared to the error exponent of the TSB. In this chapter, we address this question, and prove that the error exponents of these improved tangential-sphere bounds coincide with the error exponent of the TSB. We note however that the TSB fails to reproduce the random coding error exponent, especially for high-rate linear block codes [31].

This chapter is organized as follows: The TSB ([31], [32]), the AHP bound [46] and the ITSB [47] are presented as a preliminary material in Section 2.2. In Section 2.3, we derive the error exponents of the ITSB and the AHP, respectively and state our main result. We conclude our discussion in Section 2.4. An Appendix provides supplementary details related to the proof of our main result.

2.2 Preliminaries

We introduce in this section some preliminary material which serves as a preparatory step towards the presentation of the material in the following section. We also present notation from [11] which is useful for our analysis. The reader is referred to [35, 48] which introduce material covered in this section. However, in the following presentation, we consider boundary effects which were not taken into account in the original derivation of the two improved versions of the TSB in [46]–[48]). Though these boundary effects do not have any implication in the asymptotic case where we let the block length tend to infinity, they are addressed in this section for finite block lengths.

2.2.1 Assumption

Throughout this chapter, we assume a binary-input additive white Gaussian noise (AWGN) channel with double-sided spectral power density of $\frac{N_0}{2}$. The modulation of the transmitted signals is antipodal, and the modulated signals are coherently detected and ML decoded (with soft decision).

2.2.2 Tangential-Sphere Bound (TSB)

The TSB forms an upper bound on the decoding error probability of ML decoding of linear block code whose transmission takes place over a binary-input AWGN channel [31, 32]. Consider an (N, K) linear block code $\mathcal{C}$ of rate $R \triangleq \frac{K}{N}$ bits per channel use. Let us designate the codewords of $\mathcal{C}$ by $\{\mathbf{c}_i\}$, where $i = 0, 1, \ldots, 2^K - 1$. Assume a BPSK modulation and let $\mathbf{s}_i \in \{+\sqrt{E_\mathrm{s}}, -\sqrt{E_\mathrm{s}}\}^N$ designate the corresponding equi-energy, modulated vectors, where E_s designates the transmitted symbol energy. The transmitted vectors $\{\mathbf{s}_i\}$ are obtained from the codewords $\{\mathbf{c}_i\}$ by applying the mapping $\mathbf{s}_i = (2\mathbf{c}_i - \mathbf{1})\sqrt{E_\mathrm{s}}$, so their energy is NE_s. Since the channel is memoryless, the received vector $\mathbf{y} = (y_1, y_2, \ldots, y_N)$, given that $\mathbf{s}_i$ is transmitted, can be expressed as

$$y_j = s_{i,j} + z_j \ , \quad j = 1, 2, \ldots, N \tag{2.2}$$

where $s_{i,j}$ is the j^{th} component of the transmitted vector $\mathbf{s}_i$, and $\mathbf{z} = (z_1, z_2, \ldots, z_N)$ designates an N-dimensional Gaussian noise vector which corresponds to N orthogonal projections of the AWGN. Since $\mathbf{z}$ is a Gaussian vector and all its components are un-correlated, then the N components of $\mathbf{z}$ are i.i.d., and each component has a zero mean and variance $\sigma^2 = \frac{N_0}{2}$.

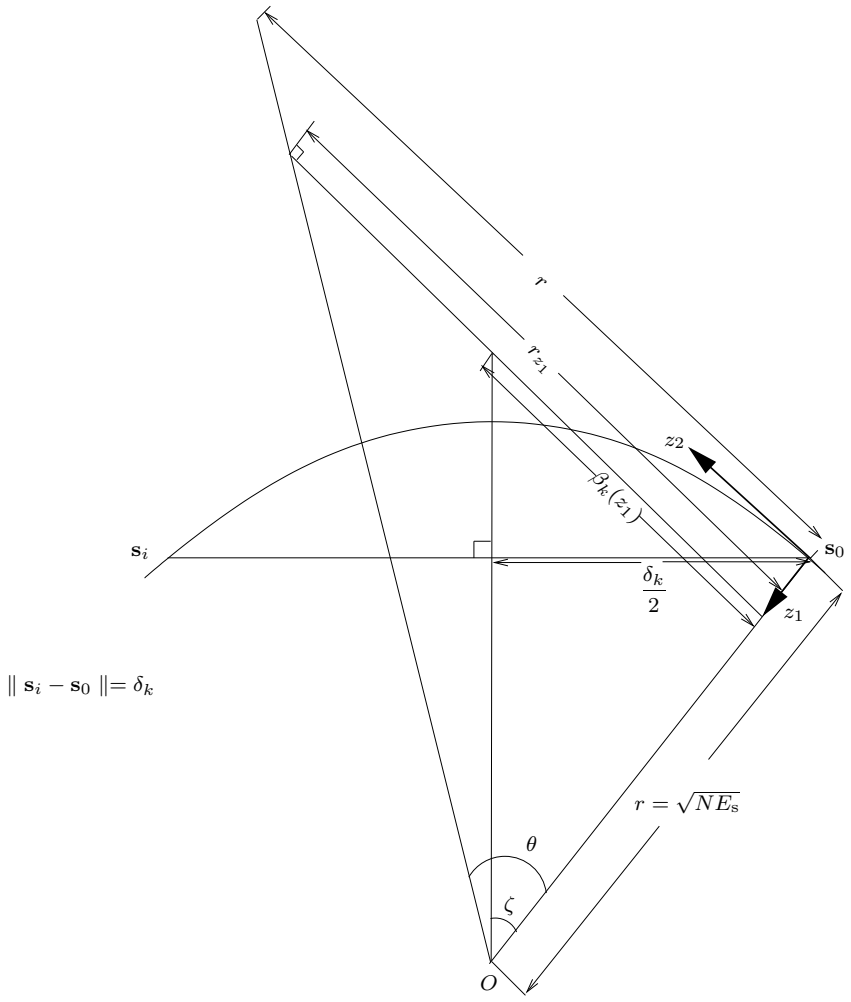


Figure 2.1: The geometric interpretation of the TSB.

Let E be the event of deciding erroneously (under ML decoding) on a codeword

other than the transmitted codeword. The TSB is based on the central inequality

$$\Pr(E|\mathbf{c}_0) \leq \Pr(E, \mathbf{y} \in \mathcal{R}|\mathbf{c}_0) + \Pr(\mathbf{y} \notin \mathcal{R}|\mathbf{c}_0) \tag{2.3}$$

where $\mathcal{R}$ is an N-dimensional circular cone with a half angle θ and a radius r, whose vertex is located at the origin and whose main axis passes through the origin and the point corresponding to the transmitted vector (see Fig. 2.1). The optimization is carried over r (r and θ are related as shown in Fig. 2.1). Let us designate this circular cone by $C_N(\theta)$. Since we deal with linear codes, the conditional error probability under ML decoding does not depend on the transmitted codeword of the code $\mathcal{C}$, so without any loss of generality, one can assume that the all-zero codeword, $\mathbf{s}_0$, is transmitted. Let z_1 be the radial component of the noise vector $\mathbf{z}$ (see Fig. 2.1) so the other $N-1$ components of $\mathbf{z}$ are orthogonal to the radial component z_1. From Fig. 2.1, we obtain that

$$\begin{aligned} r &= \sqrt{NE_{\mathrm{s}}} \tan\theta \\ r_{z_1} &= \left(\sqrt{NE_{\mathrm{s}}} - z_1\right) \tan\theta \\ \beta_k(z_1) &= \left(\sqrt{NE_{\mathrm{s}}} - z_1\right) \tan\zeta = \frac{\sqrt{NE_{\mathrm{s}}} - z_1}{\sqrt{NE_{\mathrm{s}} - \frac{\delta_k^2}{4}}} \frac{\delta_k}{2} \end{aligned} \tag{2.4}$$

The random variable $Y \triangleq \sum_{i=2}^{N} z_i^2$ is χ^2 distributed with $N-1$ degrees of freedom, so its *pdf* is given by

$$f_Y(y) = \frac{y^{\frac{N-3}{2}} e^{-\frac{y}{2\sigma^2}} U(y)}{2^{\frac{N-1}{2}} \sigma^{N-1} \Gamma\left(\frac{N-1}{2}\right)}, \quad y \geq 0 \tag{2.5}$$

where U designates the unit step function, and the function Γ is the complete Gamma function

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad \mathrm{Real}(x) > 0. \tag{2.6}$$

Conditioned on the value of the radial component of the noise, z_1, let $E(z_1)$ designate the decoding error event. The conditional error probability satisfies the inequality

$$\Pr(E(z_1) \mid z_1) \leq \Pr\left(E(z_1), \mathbf{y} \in C_N(\theta) \mid z_1\right) + \Pr\left(\mathbf{y} \notin C_N(\theta) \mid z_1\right) \tag{2.7}$$

The conditional error event $E(z_1)$ can be expressed as a union of pairwise error events, so

$$\Pr(E(z_1), \mathbf{y} \in C_N(\theta) \mid z_1) = \Pr\left(\bigcup_{i=1}^{M-1} E_{0\to i}(z_1), \mathbf{y} \in C_N(\theta) \mid z_1\right), \qquad M \triangleq 2^K \quad (2.8)$$

where $E_{0\to i}(z_1)$ designates the event of error had the only codewords been $\mathbf{c}_0$ and $\mathbf{c}_i$, given the value z_1 of the radial component noise in Fig. 2.1, and $M \triangleq 2^K$ denotes the number of codewords of the code $\mathcal{C}$. We note that for BPSK modulation, the Euclidean distance between the two signals $\mathbf{s}_i$ and $\mathbf{s}_0$ is directly linked to the Hamming weight of the codeword $\mathbf{c}_i$. Let the Hamming weight of $\mathbf{c}_i$ be h, then the Euclidean distance between $\mathbf{s}_0$ and $\mathbf{s}_i$ is equal to $\delta_h = 2\sqrt{hE_s}$. Let $\{A_h\}$ be the distance spectrum of the linear code $\mathcal{C}$, and let $E_h(z_1)$ be the event of deciding under ML decoding in favor of other codeword $\mathbf{c}_i$ whose Hamming weight is h, given the value of z_1. By applying the union bound on the RHS of (2.8), we get

$$\Pr(E(z_1), \mathbf{y} \in C_N(\theta) \mid z_1) \le \sum_{h=1}^{N} A_h \Pr(E_h(z_1), \mathbf{y} \in C_N(\theta) \mid z_1). \quad (2.9)$$

Combining (2.7) and (2.9) gives

$$\Pr\left(E(z_1) \mid z_1\right) \le \sum_{h} \left\{ A_h \Pr\left(E_h(z_1), \mathbf{y} \in C_N(\theta) \mid z_1\right) \right\} + \Pr\left(\mathbf{y} \notin C_N(\theta) \mid z_1\right). \quad (2.10)$$

The second term in the RHS of (2.10) is evaluated from (2.5)

$$\begin{aligned} \Pr(\mathbf{y} \notin C_N(\theta) \mid z_1) &= \Pr\left(Y \ge r_{z_1}^2 \mid z_1\right) \\ &= \int_{r_{z_1}^2}^{\infty} f_Y(y) dy \\ &= \int_{r_{z_1}^2}^{\infty} \frac{y^{\frac{N-2}{2}} e^{-\frac{y}{2\sigma^2}} U(y)}{2^{\frac{N-1}{2}} \sigma^{N-1} \Gamma\left(\frac{N-1}{2}\right)} dy. \end{aligned} \quad (2.11)$$

This integral can be expressed in terms of the incomplete Gamma function

$$\gamma(a, x) \triangleq \frac{1}{\Gamma(a)} \int_0^x t^{x-1} e^{-t} dt, \quad a > 0,\ x \ge 0 \quad (2.12)$$

and it is transformed to

$$\Pr(\mathbf{y} \notin C_N(\theta) \mid z_1) = 1 - \gamma\left(\frac{N-1}{2}, \frac{r_{z_1}^2}{2\sigma^2}\right). \quad (2.13)$$

Let z_2 designate the tangential component of the noise vector $\mathbf{z}$, which is on the plane that contains the signals $\mathbf{s}_0$, $\mathbf{s}_i$ and the origin of the space, and orthogonal to z_1 (see Fig. 2.1). Referring to the first term in the RHS of (2.10), it follows from the geometry in Fig. 2.1 that if $z_1 \leq \sqrt{NE_{\rm s}}$ then

$$\begin{aligned}\Pr(E_h(z_1), \mathbf{y} \in C_N(\theta) \mid z_1) &= \Pr(E_h(z_1), Y \leq r_{z_1}^2 \mid z_1) \\ &= \Pr\left(\beta_h(z_1) \leq z_2 \leq r_{z_1}, Y \leq r_{z_1}^2 \mid z_1\right). \end{aligned} \tag{2.14}$$

Let $V \triangleq \sum_{i=3}^{N} z_i^2$, then $V = Y - z_2^2$. If $z_1 \leq \sqrt{NE_{\rm s}}$, then we obtain the equality

$$\Pr\left(E_h(z_1), \mathbf{y} \in C_N(\theta) \mid z_1\right) = \Pr\left(\beta_h(z_1) \leq z_2 \leq r_{z_1}, V \leq r_{z_1}^2 - z_2^2 \mid z_1\right). \tag{2.15}$$

The random variable V is χ^2 distributed with $N-2$ degrees of freedom, so its *pdf* is

$$f_V(v) = \frac{y^{\frac{N-4}{2}} e^{-\frac{y}{2\sigma^2}} U(y)}{2^{\frac{N-2}{2}} \sigma^{N-2} \Gamma\left(\frac{N-2}{2}\right)}, \quad v \geq 0 \tag{2.16}$$

and since the random variables V and Z_2 are statistically independent, then if $z_1 \leq \sqrt{NE_{\rm s}}$

$$\Pr\left(E_h(z_1), \mathbf{y} \in C_N(\theta) \mid z_1\right) = \int_{\beta_h(z_1)}^{r_{z_1}} \frac{e^{-\frac{z_2^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \int_0^{r_{z_1}^2 - z_2^2} f_V(v) dv \, dz_2. \tag{2.17}$$

In order to obtain an upper bound on the decoding error probability, $\Pr(E)$, one should apply the statistical expectation operator on the RHS of (2.10) w.r.t. the radial noise component z_1. Referring to the upper half azimuthal cone depicted in Fig. 2.1 which corresponds to the case where the radial noise component satisfies the condition $z_1 \leq \sqrt{NE_{\rm s}}$, the inequality $\beta_h(z_1) < r_{z_1}$ holds for the values of h for which $\frac{\delta_h}{2} < \alpha_h$ where

$$\alpha_h \triangleq r\sqrt{1 - \frac{\delta_h^2}{4NE_{\rm s}}}. \tag{2.18}$$

On the other hand, if $z_1 > \sqrt{NE_{\rm s}}$, the range of integration for the component noise z_2 is $\beta_h(z_1) \leq z_2 \leq -r_{z_1}$ which is satisfied for all values of h (since for $z_1 > \sqrt{NE_{\rm s}}$, we get from (2.4) that $r_{z_1} < 0$ and $\beta_h(z_1) < 0$, so the inequality $\beta_h(z_1) \leq -r_{z_1}$ holds in this case for all values of h). Since $Z_1 \sim N(0, \sigma^2)$ where $\sigma^2 = \frac{N_0}{2}$, then the probability that the Gaussian random variable Z_1 exceeds $\sqrt{NE_{\rm s}}$ is equal to

$$Q\left(\frac{\sqrt{NE_{\rm s}}}{\sigma}\right) = Q\left(\sqrt{\frac{2NRE_{\rm b}}{N_0}}\right).$$

This results in the following upper bound on the decoding error probability under ML decoding

$$\Pr(E) \leq \int_{-\infty}^{+\sqrt{NE_{\mathrm{s}}}} \frac{e^{-\frac{z_2^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \Bigg\{ \sum_{h:\frac{\delta_h}{2}<\alpha_h} \quad \Bigg\{ A_h \int_{\beta_h(z_1)}^{r_{z_1}} \frac{e^{-\frac{z_2^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \int_0^{r_{z_1}^2 - z_2^2} f_V(v) dv \; dz_2 \Bigg\} \\ +1 - \gamma\left(\tfrac{N-1}{2}, \tfrac{r_{z_1}^2}{2\sigma^2}\right) \Bigg\} dz_1 + Q\left(\sqrt{\tfrac{2NRE_{\mathrm{b}}}{N_0}}\right) \quad (2.19)$$

The upper bound (2.19) is valid for all positive values of r. Hence, in order to achieve the tightest upper bound of the form (2.19) one should set to zero the partial derivative of the RHS of (2.19) w.r.t. r_{z_1}. After straightforward algebra the following optimization equation for the optimal value of r is obtained [31]:

$$\begin{cases} \displaystyle\sum_{h:\frac{\delta_h}{2}<\alpha_h} A_h \int_0^{\theta_h} \sin^{N-3}\phi \; d\phi = \frac{\sqrt{\pi}\,\Gamma(\frac{N-2}{2})}{\Gamma(\frac{N-1}{2})} \\ \theta_h = \cos^{-1}\left(\frac{\delta_h}{2\alpha_h}\right) \end{cases} \quad (2.20)$$

where α_h is given in (2.18). A proof for the existence and uniqueness of a solution r to the optimization equation (2.20) was provided in [33, Appendix B], together with an efficient algorithm to solve this equation numerically. In order to derive an upper bound on the bit error probability, let $A_{w,h}$ designate the corresponding coefficient in the IOWEF which is the number of codewords which are encoded by information bits whose number of ones is equal to w (where $0 \leq w \leq nR$) and whose Hamming weights (after encoding) are equal to h, and define

$$A'_h \triangleq \sum_{w=1}^{NR} \left(\frac{w}{NR}\right) A_{w,h}, \quad h = 0, \ldots, N. \quad (2.21)$$

In [33, Appendix C], Sason and Shamai derive an upper bound on the bit error probability by replacing the distance spectrum $\{A_h\}$ in (2.19) and (2.20) with the sequence $\{A'_h\}$, and show some properties on the resulting bound on the bit error probability.

2.2.3 Improved Tangential-Sphere Bound (ITSB)

In [47], Yousefi and Mehrabian derive a new upper bound on the block error probability of binary linear block codes whose transmission takes place over a binary-input

AWGN channel, and which are coherently detected and ML decoded. This upper bound, which is called improved tangential-sphere bound (ITSB) is based on inequality (2.3), where the region $\mathcal{R}$ is the same as of the TSB (i.e., an N-dimensional circular cone). To this end, the ITSB is obtained by applying a Bonferroni-type inequality of the second order [17, 23] (instead of the union bound) to get an upper bound on the joint probability of decoding error and the event that the received vector falls within the corresponding conical region around the transmitted signal vector.

The basic idea in [47] relies on Hunter's bound which states that if $\{E_i\}_{i=1}^{M}$ designates a set of M events, and E_i^{c} designates the complementary event of E_i, then

$$\begin{aligned}\Pr\left(\bigcup_{i=1}^{M} E_i\right) &= \Pr(E_1) + \Pr(E_2 \cap E_1^{c}) + \ldots + \Pr(E_M \cap E_{M-1}^{\mathrm{c}} \ldots \cap E_1^{\mathrm{c}}) \\ &\leq \Pr(E_1) + \sum_{i=2}^{M} \Pr(E_i \cap E_{\hat{i}}^{\mathrm{c}}). \end{aligned} \tag{2.22}$$

where the indices $\hat{i} \in \{1, 2, \ldots i-1\}$ are chosen arbitrarily for $i \in \{2, \ldots, M\}$. Clearly, the upper bound (2.22) is tighter than the union bound. The LHS of (2.22) is invariant to the ordering of the events (since it only depends on the union of these events), while the RHS of (2.22) depends on this ordering. Hence, the tightest bound of the form (2.22) is obtained by choosing the optimal indices ordering $i \in \{1, 2, \ldots, M\}$ and $\hat{i} \in \{1, 2, \ldots, i-1\}$. Let us designate by $\Pi(1, 2, \ldots, M) = \{\pi_1, \pi_2, \ldots, \pi_M\}$ an arbitrary permutation among the $M!$ possible permutations of the set $\{1, 2, \ldots, M\}$ (i.e., a permutation of the indices of the events E_1 to E_M), and let $\Lambda = (\lambda_2, \lambda_3, \ldots \lambda_M)$ designate an arbitrary sequence of integers where $\lambda_i \in \{\pi_1, \pi_2, \ldots \pi_{i-1}\}$. Then, the tightest form of of the bound in (2.22) is given by

$$\Pr\left(\bigcup_{i=1}^{M} E_i\right) \leq \min_{\Pi,\Lambda} \left\{ \Pr(E_{\pi_1}) + \sum_{i=2}^{M} \Pr(E_{\pi_i} \cap E_{\lambda_i}^{c}) \right\}. \tag{2.23}$$

Similar to the TSB, the derivation of the ITSB originates from the upper bound (2.7) on the conditional decoding error probability, given the radial component (z_1) of the noise vector (see Fig. 2.1). In [47], it is proposed to apply the upper bound (2.23) on the RHS of (2.8) which for an arbitrary permutation $\{\pi_1, \pi_2, \ldots, \pi_M\}$ and

a corresponding sequence of integers $(\lambda_2, \lambda_3, \ldots \lambda_{M-1})$ as above, gives

$$\Pr\left(\bigcup_{i=1}^{M-1} E_{0\to i}, \mathbf{y} \in C_N(\theta) \mid z_1\right) \leq \min_{\Pi,\Lambda}\Bigg\{\Pr(E_{0\to\pi_1}, \mathbf{y} \in C_N(\theta) \mid z_1) + \sum_{i=2}^{M-1} \Pr(E_{0\to\pi_i}, E^{\mathrm{c}}_{0\to\lambda_i}, \mathbf{y} \in C_N(\theta) \mid z_1)\Bigg\} \tag{2.24}$$

where $E_{0\to j}$ designates the pairwise error event where the decoder decides on codeword $\mathbf{c}_j$ rather than the transmitted codeword $\mathbf{c}_0$. As indicated in [45, 47], the optimization problem of (2.24) is prohibitively complex. In order to simplify it, Yousefi and Mehrabian suggest to choose $\pi_1 = \lambda_i = i_{\min}$ for all $i = 2, \ldots, M-1$, where $i_{\min}$ designates the index of a codeword which is closest (in terms of Euclidian distance) to the transmitted signal vector $\mathbf{s}_0$. Since the code is linear and the channel is memoryless and symmetric, one can assume without any loss of generality that the all-zero codeword is transmitted. Moreover, since we deal with antipodal modulation, then $w_{\mathrm{H}}(\mathbf{c}_{i_{\min}}) = d_{\min}$ where $d_{\min}$ is the minimum distance of the code. Hence, by this specific choice of π_1 and Λ (which in general loosen the tightness of the bound in (2.24)), the ordering of the indices $\{\pi_2, \ldots, \pi_M\}$ is irrelevant, and one can omit the optimization over Π and Λ. The above simplification results in the following inequality:

$$\Pr(E|z_1) \leq \Pr\left(E_{0\to i_{\min}}, \mathbf{y} \in C_N(\theta) \mid z_1\right) + \sum_{i=2}^{M-1} \Pr(E_{0\to i}, E^{\mathrm{c}}_{0\to i_{\min}}, \mathbf{y} \in C_N(\theta) \mid z_1) + \Pr\left(\mathbf{y} \notin C_N(\theta) \mid z_1\right). \tag{2.25}$$

Based on Fig. 2.1, the first and the third terms in the RHS of (2.25) can be evaluated in similarity with the TSB, and we get

$$\Pr\left(E_{0\to i_{\min}}, \mathbf{y} \in C_N(\theta) \mid z_1\right) = \Pr(\beta_{\min}(z_1) \leq z_2 \leq r_{z_1},\ V < r_{z_1}^2 - z_2^2 \mid z_1) \tag{2.26}$$

$$\Pr(\mathbf{y} \notin C_N(\theta) \mid z_1) = 1 - \gamma\left(\frac{N-1}{2}, \frac{r_{z_1}^2}{2\sigma^2}\right) \tag{2.27}$$

where

$$\beta_{\min}(z_1) = \left(\sqrt{NE_{\mathrm{s}}} - z_1\right)\sqrt{\frac{d_{\min}}{N - d_{\min}}}, \tag{2.28}$$

z_2 is the tangential component of the noise vector $\mathbf{z}$, which is on the plane that contains the signals $\mathbf{s}_0$, $\mathbf{s}_{i_{\min}}$ and the origin (see Fig. 2.1), and the other parameters are introduced in (2.4).

For expressing the probabilities of the form $\Pr(E_{0\to i}, E^c_{0\to i_{\min}}, \mathbf{y} \in C_N(\theta) \mid z_1)$ encountered in the RHS of (2.25), we use the three-dimensional geometry in Fig. 2.2-(a). The BPSK modulated signals $\mathbf{s}_0$, $\mathbf{s}_i$ and $\mathbf{s}_j$ are all on the surface of a hyper-sphere centered at the origin and with radius $\sqrt{NE_{\rm s}}$. The planes P_1 and P_2 are constructed by the points $(\mathbf{o}, \mathbf{s}_0, \mathbf{s}_i)$ and $(\mathbf{o}, \mathbf{s}_0, \mathbf{s}_j)$, respectively. In the derivation of the ITSB, Yousefi and Mehrabian choose $\mathbf{s}_j$ to correspond to codeword $\mathbf{c}_j$ with Hamming weight $d_{\min}$. Let z'_3 be the noise component which is orthogonal to z_1 and which lies on the plane P_2 (see Fig 2.2-a). Based on the geometry in Fig. 2.2-a (the probability of the event $E^c_{0\to j}$ is the probability that $\mathbf{y}$ falls in the dashed area) we obtain the following equality if $z_1 \le \sqrt{NE_{\rm s}}$:

$$\begin{aligned}&\Pr(E_{0\to i}, E^c_{0\to i_{\min}}, \mathbf{y} \in C_N(\theta) \mid z_1)\\ &= \Pr\left(\beta_i(z_1) \le z_2 \le r_{z_1},\ -r_{z_1} \le z'_3 \le \beta_{\min}(z_1),\ Y < r^2_{z_1} \mid z_1\right).\end{aligned} \tag{2.29}$$

Furthermore, from the geometry in Fig. 2.2-b, it follows that

$$z'_3 = z_3 \sin\phi + z_2 \cos\phi. \tag{2.30}$$

where z_3 is the noise component which is orthogonal to both z_1 and z_2, and which resides in the three-dimensional space that contains the signal vectors $\mathbf{s}_0$, $\mathbf{s}_i$, $\mathbf{s}_{i_{\min}}$ and the origin. Plugging (2.30) into the condition $-r_{z_1} \le z'_3 \le \beta_{\min}(z_1)$ in (2.29) yields the condition $-r_{z_1} \le z_3 \le \min\{l(z_1, z_2), r_{z_1}\}$ where

$$l(z_1, z_2) = \frac{\beta_{\min}(z_1) - \rho z_2}{\sqrt{1-\rho^2}} \tag{2.31}$$

and $\rho = \cos\phi$ is the correlation coefficient between planes P_1 and P_2. Let $W = \sum_{i=4}^{N} z_i^2$, then if $z_1 \le \sqrt{NE_{\rm s}}$

$$\begin{aligned}&\Pr(E_{0\to i}, E^c_{0\to i_{\min}}, \mathbf{y} \in C_N(\theta) \mid z_1)\\ &= \Pr\left(\beta_i(z_1) \le z_2 \le r_{z_1},\ -r_{z_1} \le z_3 \le \min\{l(z_1, z_2), r_{z_1}\},\ W < r^2_{z_1} - z_2^2 - z_3^2 \mid z_1\right).\end{aligned} \tag{2.32}$$

The random variable W is Chi-squared distributed with $N-3$ degrees of freedom, so its *pdf* is given by

$$f_W(w) = \frac{w^{\frac{N-5}{2}} e^{-\frac{w}{2\sigma^2}} U(w)}{2^{\frac{N-3}{2}} \sigma^{N-3} \Gamma\left(\frac{N-3}{2}\right)}, \quad w \geq 0. \tag{2.33}$$

Since the probabilities of the form $\Pr(E_{0\to i}, E^c_{0\to i_{\min}}, \mathbf{y} \in C_N(\theta) \mid z_1)$ depend on the correlation coefficients between the planes $(\mathbf{o}, \mathbf{s}_0, \mathbf{s}_{i_{\min}})$ and $(\mathbf{o}, \mathbf{s}_0, \mathbf{s}_i)$, the overall upper bound requires the characterization of the global geometrical properties of the code and not only the distance spectrum. To circumvent this problem and obtain an upper bound which is solely depends on the distance spectrum of the code, it is suggested in [47] to loosen the bound as follows. It is shown [46, Appendix B] that the correlation coefficient ρ, corresponding to codewords with Hamming weights d_i and d_j satisfies

$$-\min\left\{\sqrt{\frac{d_i d_j}{(N-d_i)(N-d_j)}}, \sqrt{\frac{(N-d_i)(N-d_j)}{d_i d_j}}\right\} \leq \rho \leq \frac{\min(d_i, d_j)[N - \max(d_i, d_j)]}{\sqrt{d_i d_j (N-d_i)(N-d_j)}}. \tag{2.34}$$

Moreover, the RHS of (2.32) is shown to be a monotonic decreasing function of ρ (see [47, Appendix 1]). Hence, one can omit the dependency in the geometry of the code (and loosen the upper bound) by replacing the correlation coefficients in (2.32) with their lower bounds which solely depend on the weights of the codewords. In the derivation of the ITSB, we consider the correlation coefficients between two planes which correspond to codewords with Hamming weights $d_i = h,\ h \geq N$ and $d_j = d_{\min}$. Let

$$\begin{aligned} \rho_h &\triangleq -\min\left\{\sqrt{\frac{h d_{\min}}{(N-h)(N-d_{\min})}}, \sqrt{\frac{(N-h)(N-d_{\min})}{h d_{\min}}}\right\} \\ &= -\sqrt{\frac{h d_{\min}}{(N-h)(N-d_{\min})}}, \end{aligned} \tag{2.35}$$

where the last equality follows directly from the basic property of $d_{\min}$ as the minimum distance of the code. From (2.25)–(2.26) and averaging w.r.t. Z_1, one gets the

following upper bound on the decoding error probability:

$$\begin{aligned}\Pr(E) \le \Pr\Big(z_1 \le \sqrt{NE_{\rm s}},\ \beta_{\min}(z_1) \le z_2 \le r_{z_1}, V \le r_{z_1}^2 - z_2^2\Big) \\ + \sum_{h=d_{\min}}^{N} A_h \Pr\Big(z_1 \le \sqrt{NE_{\rm s}},\ \beta_h(z_1) \le z_2 \le r_{z_1}, \\ - r_{z_1} \le z_3 \le \min\{l_h(z_1,z_2), r_{z_1}\}, W \le r_{z_1}^2 - z_2^2 - z_3^2\Big) \\ + \Pr\Big(z_1 \le \sqrt{NE_{\rm s}},\ Y \ge r_{z_1}^2\Big) + \Pr(z_1 > \sqrt{NE_{\rm s}}) \end{aligned} \tag{2.36}$$

where the parameter $l_h(z_1, z_2)$ is simply $l(z_1, z_2)$ in (2.31) with ρ replaced by ρ_h, i.e.,

$$l_h(z_1, z_2) \triangleq \frac{\beta_{\min}(z_1) - \rho_h z_2}{\sqrt{1-\rho_h^2}}. \tag{2.37}$$

Using the probability density functions of the random variables in the RHS of (2.36), and since the random variables Z_1, Z_2, Z_3 and W are statistically independent, the final form of the ITSB is given by

$$\begin{aligned} P_{\rm e} \le \int_{-\infty}^{\sqrt{NE_{\rm s}}} \Bigg[\int_{\beta_{\min}}^{r_{z_1}} f_{Z_2}(z_2) \int_0^{r_{z_1}^2 - z_2^2} f_V(v) dv \cdot dz_2 \\ + \sum_{h:\beta_h(z_1) < r_{z_1}} \Bigg(A_h \int_{\beta_h(z_1)}^{r_{z_1}} \int_{-r_{z_1}}^{\min\{l_h(z_1,z_2), r_{z_1}\}} f_{Z_2,Z_3}(z_2,z_3) \int_0^{r_{z_1}^2 - z_2^2 - z_3^2} f_W(w) dw \cdot dz_2 \cdot dz_3 \Bigg) \\ + 1 - \gamma\left(\frac{N-1}{2}, \frac{r_{z_1}^2}{2\sigma^2}\right) \Bigg] f_{Z_1}(z_1) dz_1 + Q\left(\sqrt{\frac{2NRE_{\rm b}}{N_0}}\right). \end{aligned} \tag{2.38}$$

Note that $V \triangleq \sum_{i=3}^{N} z_i^2$ and $W \triangleq \sum_{i=4}^{N} z_i^2$ are Chi-squared distributed with $(N-2)$ and $(N-3)$ degrees of freedom, respectively.

2.2.4 Added-Hyper-Plane (AHP) Bound

In [46], Yousefi and Khandani introduce a new upper bound on the ML decoding block error probability, called the added hyper plane (AHP) bound. In similarity with the ITSB, the AHP bound is based on using the Hunter bound (2.22) as an upper bound on the LHS of (2.9), which results in the inequality (2.24). The complex optimization problem in (2.24), however, is treated differently. Let us denote by $\mathcal{I}_w$ the set of the indices of the codewords of $\mathcal{C}$ with Hamming weight w. For $i \in \{1, 2, \ldots, M\} \setminus \mathcal{I}_w$,

let $\{j_i\}$ be a sequence of integers chosen from the set $\mathcal{I}_w$. Then the following upper bound holds

$$\begin{aligned}&\Pr\left(E(z_1), \mathbf{y} \in C_N(\theta) \mid z_1\right)\\ &\leq \min_{w,\mathcal{J}_w}\left\{\Pr\left(\bigcup_{j\in\mathcal{I}_w}\left\{E_{0\to j}\right\}, \mathbf{y}\in C_N(\theta)\mid z_1\right) + \sum_{i\in\{1,\ldots,M-1\}\setminus\mathcal{I}_w}\Pr\left(E_{0\to i}, E^c_{0\to j_i}, \mathbf{y}\in C_N(\theta)\mid z_1\right)\right\}.\end{aligned} \tag{2.39}$$

The probabilities inside the summation in the RHS of (2.39) are evaluated in a similar manner to the probabilities in the LHS of (2.29). From the analysis in Section 2.2.3 and the geometry in Fig. 2.2-(b), it is clear that the aforementioned probabilities depend on the correlation coefficients between the planes $(\mathbf{o}, \mathbf{s}_0, \mathbf{s}_i)$ and $(\mathbf{o}, \mathbf{s}_0, \mathbf{s}_{j_i})$. Hence, in order to compute the upper bound (2.39), one has to know the geometrical characterization of the Voronoi regions of the codewords. To obtain an upper bound requiring only the distance spectrum of the code, Yousefi and Khandani suggest to extend the codebook by adding all the $\binom{N}{w} - A_w$ N-tuples with Hamming weight w (i.e., the extended code contains all the binary vectors of length N and Hamming weight w). Let us designate the new code by $\mathcal{C}_w$ and denote its codewords by

$$\mathbf{c}_i^w, \quad i \in \left\{0, 1, \ldots, M + \binom{N}{w} - A_w - 1\right\}.$$

The new codebook is not necessarily linear, and all possible correlation coefficients between two codewords with Hamming weight i, where $i \in \{d_{\min}, \ldots d_{\max}\}$, and w are available. Thus, for each layer of the codebook, one can choose the *largest available correlation*[1] ρ with respect to any possible N-tuple binary vector of Hamming weight w. Now one may find the optimum layer at which the codebook extension is done, i.e., finding the optimum $w \in \{1, 2, \ldots n\}$ which yields the tightest upper bound within this form. We note that the resulting upper bound is not proved to be uniformly tighter than the TSB, due to the extension of the code. The maximum correlation coefficient between two codewords of Hamming weight d_i and d_j is introduced in the RHS of (2.34) (see [46]). Let us designate the maximal possible correlation coefficient between two N-tuples with Hamming weights w and h by $\rho_{w,h}$, i.e.,

$$\rho_{w,h} = \frac{\min(h,w)[N - \max(h,w)]}{\sqrt{hw(N-h)(N-w)}}, \quad w \neq h. \tag{2.40}$$

[1]The RHS of (2.39) is a monotonically decreasing function of ρ, as noted in [47].

By using the same bounding technique of the ITSB, and replacing the correlation coefficients with their respective upper bounds, $\rho_{w,h}$, (2.39) gets the form

$$\Pr\left(E(z_1), \mathbf{y} \in C_N(\theta) \mid z_1\right) \leq \min_w \Bigg\{ \Pr\left(\bigcup_{j:w_{\mathrm{H}}(\mathbf{c}_j^w)=w} \{E_{0\to j}\}, \mathbf{y} \in C_N(\theta) \mid z_1 \right) + \sum_{h\neq w} A_h \Pr\left(Y \leq r_{z_1}^2, \beta_h(z_1) \leq z_2,\ z_3 \leq l_{w,h}(z_1, z_2) \mid z_1\right) \Bigg\} \tag{2.41}$$

where

$$l_{w,h}(z_1, z_2) = \frac{\beta_w(z_1) - \rho_{w,h} z_2}{\sqrt{1-\rho_{w,h}^2}}. \tag{2.42}$$

Now, applying Hunter bound on the first term in the RHS of (2.41) yields

$$\begin{aligned} &\Pr\left(\bigcup_{j:w_{\mathrm{H}}(\mathbf{c}_j^w)=w} E_{0\to j},\ \mathbf{y} \in C_N(\theta) \mid z_1 \right) \\ &\leq \Pr(E_{0\to l_0}, \mathbf{y} \in C_N(\theta) \mid z_1) + \sum_{i=1}^{\binom{N}{w}-1} \Pr(E_{0\to l_i}, E^c_{0\to \hat{l}_i} \mathbf{y} \in C_N(\theta) \mid z_1) \end{aligned} \tag{2.43}$$

where $\{l_i\},\ i \in \left\{0, 1, \ldots, \binom{N}{w} - 1\right\}$ is a sequence which designates the indices of the codewords of $\mathcal{C}_w$ with Hamming weight w with an arbitrary order, and $\hat{l}_i \in (l_0, l_1, \ldots, l_{i-1})$. In order to obtain the tightest bound on the LHS of (2.43) in this approach, one has to order the error events such that the correlation coefficients which correspond to codewords $\mathbf{c}_{l_i}$ and $\mathbf{c}_{\hat{l}_i}$ get their maximum available value, which is $1 - \frac{N}{w(N-w)}$ [46, Appendix D]. Let us designate this value by $\rho_{w,w}$,i.e.,

$$\rho_{w,w} = 1 - \frac{N}{w(N-w)} \quad , w \notin \{0, N\}.$$

Hence, based on the geometry in Fig. 2.2, if $z_1 \leq \sqrt{NE_{\mathrm{s}}}$, we can rewrite (2.43) as

$$\begin{aligned} &\Pr\left(\bigcup_{j:w_{\mathrm{H}}(\mathbf{c}_j^w)=w} E_{0\to j}, \mathbf{y} \in C_N(\theta) \mid z_1 \right) \\ &\leq \Pr\left(\beta_w(z_1) \leq z_2 \leq r_{z_1}, V \leq r_{z_1}^2 - z_2^2 \mid z_1\right) \\ &+ \left[\binom{N}{w} - 1\right] \Pr\left(\beta_w(z_1) \leq z_2 \leq r_{z_1}, -r_{z_1} \leq z_3 \leq \min\{l_{w,w}(z_1, z_2), r_{z_1}\}, W \leq r_{z_1}^2 - z_2^2 - z_3^2 \mid z_1\right) \end{aligned} \tag{2.44}$$

where

$$l_{w,w}(z_1, z_2) = \frac{\beta_w(z_1) - \rho_{w,w} z_2}{\sqrt{1 - \rho_{w,w}^2}}. \tag{2.45}$$

By replacing the first term in the RHS of (2.41) with the RHS of (2.44), plugging the result in (2.7) and averaging w.r.t. Z_1 finally gives the following upper bound on the block error probability:

$$\begin{aligned}
\Pr(E) \leq \min_w \Bigg\{ & \Pr\left(z_1 \leq \sqrt{NE_\mathrm{s}},\ \beta_w(z_1) \leq z_2 \leq r_{z_1}, V \leq r_{z_1}^2 - z_2^2\right) \\
& + \binom{N}{w} \Pr\Big(z_1 \leq \sqrt{NE_\mathrm{s}},\ \beta_w(z_1) \leq z_2 \leq r_{z_1}\ , \\
& \qquad -r_{z_1} \leq z_3 \leq \min\{l_{w,w}(z_1, z_2), r_{z_1}\}, W \leq r_{z_1}^2 - z_2^2 - z_3^2\Big) \\
& + \sum_{h \neq w} A_h \Pr\Big(z_1 \leq \sqrt{NE_\mathrm{s}},\ \beta_h(z_1) \leq z_2 \leq r_{z_1}, \\
& \qquad -r_{z_1} \leq z_3 \leq \min\{l_{w,h}(z_1, z_2), r_{z_1}\}, W \leq r_{z_1}^2 - z_2^2 - z_3^2\Big) \Bigg\} \\
& + \Pr\left(z_1 \leq \sqrt{NE_\mathrm{s}},\ Y \geq r_{z_1}^2\right) \Bigg\} + \Pr\left(z_1 > \sqrt{NE_\mathrm{s}}\right).
\end{aligned} \tag{2.46}$$

Rewriting the RHS of (2.46) in terms of probability density functions, the AHP bound gets the form

$$\begin{aligned}
P_\mathrm{e} \leq \min_w \Bigg\{ & \int_{-\infty}^{\sqrt{NE_\mathrm{s}}} \Bigg[\int_{\beta_w(z_1)}^{r_{z_1}} f_{Z_2}(z_2) \int_0^{r_{z_1}^2 - z_2^2} f_V(v) dv \cdot dz_2 \\
& + \binom{N}{w} \int_{\beta_w(z_1)}^{r_{z_1}} \int_{-r_{z_1}}^{\min\{l_{w,w}(z_1,z_2), r_{z_1}\}} f_{Z_2,Z_3}(z_2, z_3) \int_0^{r_{z_1}^2 - z_2^2 - z_3^2} f_W(w) dw \cdot dz_2 \cdot dz_3 \\
& + \sum_{\substack{h\,:\,\beta_h(z_1) < r_{z_1} \\ h \neq w}} \Bigg(A_h \int_{\beta_h(z_1)}^{r_{z_1}} \int_{-r_{z_1}}^{\min\{l_{w,h}(z_1,z_2), r_{z_1}\}} f_{Z_2,Z_3}(z_2, z_3) \int_0^{r_{z_1}^2 - z_2^2 - z_3^2} f_W(w) dw \cdot dz_2 \cdot dz_3 \Bigg) \\
& + 1 - \gamma\left(\frac{N-1}{2}, \frac{r_{z_1}^2}{2\sigma^2}\right) \Bigg] f_{Z_1}(z_1) dz_1 \Bigg\} + Q\left(\sqrt{\frac{2NRE_\mathrm{b}}{N_0}}\right)
\end{aligned} \tag{2.47}$$

where V and W are introduced at the end of Section 2.2.3 (after Eq. (2.38)), and the last term in (2.47) follows from (2.13).

2.3 The Error Exponents of the ITSB and AHP Bounds

The ITSB and the AHP bound were originally derived in [46, 47] as upper bounds on the ML decoding error probability of *specific* binary linear block codes. In the following, we discuss the tightness of the new upper bounds for ensemble of codes, as compared to the TSB. The following lemma is also noted in [47].

Lemma 2.1 Let $\mathcal{C}$ be a binary linear block code, and let us denote by ITSB($\mathcal{C}$) and TSB($\mathcal{C}$) the ITSB and TSB, respectively, on the decoding error probability of $\mathcal{C}$. Then

$$\text{ITSB}(\mathcal{C}) \leq \text{TSB}(\mathcal{C}).$$

Proof: Since $\Pr(A, B) \leq \Pr(A)$ for arbitrary events A and B, the lemma follows immediately by comparing the bounds in the RHS of (2.10) and (2.25), reffering to the TSB and the ITSB, respectively. ■

Corollary 1 The ITSB can not exceed the value of the TSB referring to the average error probability of an arbitrary ensemble of binary linear block codes.

Lemma 2.2 The AHP bound is asymptotically (as we let the block length tend to infinity) at least as tight as the TSB.

Proof: To show this, we refer to (2.46), where we choose the layer w at which the extension of the code is done to be N. Hence, the extended code contains at most one codeword with Hamming weight N more than the original code, which has no impact on the error probability for infinitely long codes. The resulting upper bound is evidently not tighter than the AHP (which carries an optimization over w), and it is at least as tight as the TSB (since the joint probability of two events cannot exceed the probabilities of these individual events). ■

The extension of Lemma 2.2 to ensembles of codes is straightforward (by taking the expectation over the codes in an ensemble, the same conclusion in Lemma 2.2 holds also for ensembles). From the above, it is evident that the error exponents of both the AHP bound and the ITSB cannot be below the error exponent of the TSB. In the following, we introduce a lower bound on both the ITSB and the AHP bound. It serves as an intermediate stage to get our main result.

Lemma 2.3 Let $\mathcal{C}$ designate an ensemble of linear codes of length N, whose transmission takes place over an AWGN channel. Let A_h be the number of codewords of Hamming weight h, and let $\mathbb{E}_{\mathcal{C}}$ designate the statistical expectation over the codebooks of an ensemble $\mathcal{C}$. Then both the ITSB and AHP upper bounds on the average ML decoding error probability of $\mathcal{C}$ are lower bounded by the function $\psi(\mathcal{C})$ where

$$\begin{aligned}\psi(\mathcal{C}) \triangleq \min_{w} \Bigg\{ & \mathbb{E}_{\mathcal{C}} \Big[\Pr\Big(z_1 \le \sqrt{NE_{\mathrm{s}}},\ \beta_w(z_1) \le z_2 \le r_{z_1}, V \le r_{z_1}^2 - z_2^2 \Big) \\ & + \sum_h \Bigg\{ A_h \Pr\Big(z_1 \le \sqrt{NE_{\mathrm{s}}},\ \beta_h(z_1) \le z_2 \le r_{z_1}, \\ & \qquad - r_{z_1} \le z_3 \le \min\{l_{w,h}(z_1,z_2), r_{z_1}\}, W \le r_{z_1}^2 - z_2^2 - z_3^2 \Big) \Bigg\} \\ & + \Pr\Big(z_1 \le \sqrt{NE_{\mathrm{s}}},\ Y \ge r_{z_1}^2 \Big) \Big] \Bigg\} \end{aligned} \tag{2.48}$$

and $l_{w,h}(z_1, z_2)$ is defined in (2.42).

Proof: By comparing (2.46) with (2.48), it is easily verified that the RHS of (2.48) is not larger than the RHS of (2.46) (actually, the RHS of (2.48) is just the AHP *without* any extension of the code). Referring to the ITSB, we get

$$\begin{aligned}\mathrm{ITSB}(\mathcal{C}) = \mathbb{E}_{\mathcal{C}} \Big[& \Pr\Big(z_1 \le \sqrt{NE_{\mathrm{s}}},\ \beta_{\min}(z_1) \le z_2 \le r_{z_1}, V \le r_{z_1}^2 - z_2^2 \Big) \\ & + \sum_h \Bigg\{ A_h \Pr\Big(z_1 \le \sqrt{NE_{\mathrm{s}}},\ \beta_h(z_1) \le z_2 \le r_{z_1}, \\ & \qquad - r_{z_1} \le z_3 \le \min\{l_h(z_1,z_2), r_{z_1}\},\ W \le r_{z_1}^2 - z_2^2 - z_3^2 \Big) \Bigg\} \\ & + \Pr\Big(z_1 \le \sqrt{NE_{\mathrm{s}}},\ Y \ge r_{z_1}^2 \Big) \Big] + \Pr\Big(z_1 > \sqrt{NE_{\mathrm{s}}} \Big) \end{aligned}$$

$$
\begin{aligned}
&\geq \min_{w} \Bigg\{ \mathbb{E}_{\mathcal{C}} \Big[\Pr\Big(z_1 \leq \sqrt{NE_{\mathrm{s}}},\ \beta_w(z_1) \leq z_2 \leq r_{z_1}, V \leq r_{z_1}^2 - z_2^2 \Big) \\
&\qquad + \sum_{h} \Bigg\{ A_h \Pr\Big(z_1 \leq \sqrt{NE_{\mathrm{s}}},\ \beta_h(z_1) \leq z_2 \leq r_{z_1}, \\
&\qquad\qquad - r_{z_1} \leq z_3 \leq \min\{ l_{w,h}(z_1, z_2), r_{z_1} \}, W \leq r_{z_1}^2 - z_2^2 - z_3^2 \Big) \Bigg\} \\
&\qquad + \Pr\Big(z_1 \leq \sqrt{NE_{\mathrm{s}}},\ Y \geq r_{z_1}^2 \Big) \Big] \Bigg\} + \Pr\Big(z_1 > \sqrt{NE_{\mathrm{s}}} \Big) \\
&> \psi(\mathcal{C}). \qquad (2.49)
\end{aligned}
$$

The first inequality holds since the ITSB is a monotonically decreasing function w.r.t. the correlation coefficients (see Appendix C). The equality in (2.49) is due to the linearity of the function in (2.49) w.r.t. the distance spectrum, on which the expectation operator is applied, and the last transition follows directly from (2.48). ■

In [46] and [47], the RHS of (2.46) and (2.36), respectively, were evaluated by integrals, which results in the upper bounds (2.47) and (2.38). In [11, Section D], Divsalar introduced an alternative way to obtain a simple, yet asymptotically identical, version of the TSB by using the Chernoff bounding technique. Using this technique we obtain the exponential version of $\psi(\mathcal{C})$. In the following, We use the following notation [11]:

$$
c \triangleq \frac{E_{\mathrm{s}}}{N_0}, \quad \delta \triangleq \frac{h}{N}, \quad \Delta \triangleq \sqrt{\frac{\delta}{1-\delta}}, \quad r(\delta) \triangleq \frac{\ln(A_h)}{N}
$$

where for the sake of clear writing we denote the average spectrum of the ensemble by A_h. We now state the main result of this chapter.

Theorem 2.4 (The error exponent of the AHP and the ITSB bounds coincide with the error exponent of the TSB) The upper bounds ITSB, AHP and the TSB have the same error exponent, which is

$$
E(c) = \min_{0<\delta\leq 1} \left\{ \frac{1}{2} \ln\left(1 - \gamma + \gamma e^{-2r(\delta)}\right) + \frac{\gamma \Delta^2 c}{1 + \gamma \Delta^2} \right\} \qquad (2.50)
$$

where

$$
\gamma = \gamma(\delta) \triangleq \frac{1-\delta}{\delta} \left[\sqrt{\frac{c}{c_0(\delta)} + (1+c)^2 - 1} - (1+c) \right] \qquad (2.51)
$$

and

$$c_0(\delta) \triangleq \left(1 - e^{-2r(\delta)}\right) \frac{1-\delta}{2\delta}. \tag{2.52}$$

Proof: The exponential version of $\psi(\mathcal{C})$ in (2.48) is identical to the exponential version of the TSB (see Appendices A and B). Since $\psi(\mathcal{C})$ does not exceed the AHP and the ITSB, this implies that the error exponents of the AHP and the ITSB are not larger than the error exponent of the TSB. On the other hand, from Lemmas 2.1 and 2.2 it follows that asymptotically, both the AHP and the ITSB are at least as tight as the TSB, so their error exponents are at least as large as the error exponent of the TSB. Combining these results we obtain that the error exponent of the ITSB, AHP and the TSB are all identical. In [11], Divsalar shows that the error exponent of the TSB is determined by (2.50)–(2.52), which concludes the proof of the theorem. ■

Remark 1 The bound on the bit error probability in [33] is exactly the same as the TSB on the block error probability by Poltyrev [31], except that the average distance spectrum $\{A_h\}$ of the ensemble is now replaced by the sequence $\{A'_h\}$ where

$$A'_h = \sum_{w=0}^{NR} \left(\frac{w}{NR}\right) A_{w,h} \ , \quad h \in \{0, \ldots, N\}$$

and $A_{w,h}$ denotes the average number of codewords encoded by information bits of Hamming weight w and having a Hamming weight (after encoding) which is equal to h. Since $A_h = \sum_{w=0}^{NR} A_{w,h}$, then

$$\frac{A_h}{NR} \leq A'_h \leq A_h \ , \quad h \in \{0, \ldots, N\}.$$

The last inequality therefore implies that the replacement of the distance spectrum $\{A_h\}$ by $\{A'_h\}$ (for the analysis of the bit error probability) does not affect the asymptotic growth rate of $r(\delta)$ where $\delta \triangleq \frac{h}{N}$, and hence, the error exponents of the TSB on the block and bit error probabilities coincide.

Remark 2 In [51], Zangl and Herzog suggest a modification of the TSB on the bit error probability. Their basic idea is tightening the bound on the bit error probability when the received vector $\mathbf{y}$ falls outside the cone $\mathcal{R}$ in the RHS of (2.3) (see Fig. 2.1). In the derivation of the version of the TSB on the bit error probability, as suggested

by Sason and Shamai [33], the conditional bit error probability in this case was upper bounded by 1, where Zangl and Herzog [51] refine the bound and provide a tighter bound on the conditional bit error probability when the vector $\mathbf{y}$ falls in the bad region (i.e., when it is outside the cone in Fig. 2.1). Though this modification tightens the bound on the bit error probability at low SNR (as exemplified in [51] for some short linear block codes), it has no effect on the error exponent. The reason is simply because the conditional bit error probability in this case cannot be below $\frac{1}{NR}$ (i.e., one over the dimension of the code), so the bound should still possess the same error exponent. This shows that the error exponent of the TSB versions on the bit error probability, as suggested in [33] and [51], coincide.

Corollary 2 The error exponents of the TSB on the bit error probability coincides with the error exponent of the TSB on the block error probability. Moreover, the error exponents of the TSB on the bit error probability, as suggested by Sason and Shamai [33] and refined by Zangl and Herzog [51], coincide. The common value of these error exponents is explicitly given in Theorem 2.4.

2.4 Summary and Conclusions

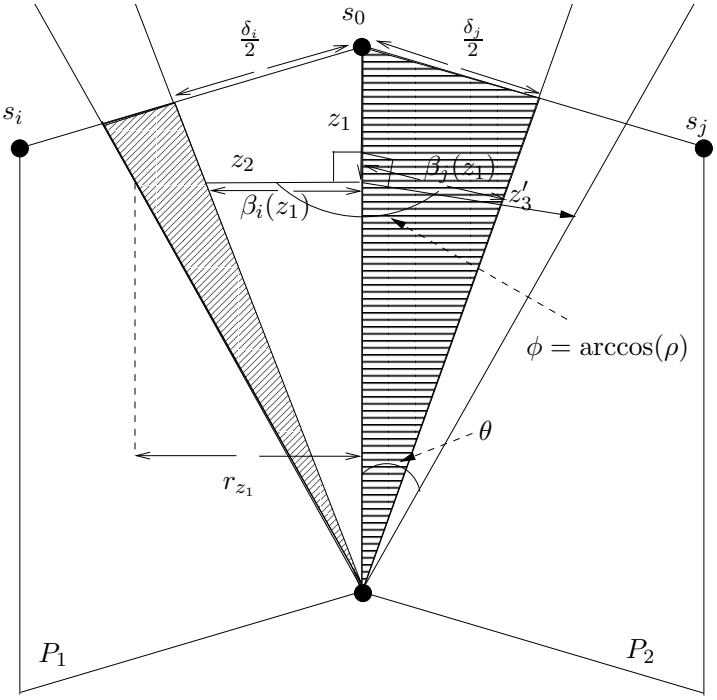
The tangential-sphere bound (TSB) of Poltyrev [31] often happens to be the tightest upper bound on the ML decoding error probability of block codes whose transmission takes place over a binary-input AWGN channel. However, in the random coding setting, it fails to reproduce the random coding error exponent [20] while the second version of the Duman and Salehi (DS2) bound does [15, 35]. The larger the code rate is, the more significant becomes the gap between the error exponent of the TSB and the random coding error exponent of Gallager [20] (see Fig. 2.3, and the plots in [**?**, Figs. 2–4]). In this respect, we note that the expression for the error exponent of the TSB, as derived by Divsalar [11], is significantly easier for numerical calculations than the original expression of this error exponent which was provided by Poltyrev [**?**, Theorem 2]. Moreover, the analysis made by Divsalar is more general in the sense that it applies to an arbitrary ensemble, and not only to the ensemble of fully random block codes.

In this chapter, we consider some recently introduced performance bounds which suggest an improvement over the TSB. These bounds rely solely on the distance spectrum of the code (or their input-output weight enumerators for the analysis of the bit error probability). We study the error exponents of these recently introduced bounding techniques. This work forms a direct continuation to the derivation of these bounds by Yousefi et al. [46, 47, 48] who also exemplified their superiority over the TSB for short binary linear block codes.

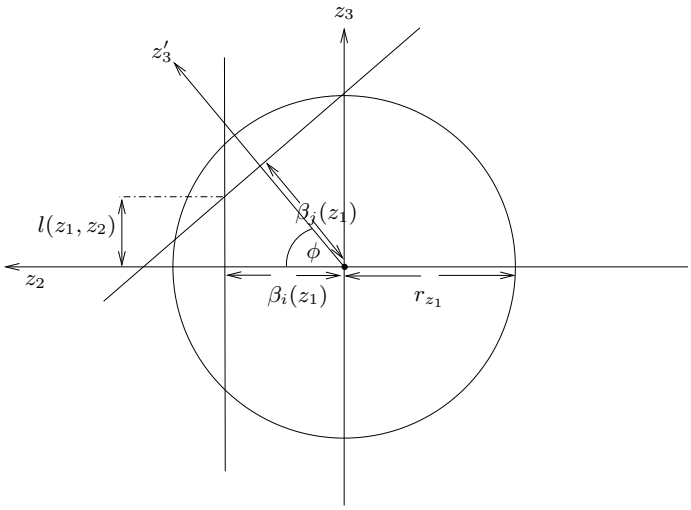
Putting the results reported by Divsalar [11] with the main result in this chapter (see Theorem 2.4), we conclude that the error exponents of the simple bound of Divsalar [11], the first version of Duman and Salehi bounds [14], the TSB [31] and its improved versions by Yousefi et al. [45, 46, 47] all coincide. This conclusion holds for any ensemble of binary linear block codes (e.g., turbo codes, LDPC codes etc.) where we let the block lengths tend to infinity, so it does not only hold for the ensemble of fully random block codes (whose distance spectrum is binomially distributed). Moreover, the error exponents of the TSB versions for the bit error probability, as provided in [33, 51], coincide and are equal to the error exponent of the TSB for the block error probability. The explicit expression of this error exponent is given in Theorem 2.4, and is identical to the expression derived by Divsalar [11] for his simple

bound. Based on Theorem 2.4, it follows that for any value of SNR, the same value of the normalized Hamming weight dominates the exponential behavior of the TSB and its two improved versions. In the asymptotic case where we let the block length tend to infinity, the dominating normalized Hamming weight can be explicitly calculated in terms of the SNR; this calculation is based on finding the value of the normalized Hamming weight δ which achieves the minimum in the RHS of (2.50), where this value clearly depends on the asymptotic growth rate of the distance spectrum of the ensemble under consideration. A similar calculation of this critical weight as a function of the SNR was done in [18], referring to the ensemble of fully random block codes and the simple union bound.

In a the next chapter, new upper bounds on the block and bit error probabilities of linear block codes are derived. These bounds improve the tightness of the Shulman and Feder bound [37] and therefore also reproduce the random coding error exponent.



(a)



(b)

Figure 2.2: (a): $\mathbf{s}_0$ is the transmitted vector, z_1 is the radial noise component, z_2 and z_3' are two (not necessarily orthogonal) noise components, which are perpendicular to z_1, and lie on planes P_1 and P_2, respectively. The doted and dashed areas are the regions where E_i and E_j^c occur, respectively. (b): A cross-section of the geometry in (a).

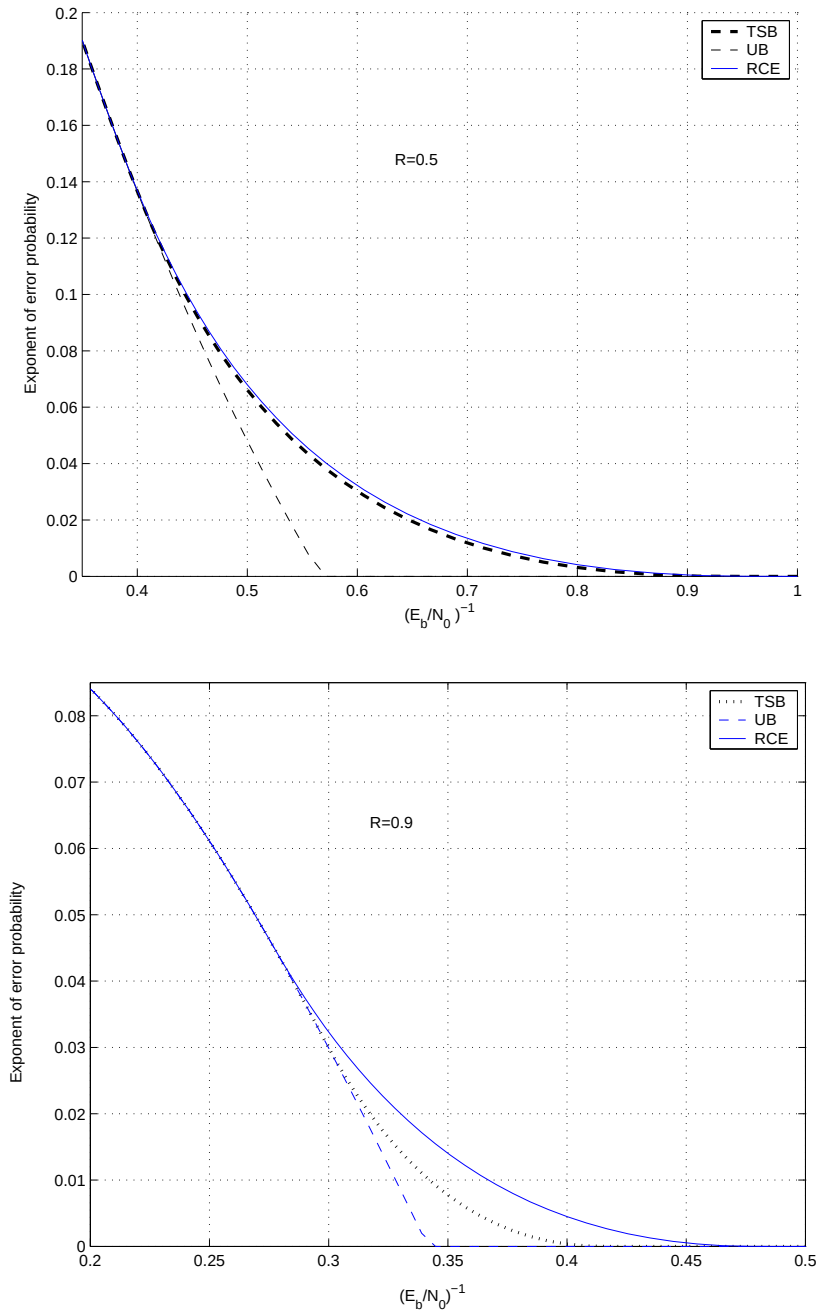


Figure 2.3: Comparison between the error exponents for random block codes which are based on the union bound (UB), the tangential-sphere bound (TSB) of Poltyrev [31] (which according to Theorem 2.4 is identical to the error exponents of the ITSB and the AHP bounds), and the random coding bound (RCE) of Gallager [19]. The upper and lower plots refer to code rates of 0.5 and 0.9 bits per channel use, respectively. The error exponents are plotted versus the reciprocal of the energy per bit to the one-sided spectral noise density.

Chapter 3

Tightened Upper Bounds on the ML Decoding Error Probability of Binary Linear Block Codes

Short overview: The performance of maximum-likelihood (ML) decoded binary linear block codes is addressed via the derivation of tightened upper bounds on their decoding error probability. The upper bounds on the block and bit error probabilities are valid for any memoryless, binary-input and output-symmetric communication channel, and their effectiveness is exemplified for various ensembles of turbo-like codes over the AWGN channel. An expurgation of the distance spectrum of binary linear block codes further tightens the resulting upper bounds.

This chapter is based on the following papers:

- M. Twitto, I. Sason and S. Shamai, "Tightened upper bounds on the ML decoding error probability of binary linear block codes," submitted to the *IEEE Trans. on Information Theory*, February 2006.

- M. Twitto, I. Sason and S. Shamai, "Tightened upper bounds on the ML decoding error probability of binary linear codes," *Proceedings 2006 IEEE International Symposium on Information Theory*, Seattle, USA, July 9–14, 2006.

3.1 Introduction

In this chapter we focus on the upper bounds which emerge from the second version of Duman and Salehi (DS2) bounding technique. The DS2 bound provides a conditional upper bound on the ML decoding error probability given an arbitrary transmitted (length-N) codeword $\mathbf{c}_m$ ($P_{\mathrm{e}|m}$). The conditional decoding error probability is upper bounded by

$$P_{\mathrm{e}|m} \leq \left(\sum_{m' \neq m} \sum_{\mathbf{y}} p_N(\mathbf{y}|\mathbf{c}_m)^{\frac{1}{\rho}} \ \psi_N^m(\mathbf{y})^{1-\frac{1}{\rho}} \ \left(\frac{p_N(\mathbf{y}|\mathbf{c}_{m'})}{p_N(\mathbf{y}|\mathbf{c}_m)} \right)^{\lambda} \right)^{\rho} \tag{3.1}$$

where $0 \leq \rho \leq 1$ and $\lambda \geq 0$ (see [15, 35]; in order to make the presentation self-contained, it will be introduced shortly in the next section as part of the preliminary material). Here, $\psi_N^m(\mathbf{y})$ is an arbitrary probability tilting measure (which may depend on the transmitted codeword $\mathbf{c}_m$), and $p_N(\mathbf{y}|\mathbf{c})$ designates the transition probability measure of the channel.

The tangential-sphere bound (TSB) of Poltyrev often happens to be the tightest upper bound on the ML decoding error probability of block codes whose transmission takes place over a binary-input AWGN channel. However, in the random coding setting, it fails to reproduce the random coding exponent [20] while the second version of the Duman and Salehi (DS2) bound, to be reviewed in the next section, does (see [38]). The Shulman-Feder bound (SFB) can be derived as a particular case of the DS2 bound (see [38]), and it achieves the random coding error exponent. Though the SFB is informative for some structured linear block codes with good Hamming properties, it appears to be rather loose when considering sequences of linear block codes whose minimum distance grows sub-linearly with the block length, as is the case with most capacity-approaching ensembles of LDPC and turbo codes. However, the tightness of this bounding technique is significantly improved by combining the SFB with the union bound; this approach was exemplified for some structured ensembles of LDPC codes (see e.g., [28] and the proof of [36, Theorem 2.2]).

In this chapter, we introduce improved upper bounds on both the bit and block error probabilities. Section 3.2 presents some preliminary material. In Section 3.3, we introduce an upper bound on the block error probability which is in general tighter than the SFB, and combine the resulting bound with the union bound. Similarly, an appropriate upper bound on the bit error probability is introduced. The effect of

an expurgation of the distance spectrum on the tightness of the resulting bounds is considered in Section 3.4. By applying the new bounds to ensembles of turbo-like codes over the binary-input AWGN channel, we demonstrate the usefulness of the new bounds in Section 3.5, especially for some coding structures of high rates. We conclude the chapter in Section 3.6.

3.2 Preliminaries

We introduce in this section some preliminary material which serves as a preparatory step towards the presentation of the material in the following sections.

3.2.1 The DS2 Bound

The bounding technique of Duman and Salehi [14, 15] originates from the 1965 Gallager bound. Let $\psi_N^m(\underline{y})$ designate an arbitrary probability measure (which may also depend on the transmitted codeword $\underline{x}^m$). The Gallager bound [20] can then be put in the form (see [38])

$$\begin{aligned} P_{\mathrm{e}|m} &\leq \sum_{\mathbf{y}} \psi_N^m(\mathbf{y})\ \psi_N^m(\mathbf{y})^{-1}\ p_N(\mathbf{y}|\mathbf{c}_m)\ \left(\sum_{m'\neq m}\left(\frac{p_N(\mathbf{y}|\mathbf{c}_{m'})}{p_N(\mathbf{y}|\mathbf{c}_m)}\right)^{\lambda}\right)^{\rho} \\ &= \sum_{\mathbf{y}} \psi_N^m(\mathbf{y})\ \left(\psi_N^m(\mathbf{y})^{-\frac{1}{\rho}}\ p_N(\mathbf{y}|\mathbf{c}_m)^{\frac{1}{\rho}}\ \sum_{m'\neq m}\left(\frac{p_N(\mathbf{y}|\mathbf{c}_{m'})}{p_N(\mathbf{y}|\mathbf{c}_m)}\right)^{\lambda}\right)^{\rho}, \qquad \forall\ \lambda,\rho\geq 0. \end{aligned} \tag{3.2}$$

By invoking the Jensen inequality in (3.2) for $0\leq\rho\leq 1$, the DS2 bound results

$$P_{\mathrm{e}|m} \leq \left(\sum_{m'\neq m}\sum_{\mathbf{y}} p_N(\mathbf{y}|\mathbf{c}_m)^{\frac{1}{\rho}}\ \psi_N^m(\mathbf{y})^{1-\frac{1}{\rho}}\ \left(\frac{p_N(\mathbf{y}|\mathbf{c}_{m'})}{p_N(\mathbf{y}|\mathbf{c}_m)}\right)^{\lambda}\right)^{\rho}, \qquad 0\leq\rho\leq 1,\ \lambda\geq 0. \tag{3.3}$$

Let $G_N^m(\mathbf{y})$ be an arbitrary non-negative function of $\mathbf{y}$, and let the probability density function $\psi_N^m(\mathbf{y})$ be

$$\psi_N^m(\mathbf{y}) = \frac{G_N^m(\mathbf{y})\ p_N(\mathbf{y}|\mathbf{c}_m)}{\sum_{\mathbf{y}} G_N^m(\mathbf{y})\ p_N(\mathbf{y}|\mathbf{c}_m)} \tag{3.4}$$

The functions $G_N^m(\mathbf{y})$ and $\psi_N^m(\mathbf{y})$ are referred to as the un-normalized and normalized tilting measures, respectively. The substitution of (3.4) into (3.3) yields the following

upper bound on the conditional ML decoding error probability

$$P_{\mathrm{e}|m} \leq \left(\sum_{\mathbf{y}} G_N^m(\mathbf{y}) \, p_N(\mathbf{y}|\mathbf{c}_m) \right)^{1-\rho} \cdot \left(\sum_{m' \neq m} \sum_{\mathbf{y}} p_N(\mathbf{y}|\mathbf{c}_m) \, G_N^m(\mathbf{y})^{1-\frac{1}{\rho}} \, \left(\frac{p_N(\mathbf{y}|\mathbf{c}_{m'})}{p_N(\mathbf{y}|\mathbf{c}_m)} \right)^{\lambda} \right)^{\rho}, \; 0 \leq \rho \leq 1, \; \lambda \geq 0. \tag{3.5}$$

The upper bound (3.5) was also derived in [11, Eq. (62)].

For the case of memoryless channels, and for the choice of $\psi_N^m(\mathbf{y})$ as $\psi_N^m(\mathbf{y}) = \prod_{i=1}^{N} \psi^m(y_i)$ (recalling that the function ψ_N^m may depend on the transmitted codeword $\mathbf{x}^m$), the upper bound (3.3) is relatively easily evaluated (similarly to the standard union bounds) for linear block codes. In that case, (3.3) is calculable in terms of the distance spectrum of the code, not requiring the fine details of the code structure. Moreover, (3.3) is also amenable to some generalizations, such as for the class of discrete memoryless channels with arbitrary input and output alphabets.

3.2.2 The Shulman and Feder bound

We consider here the transmission of a binary linear block code $\mathcal{C}$ where the communication takes place over a memoryless binary-input output-symmetric (MBIOS) channel. The analysis refers to the decoding error probability under soft-decision ML decoding.

The Shulman and Feder bound (SFB) [37] on the block error probability of an (N, K) binary linear block code $\mathcal{C}$, transmitted over a memoryless channel is given by

$$P_{\mathrm{e}} \leq 2^{-N E_{\mathrm{r}}(R + \frac{\log \alpha(\mathcal{C})}{N})} \tag{3.6}$$

where

$$E_{\mathrm{r}}(R) = \max_{0 \leq \rho \leq 1} \left(E_0(\rho) - \rho R \right) \tag{3.7}$$

$$E_0(\rho) \triangleq -\log_2 \left\{ \sum_{y} \left[\frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}} \right]^{1+\rho} \right\}. \tag{3.8}$$

E_{r} is the random coding error exponent [20], $R \triangleq \frac{K}{N}$ designates the code rate in bits per channel use, and

$$\alpha(\mathcal{C}) \triangleq \max_{1 \le l \le N} \frac{A_l}{2^{-N(1-R)} \binom{N}{l}}. \tag{3.9}$$

In the RHS of (3.9), $\{A_l\}$ denotes the distance spectrum of the code. Hence, for fully random block codes, $\alpha(\mathcal{C})$ is equal to 1, and the Shulman-Feder bound (SFB) particularizes to the random coding bound [20]. In general, the parameter $\alpha(\mathcal{C})$ in the SFB (3.6) measures the maximal ratio of the distance spectrum of a code (or ensemble) and the average distance spectrum which corresponds to fully random block codes of the same block length and rate.

The original proof of the SFB is quite involved. In [38], a simpler proof of the SFB is derived, and by doing so, the simplified proof reproduces the SFB as a particular case of the DS2 bound (see Eq. (3.3)). In light of the significance of the proof concept to the continuation of this chapter, we outline this proof briefly.

Since we deal with linear block codes and the communication channel is memoryless, binary-input output-symmetric channel (MBIOS), one can assume without any loss of generality that the all zero codeword $\mathbf{c}_0$ is the transmitted vector. In order to facilitate the expression of the upper bound (3.5) in terms of distance spectrum of the block code $\mathcal{C}$, we consider here the case where the un-normalized tilting measure $G_N^0(\mathbf{y})$ can be expressed in the following product form:

$$G_N^0(\mathbf{y}) = \prod_{i=1}^{N} g(y_i) \tag{3.10}$$

where g is an arbitrary non-negative scalar function, and the channel is by assumption MBIOS, so that the transition probability measure is expanded in the product form

$$p_N(\mathbf{y}|\mathbf{c}_{m'}) = \prod_{i=1}^{N} p(y_i|c_{m',i}) \tag{3.11}$$

where $\mathbf{c}_{m'} = (c_{m',1}, \ldots, c_{m',N})$. Hence, the upper bound on the conditional ML decoding error probability given in (3.5) can be rewritten as

$$\begin{aligned} P_{\mathrm{e}} &= P_{\mathrm{e}|0} \\ &\leq \left(\sum_y g(y)\, p(y|0) \right)^{N(1-\rho)} \\ &\quad \cdot \left\{ \sum_{l=1}^{N} A_l \left(\sum_y g(y)^{1-\frac{1}{\rho}} p(y|0) \right)^{N-l} \left(\sum_y g(y)^{1-\frac{1}{\rho}} p(y|0)^{1-\lambda} p(y|1)^{\lambda} \right)^{l} \right\}^{\rho} \quad \begin{array}{c} \lambda \geq 0, \\ 0 \leq \rho \leq 1 \end{array} \\ &\leq \left(\max_{0<l\leq N} \frac{A_l}{2^{-N(1-R)} \binom{N}{l}} \right)^{\rho} \left(\sum_y g(y)\, p(y|0) \right)^{N(1-\rho)} 2^{-N(1-R)\rho} \\ &\quad \cdot \left\{ \sum_y g(y)^{1-\frac{1}{\rho}} p(y|0) + \sum_y g(y)^{1-\frac{1}{\rho}} p(y|0)^{1-\lambda} p(y|1)^{\lambda} \right\}^{N\rho}. \end{aligned} \tag{3.12}$$

By setting

$$g(y) = \left[\frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}} \right]^{\rho} p(y|0)^{-\frac{\rho}{1+\rho}}, \; \lambda = \frac{1}{1+\rho} \tag{3.13}$$

and using the symmetry of the channel (where $p(y|0) = p(-y|1)$), the SFB follows readily.

3.3 Improved Upper Bounds

3.3.1 Upper Bound on the Block Error Probability

It is well known that at rates below the channel capacity, the block error probability of the ensemble of fully random block codes vanishes exponentially with the block length. In the random coding setting, the TSB [31] fails to reproduce the random coding exponent, while the SFB [37] particularizes to the 1965 Gallager bound for random codes, and hence, the SFB reproduces the random coding exponent. The SFB is therefore advantageous over the TSB in the random coding setting when we let the block length be sufficiently large. Equations (3.6) and (3.9) imply that for specific linear codes (or ensembles), the tightness of the SFB depends on the maximal ratio between the distance spectrum of the code (or the average distance spectrum of

the ensemble) and the average distance spectrum of fully random block codes of the same length and rate which has a binomial distribution.

In order to tighten the SFB bound for linear block codes, Miller and Burshtein [28] suggested to partition the original linear code $\mathcal{C}$ into two subcodes, namely $\mathcal{C}'$ and $\mathcal{C}''$; the subcode $\mathcal{C}'$ contains the all-zero codeword and all the codewords with Hamming weights of $l \in \mathcal{U} \subseteq \{1, 2, ..., N\}$, while $\mathcal{C}''$ contains the other codewords which have Hamming weights of $l \in \mathcal{U}^{\mathrm{c}} = \{1, 2, ..., N\} \setminus \mathcal{U}$ and the all-zero codeword. From the symmetry of the channel, the union bound provides the following upper bound on the ML decoding error probability:

$$P_{\mathrm{e}} = P_{\mathrm{e}|0} \leq P_{\mathrm{e}|0}(\mathcal{C}') + P_{\mathrm{e}|0}(\mathcal{C}'') \tag{3.14}$$

where $P_{\mathrm{e}|0}(\mathcal{C}')$ and $P_{\mathrm{e}|0}(\mathcal{C}'')$ designate the conditional ML decoding error probabilities of $\mathcal{C}'$ and $\mathcal{C}''$, respectively, given that the all zero codeword is transmitted. We note that although the code $\mathcal{C}$ is linear, its two subcodes $\mathcal{C}'$ and $\mathcal{C}''$ are in general *non-linear*. One can rely on different upper bounds on the conditional error probabilities $P_{\mathrm{e}|0}(\mathcal{C}')$ and $P_{\mathrm{e}|0}(\mathcal{C}'')$, i.e., we may bound $P_{\mathrm{e}|0}(\mathcal{C}')$ by the SFB, and rely on an alternative approach to obtain an upper bound on $P_{\mathrm{e}|0}(\mathcal{C}'')$. For example, if we consider the binary-input AWGN channel, then the TSB (or even union bounds) may be used in order to obtain an upper bound on the conditional error probability $P_{\mathrm{e}|0}(\mathcal{C}'')$ which corresponds to the subcode $\mathcal{C}''$. In order to obtain the tightest bound in this approach, one should look for an optimal partitioning of the original code $\mathcal{C}$ into two sub-codes, based on the distance spectrum of $\mathcal{C}$. The solution of the problem is quite tedious, because in general, if the subset $\mathcal{U}$ can be an arbitrary subset of the set of integers $\{1, \ldots, N\}$, then one has to compare $\sum_{i=0}^{N} \binom{N}{i} = 2^N$ different possibilities for $\mathcal{U}$. However, we may use practical optimization schemes to obtain good results which may improve the tightness of both the SFB and TSB.

An easy way to make an efficient partitioning of a linear block code $\mathcal{C}$ is to compare its distance spectrum (or the average distance spectrum for an ensemble of linear codes) with the average distance spectrum of the ensemble of fully random block codes of the same rate and block length. Let us designate the average distance spectrum of the ensemble of fully random block codes of length N and rate R by

$$B_l \triangleq 2^{-N(1-R)} \binom{N}{l} \quad l = 0, 1, \ldots, N. \tag{3.15}$$

Then, it is suggested to partition $\mathcal{C}$ in a way so that all the codewords with Hamming weight l for which $\frac{A_l}{B_l}$ is greater than some threshold (which should be larger than 1 but close to it) are associated with $\mathcal{C}''$, and the other codewords are associated with $\mathcal{C}'$. The following algorithm is suggested for the calculation of the upper bound on the block error probability under ML decoding:

Algorithm 1

1. Set

$$\mathcal{U} = \Phi, \quad \mathcal{U}^{\mathrm{c}} = \{1, 2, ...N\}, \quad l = 1$$

where Φ designates an empty set, and set the initial value of the upper bound to be 1.

2. Compute the ratio $\frac{A_l}{B_l}$ where $\{A_l\}$ is the distance spectrum of the binary linear block code (or the average distance of an ensemble of such codes), and $\{B_l\}$ is the binomial distribution introduced in (3.15).

3. If this ratio is smaller than some threshold (where the value of the threshold is typically set to be slightly larger than 1), then the element l is added to the set $\mathcal{U}$, i.e.,

$$\mathcal{U} := \mathcal{U} + \{l\}, \quad \mathcal{U}^{\mathrm{c}} := \mathcal{U}^{\mathrm{c}} \setminus \{l\}.$$

4. Update correspondingly the upper bound in the RHS of (3.14) (we will derive later the appropriate upper bounds on $P_{\mathrm{e}|0}(\mathcal{C}')$ and $P_{\mathrm{e}|0}(\mathcal{C}'')$.

5. Set the bound to be the minimum between the RHS from Step 4 and its previous value.

6. Set $l = l + 1$ and go to Step 2.

7. The algorithm terminates when l gets the value N (i.e., the block length of the code) or actually, the maximal value of l for which A_l does not vanish.[1]

[1]The number of steps can be reduced by factor of 2 for binary linear codes which contain the all-ones codeword (hence maintain the property $A_l = A_{N-l}$). For such codes, the update equation in Step 3 becomes: $\mathcal{U} := \mathcal{U} + \{l, N-l\}, \quad \mathcal{U}^{\mathrm{c}} := \mathcal{U}^{\mathrm{c}} - \{l, N-l\}$ and the algorithm terminates when l gets the value $\lceil \frac{N}{2} \rceil$.

Fig. **??**(a) shows a plot of the ratio $\frac{A_l}{B_l}$ as a function of $\delta \triangleq \frac{l}{N}$ for an ensemble of uniformly interleaved turbo-random codes. The calculation of the average distance spectrum of these ensemble relies on the results of Soljanin and Urbanke in [40].

From the discussion above, it is clear that the combination of the SFB with another upper bound has the potential to tighten the overall upper bound on the ML decoding probability; this improvement is expected to be especially pronounced for ensembles whose average distance spectrum resembles the binomial distribution of fully random block codes over a relatively large range of Hamming weights, but deviates significantly from the binomial distribution for relatively low and large Hamming weights (e.g., ensembles of uniformly interleaved turbo codes possess this property, as indicated in [33, Section 4]). This bounding technique was successfully applied by Miller and Burshtein [28] and also by Sason and Urbanke [36] to ensembles of regular LDPC codes where the SFB was combined with union bounds. If the range of Hamming weights where the average distance spectrum of an ensemble resembles the binomial distribution is relatively large, then according to the above algorithm, one would expect that $\mathcal{C}'$ typically contains a very large fraction of the overall number of the codewords of a code from this ensemble. Hence, in order to obtain an upper bound on $P_{\mathrm{e}|0}(\mathcal{C}'')$, where $\mathcal{C}''$ is expected to contain a rather small fraction of the codewords in $\mathcal{C}$, we may use a simple bound such as the union bound while expecting not to pay a significant penalty in the tightness of the overall bound on the decoding error probability (P_{e}).

The following bound introduced in Theorem 3.1 is derived as a particular case of the DS2 bound [15]. The beginning of its derivation is similar to the steps in [38, Section 4A], but we later deviate from the analysis there in order to modify the SFB. We finally obtain a tighter version of this bound.

Theorem 3.1 (Modified Shulman and Feder Bound) Let $\mathcal{C}$ be a binary linear block code of length N and rate R, and let $\{A_l\}$ designate its distance spectrum. Let this code be partitioned into two subcodes, $\mathcal{C}'$ and $\mathcal{C}''$, where $\mathcal{C}'$ contains the all-zero codeword and all the other codewords of $\mathcal{C}$ whose Hamming weights are in an arbitrary set $\mathcal{U} \subseteq \{1, 2, , \ldots, N\}$; the second subcode $\mathcal{C}''$ contains the all-zero codeword and the other codewords of $\mathcal{C}$ which are not included in $\mathcal{C}'$. Assume that the communication takes place over a memoryless binary-input output-symmetric (MBIOS) channel with transition probability measure $p(y|x),\ x \in \{0, 1\}$. Then, the block error probability

of $\mathcal{C}$ under ML decoding is upper bounded by

$$P_{\mathrm{e}} \leq P_{\mathrm{e}|0}(\mathcal{C}') + P_{\mathrm{e}|0}(\mathcal{C}'')$$

where

$$P_{\mathrm{e}|0}(\mathcal{C}') \leq \mathrm{SFB}(\rho) \cdot \left[\sum_{l \in \mathcal{U}} \binom{N}{l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)} \right)^{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)} \right)^{N-l} \right]^{\rho}, \ 0 \leq \rho \leq 1 \tag{3.16}$$

$$A(\rho) \triangleq \sum_{y} \left\{ [p(y|0)p(y|1)]^{\frac{1}{1+\rho}} \left[\frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}} \right]^{\rho-1} \right\} \tag{3.17}$$

$$B(\rho) \triangleq \sum_{y} \left\{ p(y|0)^{\frac{2}{1+\rho}} \left[\frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}} \right]^{\rho-1} \right\}. \tag{3.18}$$

The multiplicative term, $\mathrm{SFB}(\rho)$, in the RHS of (3.16) designates the conditional Shulman-Feder upper bound of the subcode $\mathcal{C}'$ given the transmission of the all-zero codeword, i.e.,

$$\mathrm{SFB}(\rho) = 2^{-N\left(E_0(\rho) - \rho\left(R + \frac{\log(\alpha(\mathcal{C}'))}{N}\right)\right)}, \quad 0 \leq \rho \leq 1 \tag{3.19}$$

and E_0 is introduced in (3.8). An upper bound on the conditional block error probability for the subcode $\mathcal{C}''$, $P_{\mathrm{e}|0}(\mathcal{C}'')$, can be either a standard union bound or any other bound.

Proof: Since the block code $\mathcal{C}$ is linear and the channel is MBIOS, the conditional block error probability of $\mathcal{C}$ is independent of the transmitted codeword. Hence, the union bound gives the following upper bound on the block error probability: $P_{\mathrm{e}} \leq P_{\mathrm{e}|0}(\mathcal{C}') + P_{\mathrm{e}|0}(\mathcal{C}'')$.

In order to prove the theorem, we derive an upper bound on $P_{\mathrm{e}|0}(\mathcal{C}')$. Let $\{A_l(\mathcal{C}')\}$ denote the weight spectrum of the subcode $\mathcal{C}'$, and let $G_N(\mathbf{y})$ be an arbitrary non-negative function of the received vector $\mathbf{y} = (y_1, y_2, \ldots, y_N)$ where this function is assumed to be expressible in the product form (3.10). Then, we get from (3.5) and (3.10) the following upper bound on the conditional ML decoding error probability

of the subcode $\mathcal{C}'$:

$$\begin{aligned}
P_{\mathrm{e}|0}(\mathcal{C}') &\leq \left(\sum_y g(y)\, p(y|0)\right)^{N(1-\rho)} \\
&\quad \cdot \left\{\sum_l A_l(\mathcal{C}')\left(\sum_y g(y)^{1-\frac{1}{\rho}} p(y|0)\right)^{N-l}\left(\sum_y g(y)^{1-\frac{1}{\rho}} p(y|0)^{1-\lambda} p(y|1)^{\lambda}\right)^{l}\right\}^{\rho} \quad \begin{array}{l}\lambda \geq 0, \\ 0 \leq \rho \leq 1\end{array} \\
&= \left(\sum_y g(y)\, p(y|0)\right)^{N(1-\rho)} 2^{-N(1-R)\rho} \\
&\quad \cdot \left\{\sum_{l \in \mathcal{U}} \left(\frac{A_l}{2^{-N(1-R)}\binom{N}{l}}\right)\binom{N}{l}\left(\sum_y g(y)^{1-\frac{1}{\rho}} p(y|0)\right)^{N-l}\right. \\
&\qquad \left.\left(\sum_y g(y)^{1-\frac{1}{\rho}} p(y|0)^{1-\lambda} p(y|1)^{\lambda}\right)^{l}\right\}^{\rho} \\
&\leq \left(\max_{l \in \mathcal{U}} \frac{A_l}{2^{-N(1-R)}\binom{N}{l}}\right)^{\rho}\left(\sum_y g(y)\, p(y|0)\right)^{N(1-\rho)} 2^{-N(1-R)\rho} \\
&\quad \cdot \left\{\sum_{l \in \mathcal{U}} \binom{N}{l}\left(\sum_y g(y)^{1-\frac{1}{\rho}} p(y|0)\right)^{N-l}\left(\sum_y g(y)^{1-\frac{1}{\rho}} p(y|0)^{1-\lambda} p(y|1)^{\lambda}\right)^{l}\right\}^{\rho}. \quad (3.20)
\end{aligned}$$

The transition in the first equality above follows since $A_l(\mathcal{C}') \equiv 0$ for $l \notin \mathcal{U}$, and $A_l(\mathcal{C}')$ coincide with the distance spectrum of the code $\mathcal{C}$ for all $l \in \mathcal{U}$. Note that (3.20) is a tighter version of the bound in [38, Eq. (32)]. The difference between the modified and the original bounds is that in the former, we only sum over the indices $l \in \mathcal{U}$ while in the latter, we sum over the whole set of indices, i.e., $l \in \{1, 2, \ldots, N\}$. By setting the tilting measure in (3.13), the symmetry of the MBIOS channel gives the equality

$$\sum_y g(y) p(y|0) = \sum_y \left[\frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}}\right]^{\rho+1} \quad (3.21)$$

and from (3.17) and (3.18)

$$\begin{aligned}
&\sum_y p(y|0)^{1-\lambda} p(y|1)^{\lambda} g(y)^{1-\frac{1}{\rho}} \\
&= \sum_y \left\{ [p(y|0)p(y|1)]^{\frac{1}{1+\rho}} \left[\frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}} \right]^{\rho-1} \right\} \\
&= A(\rho) \qquad (3.22)
\end{aligned}$$

$$\begin{aligned}
&\sum_y p(y|0) g(y)^{1-\frac{1}{\rho}} \\
&= \sum_y \left\{ p(y|0)^{\frac{2}{1+\rho}} \left[\frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}} \right]^{\rho-1} \right\} \\
&= B(\rho). \qquad (3.23)
\end{aligned}$$

where the RHS of (3.22) and (3.23) are obtained by setting $\lambda = \frac{1}{1+\rho}$. Finally, based on (3.13) and the symmetry of the channel, one can verify that

$$\sum_y g(y) p(y|0) = \frac{A(\rho) + B(\rho)}{2}. \qquad (3.24)$$

Substituting (3.21)–(3.24) into (3.20) gives the following conditional upper bound on the ML decoding error probability of the subcode $\mathcal{C}'$:

$$P_{\mathrm{e}|0}(\mathcal{C}') \leq \alpha(\mathcal{C}_2)^{\rho} \left(\frac{A(\rho) + B(\rho)}{2} \right)^{N(1-\rho)} 2^{-N(1-R)\rho} \cdot \left(\sum_{l \in \mathcal{U}} \binom{N}{l} A^l(\rho) B^{N-l}(\rho) \right)^{\rho} \qquad (3.25)$$

where we use the notation

$$\alpha(\mathcal{C}') \triangleq \max_{l \in \mathcal{U}} \frac{A_l}{2^{-N(1-R)} \binom{N}{l}}.$$

The latter parameter measures by how much the (expected) number of codewords in the subcode $\mathcal{C}'$ deviates from the binomial distribution which characterizes the average distance spectrum of the ensemble of fully random block codes of length N

and rate R. By straightforward algebra, we obtain that

$$\begin{aligned}
P_{\mathrm{e}|0}(\mathcal{C}') &\leq \alpha(\mathcal{C}')^{\rho} \left(\frac{A(\rho)+B(\rho)}{2}\right)^{N} 2^{-N(1-R)\rho} \left(\frac{1}{2}\right)^{-N\rho} \\
&\quad \cdot \left[\sum_{l\in\mathcal{U}} \binom{N}{l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l}\right]^{\rho} \\
&= \alpha(\mathcal{C}')^{\rho} \left(\frac{A(\rho)+B(\rho)}{2}\right)^{N} 2^{NR\rho} \left[\sum_{l\in\mathcal{U}} \binom{N}{l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l}\right]^{\rho} \\
&= \mathrm{SFB}(\rho) \cdot \left[\sum_{l\in\mathcal{U}} \binom{N}{l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l}\right]^{\rho}, \ 0 \leq \rho \leq 1.
\end{aligned} \tag{3.26}$$

The second equality follows from (3.19) and (3.8), and since

$$\begin{aligned}
E_0(\rho) &\triangleq -\log_2 \left\{\sum_{y} \left[\frac{1}{2} p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2} p(y|1)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right\} \\
&= -\log_2 \left(\frac{A(\rho)+B(\rho)}{2}\right).
\end{aligned} \tag{3.27}$$

This concludes the proof of the theorem. ∎

Discussion: The improvement of the bound introduced in Theorem 3.1 over the standard combination of the SFB and the union bound [28, 36] stems from the introduction of the factor which multiplies $\mathrm{SFB}(\rho)$ in the RHS of (3.16); this multiplicative term cannot exceed 1 since

$$\begin{aligned}
&\sum_{l\in\mathcal{U}} \binom{N}{l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l} \\
&\leq \sum_{l=0}^{N} \binom{N}{l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l} = 1.
\end{aligned}$$

This multiplicative factor which appears in the new bound is useful for finite-length codes with small to moderate block lengths. The upper bound (3.16) on $P_{\mathrm{e}|0}(\mathcal{C}')$ is clearly at least as tight as the corresponding conditional SFB. We refer to the upper bound (3.16) as the modified SFB (MSFB). The conditional block error probability of the subcode $\mathcal{C}''$, given that the all-zero codeword is transmitted, can be bounded

by a union bound or any improved upper bound conditioned on the transmission of the all-zero codeword (note that the subcode $\mathcal{C}''$ is in general a non-linear code). In general, one is looking for an appropriate balance between the two upper bounds on $P_{\mathrm{e}|0}^{(1)}$ and $P_{\mathrm{e}|0}^{(2)}$ (see Algorithm 1). The improvement that is achieved by using the MSFB instead of the corresponding SFB is exemplified in Section 3.5 for ensembles of uniformly interleaved turbo-Hamming codes.

3.3.2 Upper Bounds on Bit Error Probability

Let $\mathcal{C}$ be a binary linear block code whose transmission takes place over an arbitrary MBIOS channel, and let P_{b} designate the bit error probability of $\mathcal{C}$ under ML decoding. In [34, Appendix A], Sason and Shamai derived an upper bound on the bit error probability of systematic, binary linear block codes which are transmitted over fully interleaved fading channels with perfect channel state information at the receiver. Here we generalize the result of [34] for arbitrary MBIOS channels. In order to derive the desired upper bound we use the following lemma due to Divsalar [11], and provide a simplified proof to this lemma:

Lemma 3.2 [11, Section III.C] Let $\mathcal{C}$ be a binary block code of dimension K whose transmission takes place over an MBIOS channel. Let $\mathcal{C}(w)$ designate a sub-code of $\mathcal{C}$ which includes the all-zero codeword and all the codewords of $\mathcal{C}$ which are encoded by *information bits* whose Hamming weight is w. Then the conditional bit error probability of $\mathcal{C}$ under ML decoding, given that the all-zero codeword is transmitted, is upper bounded by

$$P_{\mathrm{b}|0} \leq \sum_{\mathbf{y}} p_N(\mathbf{y}|\mathbf{0})^{1-\lambda\rho} \left\{ \sum_{w=1}^{K} \left(\frac{w}{K}\right) \sum_{\substack{\mathbf{c} \in \mathcal{C}(w) \\ \mathbf{c} \neq \mathbf{0}}} p_N(\mathbf{y}|\mathbf{c})^{\lambda} \right\}^{\rho}, \qquad \lambda > 0, \;\; 0 \leq \rho \leq 1. \quad (3.28)$$

We introduce here a somewhat simpler proof than in [11].

Proof: The conditional bit error probability under ML decoding admits the form

$$P_{\mathrm{b}|0} = \sum_{\mathbf{y}} \left(\frac{w_0(\mathbf{y})}{K}\right) p_N(\mathbf{y}|\mathbf{0}) \quad (3.29)$$

where $w_0(\mathbf{y}) \in \{0, 1, ..., K\}$ designates the weight of the information bits in the decoded codeword, given the all-zero codeword is transmitted and the received vector

is $\mathbf{y}$. In particular, if the received vector $\mathbf{y}$ is included in the decision region of the all-zero codeword, then $w_0(\mathbf{y}) = 0$. The following inequalities hold:

$$\begin{aligned}
\frac{w_0(\mathbf{y})}{K} &\leq \left(\frac{w_0(\mathbf{y})}{K}\right)^{\rho}, \qquad 0 \leq \rho \leq 1 \\
&\overset{\text{(a)}}{\leq} \left\{ \left(\frac{w_0(\mathbf{y})}{K}\right) \sum_{\substack{\mathbf{c} \in \mathcal{C}(w_0(\mathbf{y})) \\ \mathbf{c} \neq \mathbf{0}}} \left[\frac{p_N(\mathbf{y}|\mathbf{c})}{p_N(\mathbf{y}|\mathbf{0})}\right]^{\lambda} \right\}^{\rho} \qquad \lambda \geq 0 \\
&\leq \left\{ \sum_{w=1}^{K} \left(\frac{w}{K}\right) \sum_{\substack{\mathbf{c} \in \mathcal{C}(w) \\ \mathbf{c} \neq \mathbf{0}}} \left[\frac{p_N(\mathbf{y}|\mathbf{c})}{p_N(\mathbf{y}|\mathbf{0})}\right]^{\lambda} \right\}^{\rho}. \qquad (3.30)
\end{aligned}$$

Inequality (a) holds since the received vector $\mathbf{y}$ falls in the decision region of a codeword $\tilde{\mathbf{c}}$ which is encoded by information bits of total Hamming weight $w_0(\mathbf{y})$; hence, the quotient $\frac{p_N(\mathbf{y}|\tilde{\mathbf{c}})}{p_N(\mathbf{y}|\mathbf{0})}$ is larger than 1 while the other terms in the sum are simply non-negative. The third inequality holds because of adding non-negative terms to the sum. The lemma follows by substituting (3.30) into the RHS of (3.29). ∎

Theorem 3.3 (The SFB Version on the BER) Let $\mathcal{C}$ be a binary linear block code of length N and dimension K, and assume that the transmission of the code takes place over an MBIOS channel. Let $A_{w,l}$ designate the number of codewords in $\mathcal{C}$ which are encoded by information bits whose Hamming weight is w and their Hamming weight after encoding is l. Then, the bit error probability of $\mathcal{C}$ under ML decoding is upper bounded by

$$P_{\mathrm{b}} \leq 2^{-NE_{\mathrm{r}}(R + \frac{\log \alpha_{\mathrm{b}}(\mathcal{C})}{N})} \qquad (3.31)$$

where $R = \frac{K}{N}$ is the code rate of $\mathcal{C}$, and

$$\alpha_{\mathrm{b}}(\mathcal{C}) \triangleq \max_{0<l\leq N} \frac{A'_l}{2^{-N(1-R)}\binom{N}{l}}, \qquad A'_l \triangleq \sum_{w=1}^{K} \left(\frac{w}{K}\right) A_{w,l}.$$

Proof: Due to the linearity of the code $\mathcal{C}$ and the symmetry of the channel, the conditional bit error probability of the code is independent on the transmitted codeword; hence, without any loss of generality, it is assumed that the all-zero codeword

is transmitted. From (3.28), the following upper bound on the bit error probability of $\mathcal{C}$ follows:

$$\begin{aligned} P_{\mathrm{b}} = P_{\mathrm{b}|0} &\leq \sum_{\mathbf{y}} p_N(\mathbf{y}|\mathbf{0})^{1-\lambda\rho} \left\{ \sum_{w=1}^{K} \left(\frac{w}{K}\right) \sum_{\substack{\mathbf{c} \in \mathcal{C}(w) \\ \mathbf{c} \neq \mathbf{0}}} p_N(\mathbf{y}|\mathbf{c})^{\lambda} \right\}^{\rho}, \qquad \lambda > 0, \;\; 0 \leq \rho \leq 1 \\ &= \sum_{\mathbf{y}} \psi_N^0(\mathbf{y}) \left\{ \psi_N^0(\mathbf{y})^{-\frac{1}{\rho}} p_N(\mathbf{y}|\mathbf{0})^{\frac{1}{\rho}} \sum_{w=1}^{K} \left(\frac{w}{K}\right) \sum_{\substack{\mathbf{c} \in \mathcal{C}(w) \\ \mathbf{c} \neq \mathbf{0}}} \left[\frac{p_N(\mathbf{y}|\mathbf{c})}{p_N(\mathbf{y}|\mathbf{0})}\right]^{\lambda} \right\}^{\rho} \end{aligned} \tag{3.32}$$

where ψ_N^0 is an arbitrary probability tilting measure. By invoking Jensen inequality in the RHS of (3.32) and replacing $\psi_N^0(\mathbf{y})$ with the un-normalized tilting measure $G_N^0(\mathbf{y})$ which appears in the RHS of (3.4), the upper bound in (3.32) transforms to

$$\begin{aligned} P_{\mathrm{b}|0} \leq & \left(\sum_{\mathbf{y}} G_N^0(\mathbf{y}) \; p_N(\mathbf{y}|\mathbf{0}) \right)^{1-\rho} \\ & \cdot \left\{ \sum_{w=1}^{k} \left(\frac{w}{k}\right) \sum_{\substack{\mathbf{c} \in \mathcal{C}(w)}} \sum_{\mathbf{y}} p_N(\mathbf{y}|\mathbf{0}) G_N^0(\mathbf{y})^{1-\frac{1}{\rho}} \left[\frac{p_N(\mathbf{y}|\mathbf{c})}{p_N(\mathbf{y}|\mathbf{0})}\right]^{\lambda} \right\}^{\rho}, \; 0 \leq \rho \leq 1, \; \lambda > 0. \end{aligned} \tag{3.33}$$

We consider an un-normalized tilting measure $G_N^0(\mathbf{y})$ which is expressible in the product form (3.10). Since the communication channel is MBIOS and $\mathcal{C}$ is a binary linear block code, one obtains the following upper bound on the bit error probability:

$$\begin{aligned} P_{\mathrm{b}|0} &\leq \left(\sum_{y} g(y) \; p(y|0) \right)^{N(1-\rho)} \qquad 0 \leq \rho \leq 1, \; \lambda > 0 \\ & \quad \cdot \left\{ \sum_{w=1}^{K} \left(\frac{w}{K}\right) \sum_{l=0}^{N} A_{w,l} \left(\sum_{y} p(y|0) g(y)^{1-\frac{1}{\rho}} \right)^{N-l} \left(\sum_{y} p(y|1)^{\lambda} p(y|0)^{1-\lambda} g(y)^{1-\frac{1}{\rho}} \right)^{l} \right\}^{\rho} \\ &= \left(\sum_{y} g(y) \; p(y|0) \right)^{N(1-\rho)} \\ & \quad \cdot \left\{ \sum_{l=0}^{N} A'_{l} \left(\sum_{y} p(y|0) g(y)^{1-\frac{1}{\rho}} \right)^{N-l} \left(\sum_{y} p(y|1)^{\lambda} p(y|0)^{1-\lambda} g(y)^{1-\frac{1}{\rho}} \right)^{l} \right\}^{\rho} \end{aligned}$$

$$\leq \left(\sum_y g(y)\, p(y|0)\right)^{N(1-\rho)} \left(\max_{0\leq l\leq N} \frac{A'_l}{2^{-N(1-R)}\binom{n}{l}}\right)^{\rho} \cdot 2^{-N(1-R)\rho}$$
$$\cdot \left(\sum_y p(y|1)^{\lambda} p(y|0)^{1-\lambda} g(y)^{1-\frac{1}{\rho}} + \sum_y p(y|1)^{\lambda} p(y|0)^{1-\lambda} g(y)^{1-\frac{1}{\rho}}\right)^{N\rho} \quad (3.34)$$

By setting $g(y)$ as in (3.13), we obtain an upper bound which is the same as the original SFB, except that the distance spectrum $\{A_l\}$ is replaced by $\{A_l{}'\}$. This provides the bound introduced in (3.31), and concludes the proof of the theorem. ■

Similarly to the derivation of the combined upper bound on the block error probability in Theorem 3.1, we suggest to partition the code into two subcodes in order to get improved upper bounds on the bit error probability; however, since we consider the bit error probability instead of block error probability, the threshold in Algorithm 1 is typically modified to a value which is slightly above $\frac{1}{2}$ (instead of 1). Since the code is linear and the channel is MBIOS, the conditional decoding error probability is independent of the transmitted codeword (so, we assume again that the all-zero codeword is transmitted). By the union bound

$$P_{\mathrm{b}} = P_{\mathrm{b}|0} \leq P_{\mathrm{b}|0}(\mathcal{C}') + P_{\mathrm{b}|0}(\mathcal{C}'') \quad (3.35)$$

where $P_{\mathrm{b}|0}(\mathcal{C}')$ and $P_{\mathrm{b}|0}(\mathcal{C}'')$ denote the conditional ML decoding bit error probabilities of two disjoint subcodes $\mathcal{C}'$ and $\mathcal{C}''$ which partition the block code $\mathcal{C}$ (except that these two subcodes have the all-zero vector in common), given that the all-zero codeword is transmitted. The construction of the subcodes $\mathcal{C}'$ and $\mathcal{C}''$ is characterized later.

Upper bound on $P_{\mathrm{b}|0}(\mathcal{C}')$: Let $A_{w,l}$ designate the number of codewords of Hamming weight l which are encoded by a sequence of information bits of Hamming weight w. Similarly to the discussion on the block error probability, we use the bit-error version of the SFB (see Eq. (3.31)) as an upper bound on $P_{\mathrm{b}|0}(\mathcal{C}')$. From Theorem 3.3, it follows that the conditional bit error probability of the subcode $\mathcal{C}'$, given that the all-zero codeword is transmitted is upper bounded by

$$P_{\mathrm{b}|0}(\mathcal{C}') \leq 2^{-NE_{\mathrm{r}}\left(R+\frac{\log \alpha_{\mathrm{b}}(\mathcal{C}')}{N}\right)} \quad (3.36)$$

where

$$\alpha_{\mathrm{b}}(\mathcal{C}') \triangleq \max_{l\in\mathcal{U}} \frac{A'_l(\mathcal{C}')}{B_l} \qquad , A_l{}'(\mathcal{C}') \triangleq \begin{cases} \sum_{w=1}^{NR} \left(\frac{w}{NR}\right) A_{w,l} & \text{if } l \in \mathcal{U} \\ 0 & \text{otherwise} \end{cases} \quad (3.37)$$

and the set $\mathcal{U}$ in (3.37) stands for an arbitrary subset of $\{1, \ldots, N\}$.

Upper bound on $P_{\mathrm{b}|0}(\mathcal{C}'')$: We may bound the conditional bit error probability of the subcode $\mathcal{C}''$, $P_{\mathrm{b}|0}(\mathcal{C}'')$, by an improved upper bound. For the binary-input AWGN, the modified version of the TSB, as shown in [33] is an appropriate bound. This bound is the same as the original TSB in (2.19), except that the distance spectrum $\{A_l\}$ is replaced by $\{A_l{}'(\mathcal{C}'')\}$ where

$$A_l{}'(\mathcal{C}'') \triangleq \begin{cases} \displaystyle\sum_{w=1}^{NR} \left(\frac{w}{NR}\right) A_{w,l} & \text{if } l \in \mathcal{U}^{\mathrm{c}} \\ 0 & \text{otherwise} \end{cases} \tag{3.38}$$

and $\mathcal{U}^{\mathrm{c}}$ stands for an complementary set of $\mathcal{U}$ in (3.37), i.e., $\mathcal{U}^{\mathrm{c}} \triangleq \{1, \ldots, N\} \setminus \mathcal{U}$. For the binary-input AWGN channel, the TSB on the conditional bit error probability admits the following final form (see [33]):

$$P_{\mathrm{b}|0}(\mathcal{C}'') \leq \int_{-\infty}^{\infty} \frac{dz_1}{\sqrt{2\pi}\sigma} e^{-\frac{z_1^2}{2\sigma^2}} \left\{ \begin{array}{ll} \sum_{l:\frac{\delta_l}{2} \leq \alpha_l} & \left\{ A_l{}'(\mathcal{C}'') \int_{\beta_l(z_1)}^{r_{z_1}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z_2^2}{2\sigma^2}} \bar{\gamma}\left(\frac{N-2}{2}, \frac{{r_{z_1}}^2 - z_2^2}{2\sigma^2}\right) dz_2 \right\} \\ & +1 - \bar{\gamma}\left(\frac{N-1}{2}, \frac{{r_{z_1}}^2}{2\sigma^2}\right) \end{array} \right\} \tag{3.39}$$

where the incomplete Gamma function $\bar{\gamma}$ is introduced in (2.12). As the simplest alternative to obtain an upper bound on the conditional bit error probability of the subcode $\mathcal{C}'$ given that the all-zero codeword is transmitted, one may use the union bound (UB) for the binary-input AWGN channel

$$\begin{aligned} P_{\mathrm{b}|0}(\mathcal{C}'') &\leq \sum_{w=1}^{NR} \left(\frac{w}{NR}\right) \sum_{l \in \mathcal{U}^{\mathrm{c}}} A_{w,l}\, Q\left(\sqrt{\frac{2lRE_{\mathrm{b}}}{N_0}}\right) \\ &= \sum_{l=1}^{N} A'_l(\mathcal{C}'')\, Q\left(\sqrt{\frac{2lRE_{\mathrm{b}}}{N_0}}\right) \end{aligned} \tag{3.40}$$

where E_{b} is the energy per information bit and $\frac{N_0}{2}$ is the two-sided spectral power density of the additive noise.

In order to tighten the upper bound (3.36), we obtain the bit-error version of the MSFB (see Eq. (3.16)), by following the steps of the proof of Theorem 3.1. In a similar manner to the transition from (3.6) to (3.31), we just need to replace the terms $A_l(\mathcal{C}')$ in (3.16) with $A'_l(\mathcal{C}')$ to get the conditional modified SFB (MSFB) on the bit error probability of $\mathcal{C}'$, given the all-zero codeword is transmitted. The resulting upper bound is expressed in the following theorem:

Theorem 3.4 (Modified SFB on the Bit Error Probability) Let $\mathcal{C}$ be a binary linear block code of length N and rate R, and let $A_{w,l}$ be the number of codewords of $\mathcal{C}$ which are encoded by information bits whose Hamming weight is w and their Hamming weight after encoding is l (where $0 \leq w \leq NR$ and $0 \leq l \leq N$). Let the code $\mathcal{C}$ be partitioned into two subcodes, $\mathcal{C}'$ and $\mathcal{C}''$, where $\mathcal{C}'$ contains all codewords of $\mathcal{C}$ with Hamming weight $l \in \mathcal{U} \subseteq \{1, 2, , \ldots, N\}$ and the all-zero codeword, and $\mathcal{C}''$ contains the all-zero codeword and all the other codewords of $\mathcal{C}$ which are not in $\mathcal{C}'$. Assume that the communication takes place over an MBIOS channel. Then, the bit error probability of $\mathcal{C}$ under ML decoding is upper bounded by

$$P_{\mathrm{b}} \leq P_{\mathrm{b}|0}(\mathcal{C}') + P_{\mathrm{b}|0}(\mathcal{C}'')$$

where

$$P_{\mathrm{b}|0}(\mathcal{C}') \leq 2^{-N\left(E_0(\rho)-\rho\left(R+\frac{\log(\alpha_{\mathrm{b}}(\mathcal{C}'))}{N}\right)\right)} \left[\sum_{l\in\mathcal{U}} \binom{N}{l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l}\right]^{\rho},$$
$$0 \leq \rho \leq 1 \tag{3.41}$$

$$\alpha_{\mathrm{b}}(\mathcal{C}') \triangleq \max_{l\in\mathcal{U}} \frac{A'_l}{2^{-N(1-R)}\binom{N}{l}}, \qquad A'_l \triangleq \sum_{w=1}^{NR} \left(\frac{w}{NR}\right) A_{w,l}$$

and the functions A, B, E_0 are introduced in (3.17), (3.18) and (3.8), respectively. An upper bound on the conditional bit error probability for the subcode $\mathcal{C}''$, $P_{\mathrm{b}|0}(\mathcal{C}'')$, can be either a union bound (3.40), the TSB (3.39) or any other improved bound.

Discussion: Note that $\alpha_{\mathrm{b}}(\mathcal{C}') \leq \alpha(\mathcal{C}')$, therefore the bound on the bit error probability in (3.41) is always smaller than the bound on the block error probability in (3.16), as one could expect.

In the derivation of the MSFB on the conditional block and bit error probabilities (see Eqs. (3.16) and (3.41), respectively), we obtain simplified expressions by taking out the maximum of $\left\{\frac{A_l(\mathcal{C}')}{B_l}\right\}$ and $\left\{\frac{A'_l(\mathcal{C}')}{B_l}\right\}$ from the corresponding summations in (3.20) and (3.34). This simplification was also done in [38] for the derivation of the SFB as a particular case of the DS2 bound. When considering the case of an upper bound on the block error probability, this simplification is reasonable because we consider the terms $\left\{\frac{A_l(\mathcal{C}')}{B_l}\right\}$ which vary slowly over a large range of the Hamming

weights l (see Fig. **??**(a) when referring to ensembles of turbo-like codes whose average distance spectrum resembles the binomial distribution). However, by considering the terms $\left\{\frac{A'_l(\mathcal{C}')}{B_l}\right\}$ whose values change considerably with l and almost grow linearly with l (see Fig. **??**(b)), such simplification previously done for the block error analysis (i.e., taking out the maximal value of $\frac{A'_l(\mathcal{C}')}{B_l}$ from the summation) is expected to significantly reduce the tightness of the bound on the *bit error probability*. Thus, the modification which results in (3.41) does not seem to yield a good upper bound.[2] In order to get a tighter upper bound on the bit error probability we introduce the following theorem:

Theorem 3.5 (Simplified DS2 Bound) Let $\mathcal{C}$ be a binary linear block code of length N and rate R, and let $A_{w,l}$ designate the number of codewords which are encoded by information bits whose Hamming weight is w and their Hamming weight after encoding is l (where $0 \leq w \leq NR$ and $0 \leq l \leq N$). Let the code $\mathcal{C}$ be partitioned into two subcodes, $\mathcal{C}'$ and $\mathcal{C}''$, where $\mathcal{C}'$ contains all the codewords in $\mathcal{C}$ with Hamming weight $l \in \mathcal{U} \subseteq \{1, 2, , \ldots, N\}$ and the all-zero codeword, and $\mathcal{C}''$ contains all the other codewords of $\mathcal{C}$ and the all-zero codeword. Let

$$A_l{}'(\mathcal{C}') \triangleq \begin{cases} \sum_{w=1}^{NR} \left(\frac{w}{NR}\right) A_{w,l} & \text{if } l \in \mathcal{U} \\ 0 & \text{otherwise} \end{cases} .$$

Assume that the communication takes place over an MBIOS channel. Then, under ML decoding, the bit error probability of $\mathcal{C}$, is upper bounded by

$$P_{\mathrm{b}} \leq P_{\mathrm{b}|0}(\mathcal{C}') + P_{\mathrm{b}|0}(\mathcal{C}'')$$

where

$$P_{\mathrm{b}|0}(\mathcal{C}') \;\leq\; 2^{-N\left(E_0(\rho)-\rho\left(R+\frac{\log \bar{\alpha}_\rho(\mathcal{C}')}{N}\right)\right)}, \qquad 0 \leq \rho \leq 1 \tag{3.42}$$

$$\bar{\alpha}_\rho(\mathcal{C}') \triangleq \sum_{l=0}^{N} \left\{ \frac{A'_l(\mathcal{C}')}{2^{-N(1-R)}\binom{N}{l}} \cdot \binom{N}{l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l} \right\}. \tag{3.43}$$

$A(\rho), B(\rho)$ and E_0 are defined in (3.17), (3.18) and (3.8), respectively. As before, an upper bound on the conditional bit error probability for the subcode $\mathcal{C}''$, $P_{\mathrm{b}|0}(\mathcal{C}'')$, can be either a union bound or any other improved bound.

[2]Note that for an ensemble of fully random block codes, all the terms $\frac{A'_l}{B_l}$ are equal to $\frac{1}{2}$; hence, the simplification above does not reduce the tightness of the bound at all when considering this ensemble.

Proof: Starting from the first equality in (3.34), and using the definition for $A(\rho)$, $B(\rho)$ in (3.17) and (3.18) we get

$$\begin{aligned} P_{\mathrm{b}|0} &\leq \left(\frac{A(\rho)+B(\rho)}{2}\right)^N 2^{N\rho} \cdot \left\{\sum_{l=0}^{N} A'_l(\mathcal{C}') \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l}\right\}^{\rho} \\ &= \left(\frac{A(\rho)+B(\rho)}{2}\right)^N 2^{NR\rho} \cdot 2^{N\rho(1-R)} \cdot \left\{\sum_{l=0}^{N} A'_l(\mathcal{C}') \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l}\right\}^{\rho} \\ &= 2^{-N(E_0(\rho)-\rho R)} \cdot \left\{\sum_{l=0}^{N} \frac{A'_l(\mathcal{C}')}{B_l} \binom{n}{l} \left(\frac{B(\rho)}{A(\rho)+B(\rho)}\right)^{N-l} \left(\frac{A(\rho)}{A(\rho)+B(\rho)}\right)^{l}\right\}^{\rho} \end{aligned} \tag{3.44}$$

where

$$B_l \triangleq 2^{-N(1-R)} \binom{N}{l}, \qquad l = 0, \ldots, N$$

designates the distance spectrum of fully random block codes of length N and rate R. Using the definition for $\bar{\alpha}_\rho(\mathcal{C}')$ in (3.43) we get the upper bound (3.42). ■

Evidently, the upper bound (3.42) is tighter than the bit-error version of the SFB in (3.36), because $\bar{\alpha}_\rho(\mathcal{C}')$ which is the expected value of $\frac{A'_l(\mathcal{C}')}{B_l}$ is not larger than $\alpha_{\mathrm{b}}(\mathcal{C}')$ which is the maximal value of $\frac{A'_l(\mathcal{C}')}{B_l}$. We note that the upper bound (3.42) is just the DS2 bound [15], with the un-normalized tilting measure (3.13). This tilting measure is optimal only for the ensemble of fully random block codes, and is sub-optimal for other codes. We refer to the upper bound (3.42) as the *simplified DS2*. From the discussion above, we conclude that the simplified DS2 bound (which is also valid as an upper bound on the conditional *block* error probability if we replace $A'_l(\mathcal{C}')$ in (3.44) by $A_l(\mathcal{C}')$) is advantageous over the MSFB when A'_l (or A_l for the case of block error probability) changes dramatically over the Hamming weight range of interest. This is demonstrates for the block error probability of the ensemble of multiple turbo-Hamming codes where there is no noticeable improvement if we use the simplified DS2 to bound $P_{\mathrm{e}|0}(\mathcal{C}')$ instead of the MSFB, where for the case of bit-error probability we get tighter upper bound when using the simplified DS2 to upper bound $P_{\mathrm{b}|0}(\mathcal{C}')$ rather than the MSFB.

3.4 Expurgation

In this section we consider a possible expurgation of the distance spectrum which yields in general tighter upper bounds on the ML decoding error probability when transmission takes place over a binary-input AWGN (BIAWGN) channel. To this end, we rely on some properties of the Voronoi regions of binary linear block codes, as presented in [1, 2, 3].

Let $\mathcal{C}$ be a binary linear block code of length N and rate R. Without any loss of generality, let us assume that the all-zero codeword, $\mathbf{c}_0$, was transmitted over the BIAWGN channel. For any received vector $\mathbf{y}$, an ML decoder checks whether it falls within the decision region of the all zero vector. This decision region (which is also called the Voronoi region of $\mathbf{c}_0$) is defined as the set $\mathcal{V}_0$ of vectors in $\mathbb{R}^N$ that are closest (in terms of Euclidian distance) to the all-zero codeword, i.e.,

$$\mathcal{V}_0 = \left\{ \mathbf{x} \in \mathbb{R}^N : d(\mathbf{x}, \mathbf{c}_0) \leq d(\mathbf{x}, \mathbf{c}), \quad \forall\, \mathbf{c} \in \mathcal{C} \right\}. \tag{3.45}$$

Not all of the 2^{NR} inequalities in (3.45) are necessarily required to define the Voronoi region. The minimal set of codewords that determine the Voronoi region of $\mathbf{c}_0$, forms the set of Voronoi neighbors of $\mathbf{c}_0$ (to be designated by $\mathcal{N}_0$). So the region (3.45) can be defined by

$$\mathcal{V}_0 = \left\{ \mathbf{x} \in \mathbb{R}^N : d(\mathbf{x}, \mathbf{c}_0) \leq d(\mathbf{x}, \mathbf{c}), \quad \forall\, \mathbf{c} \in \mathcal{N}_0 \right\}. \tag{3.46}$$

It is clear that the block error probability of $\mathcal{C}$ is equal to the conditional block error probability of the expurgated subcode $\mathcal{C}^{\text{ex}}$, assuming the all-zero codeword is transmitted, where $\mathcal{C}^{\text{ex}}$ designates the subcode of $\mathcal{C}$ which contains the all-zero codeword and all its (Voronoi) neighbors. Hence, any upper bound that solely depends on the code distance spectrum of the code can be tightened by replacing the original distance spectrum with *the distance spectrum of the expurgated code*. It should be noted, however, that the argument above cannot be applied to the *bit* error probability. This stems from the fact that while the block error event is solely defined by the Voronoi region of the transmitted codeword, the bit error event also depends on the Hamming weight of the information bits of each decoded codeword; hence, the above expurgation cannot be applied to the analysis of the bit error probability. The distance spectrum of the Voronoi neighbors of an arbitrary codeword of some popular linear block codes (e.g., Hamming, BCH and Golay codes) is given in [1]. A simple

way to find a subcode of $\mathcal{C}$ which contains the subcode $\mathcal{C}^{\text{ex}}$ is given in the following theorem from [2]:

Theorem 3.6 (On the Voronoi Regions of Binary Linear Block Codes [2]) For any binary linear block code $\mathcal{C}$ with rate R and length N

$$\mathcal{N}_0 \supseteq \{\mathbf{c} \in \mathcal{C} : 1 \leq W_{\text{H}}(\mathbf{c}) \leq 2d_{\min} - 1\}$$

and

$$\mathcal{N}_0 \subseteq \{\mathbf{c} \in \mathcal{C} : 1 \leq W_{\text{H}}(\mathbf{c}) \leq N(1-R) + 1\}$$

where $d_{\min}$ is the minimal Hamming weight of the codewords in $\mathcal{C}$.

Note that according to the theorem above, one should expect the expurgation to have maximal impact on the tightness of an upper bound for high rate codes, where most of the codewords can be expurgated. We should also observe that the expurgated codewords have large distances from the all-zero codeword (all the expurgated codewords have a Hamming weight larger than $2d_{\min} - 1$). Thus, the improvement due to the expurgation process is especially substantial at low SNRs. One can use this theorem to achieve an immediate improvement of an arbitrary upper bound by expurgating all the codewords whose Hamming weight is greater than $N(1-R)+1$. We refer to this kind of expurgation as the *trivial* expurgation. The trivial expurgation, though very simple to apply, does not produce satisfactory results in many cases, because in many cases, the portion of the distance spectrum which corresponds to Hamming weights above $N(1-R)+1$ has a negligible effect on the overall bound. In [2], Agrell introduces a method (called *C rule*) in order to determine whether a codeword $\mathbf{c}$ is a zero-neighbor.

C rule: A codeword is a 0-neighbor if and only if it covers[3] no other nonzero codeword.

In [3] , Ashikmin and Barg used this rule to derive explicit formulas for the weight spectrums of zero-neighbors for various codes. This includes the families of Hamming codes and second-order Reed-Muller codes.

In order to upper bound the block error probability using the bounding technique introduced in this chpater, we split the subcode $\mathcal{C}_{\text{ex}}$ into two subcodes, $\mathcal{C}'_{\text{ex}}$

[3]A binary codeword $\mathbf{c}_1$ is said to *cover* another codeword, $\mathbf{c}_2$, if $\mathbf{c}_2$ has zeros in all the positions where $\mathbf{c}_1$ has a zero.

and $\mathcal{C}''_{\rm ex}$, where $\mathcal{C}'_{\rm ex}$ contains all the codewords of $\mathcal{C}_{\rm ex}$ with Hamming weight $l \in \mathcal{U} \subseteq \{1, 2, ..., N(1-R)+1\}$, and $\mathcal{C}''_{\rm ex}$ contains the all-zero codeword and all the other codewords. The following upper bound holds:

$$P_{\rm e}(\mathcal{C}) = P_{\rm e|0}(\mathcal{C}_{\rm ex}) \leq P_{\rm e|0}(\mathcal{C}'_{\rm ex}) + P_{\rm e|0}(\mathcal{C}''_{\rm ex}) \tag{3.47}$$

were $P_{\rm e|0}(\mathcal{C}'_{\rm ex})$ and $P_{\rm e|0}(\mathcal{C}''_{\rm ex})$ are the conditional block error probabilities of the subcodes $\mathcal{C}'_{\rm ex}$ and $\mathcal{C}''_{\rm ex}$, respectively, given that the all-zero codeword was transmitted. We can upper bound $P_{\rm e|0}(\mathcal{C}''_{\rm ex})$ by the union bound or the TSB, and we upper bound $P_{\rm e|0}(\mathcal{C}'_{\rm ex})$ by the MSFB (3.16). The partitioning of the subcode $\mathcal{C}_{\rm ex}$ into two subcodes $\mathcal{C}'_{\rm ex}$ and $\mathcal{C}''_{\rm ex}$ is done following the adaptive algorithm introduced in Section 3.3.

3.5 Applications

This section demonstrates some numerical results of the improved upper bounds on the ML decoding error probability of linear block codes. We apply the bounds introduced in Sections 3.3 and 3.4 to various ensembles of parallel and serially concatenated codes. Throughout this section, it is assumed that the encoded bits are BPSK modulated, transmitted over an AWGN channel, and coherently detected. The effect of an expurgation of the distance spectrum on the tightness of some upper bounds on the decoding error probability is exemplified as well.

For the binary-input additive white Gaussian noise (BIAWGN) channel with BPSK modulation, the conditional probability density function (*pdf*) for a single letter input is:

$$\begin{aligned} p(y|0) &= \frac{1}{\sqrt{\pi N_0}} \exp\left\{ -\left(y + \sqrt{E_{\rm s}}\right)^2 / N_0 \right\}, \\ p(y|1) &= \frac{1}{\sqrt{\pi N_0}} \exp\left\{ -\left(y - \sqrt{E_{\rm s}}\right)^2 / N_0 \right\} \end{aligned} \tag{3.48}$$

where $E_{\rm s}$ designates the energy of the symbol, and $\frac{N_0}{2}$ is the two-sided spectral power density of the channel. In order to calculate the SFB on $P_{\rm e|0}(\mathcal{C}')$, we first calculate the terms $A(\rho)$ and $B(\rho)$, as defined in (3.17) and (3.18), respectively. Clearly, for a

continuous-output channel, the sums in (3.17) and (3.18) are replaced by integrals.

$$\begin{aligned}
B(\rho) &= \int_{-\infty}^{\infty} p(y|0)^{\frac{2}{\rho+1}} \left[\frac{1}{2}p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2}p(y|1)^{\frac{1}{1+\rho}}\right]^{\rho-1} dy \\
&= \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{\pi N_0}}\right)^{\frac{2}{\rho+1}} e^{-\frac{2\left(y+\sqrt{E_\mathrm{s}}\right)^2}{N_0(1+\rho)}} \left(\frac{1}{\sqrt{\pi N_0}}\right)^{\frac{\rho-1}{\rho+1}} \left[\frac{1}{2}e^{-\frac{\left(y+\sqrt{E_\mathrm{s}}\right)^2}{N_0(1+\rho)}} + \frac{1}{2}e^{-\frac{\left(y-\sqrt{E_\mathrm{s}}\right)^2}{N_0(1+\rho)}}\right]^{\rho-1} dy \\
&= \exp\left(-\frac{E_\mathrm{s}}{N_0}\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} e^{-\frac{y^2}{N_0}} \cdot e^{-\frac{4y\sqrt{E_\mathrm{s}}}{N_0(\rho+1)}} \left[\frac{1}{2}e^{\frac{2y\sqrt{E_\mathrm{s}}}{N_0(1+\rho)}} + \frac{1}{2}e^{-\frac{2y\sqrt{E_\mathrm{s}}}{N_0(1+\rho)}}\right]^{\rho-1} dy \\
&= \exp\left(-\frac{E_\mathrm{s}}{N_0}\right) \mathbb{E}\left[e^{-\frac{2X\sqrt{2E_\mathrm{s}/N_0}}{\rho+1}} \cosh^{\rho-1}\left(\frac{\sqrt{2E_\mathrm{s}/N_0}X}{1+\rho}\right)\right] \qquad (3.49)
\end{aligned}$$

where $\mathbb{E}$ denotes the statistical expectation, and $X \sim \mathrm{N}(0,1)$. We also obtain that

$$\begin{aligned}
A(\rho) &= \int_{-\infty}^{\infty} [p(y|0)p(y|1)]^{\frac{1}{1+\rho}} \left[\frac{1}{2}p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2}p(y|1)^{\frac{1}{1+\rho}}\right]^{\rho-1} dy \\
&= \exp\left(-\frac{E_\mathrm{s}}{N_0}\right) \mathbb{E}\left[\cosh^{\rho-1}\left(\frac{\sqrt{2E_\mathrm{s}/N_0}X}{1+\rho}\right)\right] \qquad (3.50)
\end{aligned}$$

and

$$A(\rho) + B(\rho) = 2\exp\left(-\frac{E_\mathrm{s}}{N_0}\right) \mathbb{E}\left[\cosh^{1+\rho}\left(\frac{\sqrt{2E_\mathrm{s}/N_0}X}{1+\rho}\right)\right] \qquad (3.51)$$

Plugging (3.49) – (3.51) into (3.16), and (3.41) and minimizing over the interval $0 \leq \rho \leq 1$ will give us the desired bounds for $P_{\mathrm{e}|0}(\mathcal{C}')$ and $P_{\mathrm{b}|0}(\mathcal{C}')$, respectively.

3.5.1 Ensemble of Serially Concatenated Codes

The scheme in Fig. 3.2 depicts the encoder of an ensemble of serially concatenated codes where the outer code is a (127, 99, 29) Reed-Solomon (RS) code, and the inner code is chosen uniformly at random from the ensemble of (8, 7) binary linear block codes. Thus, the inner code extends every symbol of 7 bits from the Galois field $\mathrm{GF}(2^7)$ to a sequence of 8 bits. The decoding is assumed to be performed in two stages: the inner (8, 7) binary linear block code is soft-decision ML decoded, and then a hard decision ML decoding is used for the outer (129, 99, 29) RS code. Due to the hard-decision ML decoding of the (127, 99, 29) RS code, its decoder can correct up to $t = \lfloor \frac{d_{\min}-1}{2} \rfloor = 14$ erroneous symbols. Hence, an upper bound on the average

block error probability of the considered serially concatenated ensemble is given by

$$P_{\mathrm{e}} \leq \sum_{i=t+1}^{127} \binom{127}{i} p_{\mathrm{s}}^{i} (1-p_{\mathrm{s}})^{127-i} \tag{3.52}$$

where p_{s} is the average symbol error probability of the inner code under soft-decision ML decoding. The symbol error probability p_{s} of the inner code is either upper bounded by the ubiquitous union bound or the TSB, and this upper bound is substituted in the RHS of (3.52). Since the rate of the inner code is rather high (it is equal to $\frac{7}{8}$ bits per channel use), an expurgation of the distance spectrum seems to be attractive in order to tighten the upper bound on the overall performance of the concatenated ensemble. Ashikmin and Barg [3] show that the average expurgated distance spectrum of the ensemble of random linear block codes of length N and dimension K is given by

$$E[A_l] = \begin{cases} \binom{N}{l} 2^{-(N-K)} \prod_{i=0}^{l-2} \left(1 - 2^{-(N-K-i)}\right) & l = 0, 1, \ldots, N-K+1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.53}$$

We rely on the expurgated distance spectrum in (3.53) in order to get a tighter version of the union bound or the TSB on the symbol error probability p_{s} of the inner code (where $N = 8$ and $K = 7$)[4]. The expurgated union bound in Fig. 3.3 provides a gain of 0.1 dB over the union bound or TSB at block error probability of 10^{-4}, and the improvement in the tightness of the bound due to the distance spectrum expurgation is especially prominent at low values of SNR. Clearly, we take 1 as the trivial bound on p_{s} (as otherwise, for low values of SNR, the union bound on p_{s} may exceed 1, which gives in turn a useless upper bound on the decoding error probability of the ensemble).

3.5.2 Turbo-Hamming Codes

Let us consider an ensemble of uniformly interleaved parallel concatenated turbo-Hamming codes. The encoder consists of two identical $(2^m - 1, 2^m - m - 1)$ Hamming codes as component codes, and a uniform interleaver operating on the $2^m - m - 1$

[4]In order to calculate the average distance spectrum of the ensemble of random binary linear block codes, see Appendix D.

information bits. The comparison here refers to the case where $m = 10$, so the two component codes are (1023, 1013) Hamming codes, and the overall rate of the ensemble is $R = \frac{2^m-m-1}{2^m+m-1} = 0.9806$ bits per channel use. The value of the energy per bit to one-sided spectral noise density ($\frac{E_\mathrm{b}}{N_0}$) which corresponds to this coding tare is 5.34 dB, assuming that communication takes place over a binary-input AWGN channel. In order to obtain performance bounds for the ensemble of uniformly interleaved turbo-Hamming codes, we rely on an algorithm for the calculation of the average input-output weight enumerator function (IOWEF) of this ensemble, as provided in [32, Section 5.2]. As noted in [32], the average distance spectrum of this ensemble is very close to the binomial distribution for a rather large range of Hamming weights (see Fig. **??**(a)). Hence, one can expect that the upper bound introduced in Theorem 3.1 provides a tight bounding technique on the average block error probability of this ensemble. For this coding scheme, we note that regarding P_e, there is no substantial improvement in the tightness of the overall upper bound if we upper bound $P_{\mathrm{e}|0}(\mathcal{C}'')$ by the TSB instead of the simple union bound (see Fig. 3.5). Among the bounds introduced in Section 3.3, the upper bound which combines the TSB and the MSFB is the tightest bound, especially for the low SNR range (see Fig. 3.5); referring to the bound in Theorem 3.1, the partitioning of codes in the considered ensemble relies on Algorithm 1 (see Section 3.3). In Fig. 3.6, we provide a comparison between various upper bound on the *bit* error probability of this turbo-like ensemble. The tightest bound for the bit error analysis is the one provided in Theorem 3.5, combining the simplified DS2 bound with the union bound. It is shown in Fig. 3.6 that the simplified DS2 provides gains of 0.16 dB and 0.05 dB over the MSFB at bit error probabilities of 10^{-1} and 10^{-2}, respectively. The simplified DS2 also provides gain of 0.08 dB over the TSB at bit error probability of 10^{-1}. Unfortunately, a trivial expurgation of the average distance spectrum of uniformly interleaved turbo codes with two identical $(2^m - 1, 2^m - m - 1)$ Hamming codes as components (i.e., by nullifying the average distance spectrum at Hamming weights above $2m + 1$) has no impact on tightening the performance bounds of this ensemble.

3.5.3 Multiple Turbo-Hamming Codes

Multiple turbo codes are known to yield better performance, and hence, it is interesting to apply the new bounding techniques in Section 3.3 to these ensembles. The encoder of a multiple turbo-Hamming code is depicted in Fig. 3.7.

Consider the ensemble of uniformly and independently interleaved multiple-turbo codes, where the components codes are identical systematic binary linear block codes of length N. Let S_{w,h_i} denote the number of codewords of the i^{th} component code with weight of the systematic bits equal to w and the weight of the parity bits equal to h_i. The average number of codewords of the ensemble of multiple-turbo codes, with systematic-bits weight of w and overall weight l is given by

$$A_{w,l} = \sum_{\substack{h_1, h_2, h_3 \quad \text{s.t.} \\ w + h_1 + h_2 + h_3 = l}} \frac{S_{w,h_1} S_{w,h_2} S_{w,h_3}}{\binom{N}{w}^2}. \tag{3.54}$$

From (3.54) and the algorithm to calculate the input-output weight enumerators of Hamming codes (see [32, Appendix A]), it is possible to verify that the average distance spectrum of the ensemble of multiple turbo-Hamming codes with two independent uniform interleavers is very close to the binomial distribution for a relatively large range of Hamming weights (similarly to the plot in Fig. **??**(a)). Hence, the improved bounds provided in Section 3.3 are expected to yield good upper bounds on the decoding error probability. The comparison here refers to the case of $m = 10$, so the three component codes are (1023, 1013) Hamming codes. The overall rate of the ensemble is $\frac{2^m - m - 1}{2^m + 2m - 1} = 0.9712$ bits per channel use, and the channel capacity for this coding rate corresponds to $\frac{E_{\text{b}}}{N_0} = 5$ dB. All the improved bounds that are evaluated here, incorporate the union bound as an upper bound on $P_{\text{e}}(\mathcal{C}'')$ (or $P_{\text{b}}(\mathcal{C}'')$ for bit error probabilities). The numerical results of various upper bounds are shown in Fig. 3.8 for the block and bit error probabilities. As expected, the improvements that were obtained by the improved bounds (Theorems 3.1–3.5) are more pronounced here than for the ensemble of turbo-Hamming code. For example, at bit error rate of 10^{-1}, the simplified DS2 bound yields a gain of 0.12 dB over the TSB. A modest improvement of 0.05 dB was obtained at bit error rate of 10^{-2}.

3.5.4 Random Turbo-Block Codes with Systematic Binary Linear Block Codes as Components

Finally, we evaluate improved upper bound for the ensemble of uniformly interleaved parallel concatenated (turbo) codes, having two identical component codes chosen uniformly at random and independently from the ensemble of systematic binary linear block codes. We assume that the parameters of the overall code are (N, K), so the parameters of its component codes are $(\frac{N+K}{2}, K)$. In addition, the length of the uniform interleaver is K.

According to the analysis in [40], the input-output weight enumeration of the considered ensemble is given by

$$\begin{aligned} S(W,Z) &= \sum_{w,j} S_{w,j} W^w Z^j \\ &= 1 + \sum_{w=1}^{K} \left\{ W^w \left[2^{-(N-K)} \left(\binom{K}{w} - 1 \right) \sum_{j=0}^{N-K} \binom{N-K}{j} Z^j + 2^{-\frac{N-K}{2}} \sum_{j=0}^{\frac{N-K}{2}} \binom{\frac{N-K}{2}}{j} Z^{2j} \right] \right\} \end{aligned}$$

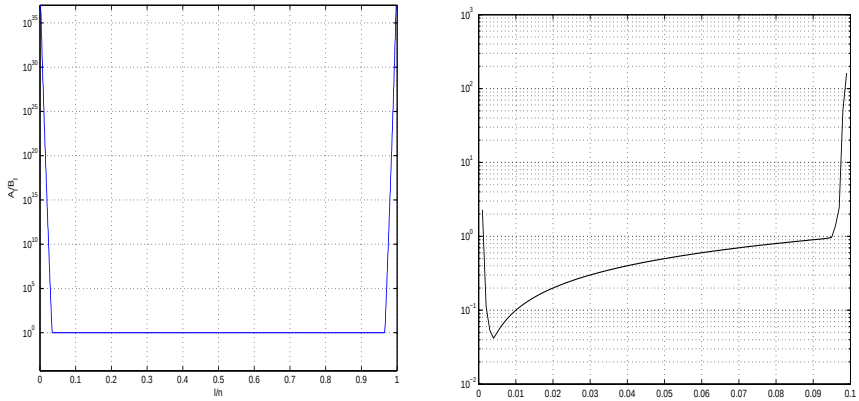
where $S_{w,j}$ denotes the number of codewords whose information sub-words have Hamming weight of w and the parity sub-word has Hamming weight j. We apply the improved bounds introduced in Section 3.3 to this ensemble where the parameters are set to $(N, K) = (1144, 1000)$ (hence, the rate of the parallel concatenated ensemble is $R = 0.8741$ bits per channel use). The plots of various upper bounds on the block and bit error probabilities are shown in Fig. 3.9. The improved bounds yield the best reported upper bound on the block and bit error probabilities. For the block error probability, the upper bound which combines the MSFB with the union bound is the tightest bound; it achieve a gain of 0.1 dB over the TSB, referring to a block error probability of 10^{-4}. A similar gain of 0.11 dB is obtained for the bit error probability, referring to a BER of 10^{-4}, referring to the bound which combined the union bound with the simplified DS2 bound (see Theorem 3.5).

3.6 Conclusions

We derive in this chpater tightened versions of the Shulman and Feder bound. The new bounds apply to the bit and block error probabilities of binary linear block codes

under ML decoding. The effectiveness of these bounds is exemplified for various ensembles of turbo-like codes over the AWGN channel. An expurgation of the distance spectrum of binary linear block codes further tightens in some cases the resulting upper bounds.

Figures



(a) (b)

Figure 3.1: Plots of $\frac{A_l}{B_l}$ and $\frac{A'_l}{B_l}$ as a function of the normalized Hamming weight $\left(\frac{l}{N}\right)$, on a logarithmic scale. The plots refer to ensembles of random turbo-block codes with two identical systematic binary linear block codes as components; (a) A plot of $\frac{A_l}{B_l}$ with $N = 1000$ and $R = 0.72$ bits/Symbol, referring to the analysis of the block error probability, (b) A plot of $\frac{A'_l}{B_l}$ with $N = 100$ and $R = 0.72$ bits/Symbol, referring to the analysis of the bit error probability.

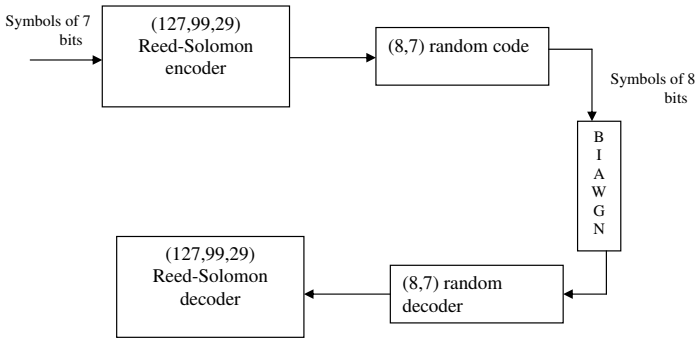


Figure 3.2: A scheme for an ensemble of serially concatenated codes where the outer code is a (127, 99, 29) Reed-Solomon (RS) code, and the inner code is chosen uniformly at random from the ensemble of (8,7) binary linear block codes.

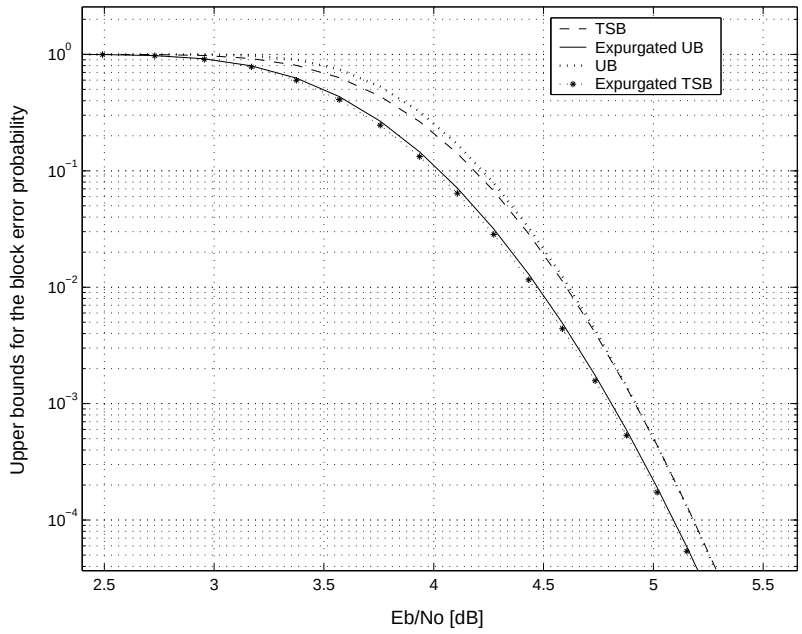


Figure 3.3: Various upper bounds on the block error probability of the ensemble of serially concatenated codes depicted in Fig. 3.2. The compared bounds are the tangential-sphere bound (TSB) and the union bound with and without expurgation of the distance spectrum; this expurgation refers to the ensemble of inner codes, chosen uniformly at random from the ensemble of (8,7) binary linear block codes.

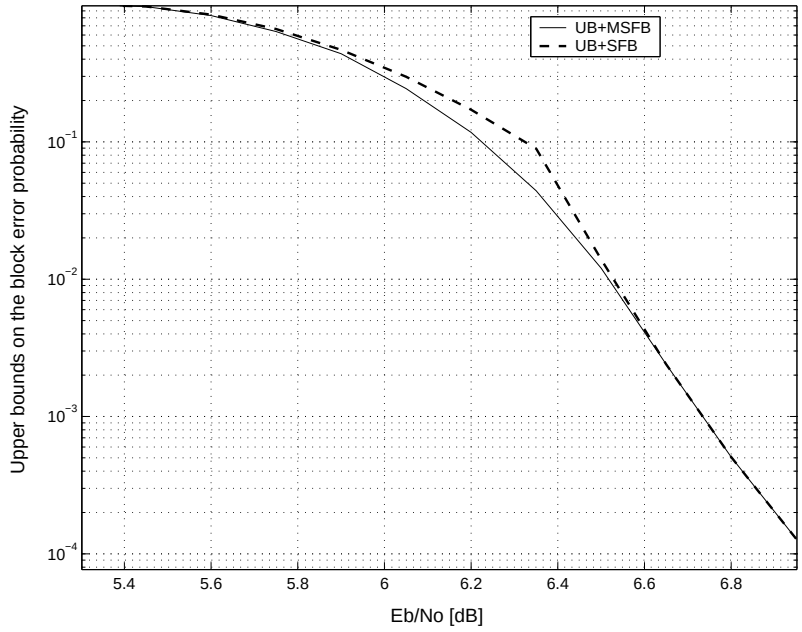


Figure 3.4: A comparison between the upper bound which combines the UB with the SFB bound in it original form (Eq. (3.6)) and the upper bound which combines the UB with the MSFB bound in (3.16). The comparison refers to the ensemble of uniformly interleaved turbo-Hamming codes where the two component codes are (1023, 1013) Hamming codes. The overall rate of the code is 0.973 bits per channel use.

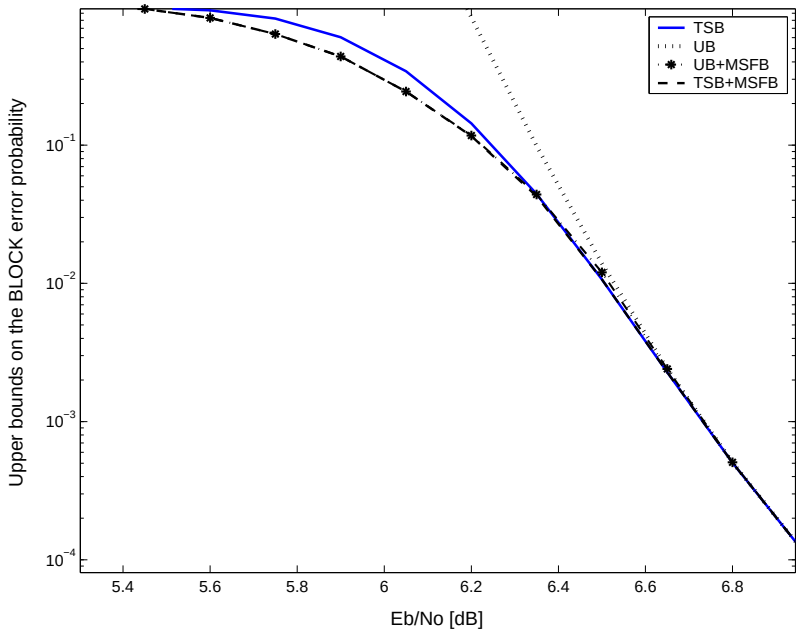


Figure 3.5: Comparison between various upper bounds on the ML decoding block error probability where the comparison refers to the ensemble of uniformly interleaved turbo-Hamming codes whose two component codes are (1023, 1013) Hamming codes. The compared bounds are the union bound (UB), the tangential-sphere bound (TSB), and two instances of the improved upper bound from Theorem 3.1: the UB+MSFB combines the MSFB with the union bound, and the TSB+MSFB is the upper bound which combines the MSFB with the tangential-sphere bound.

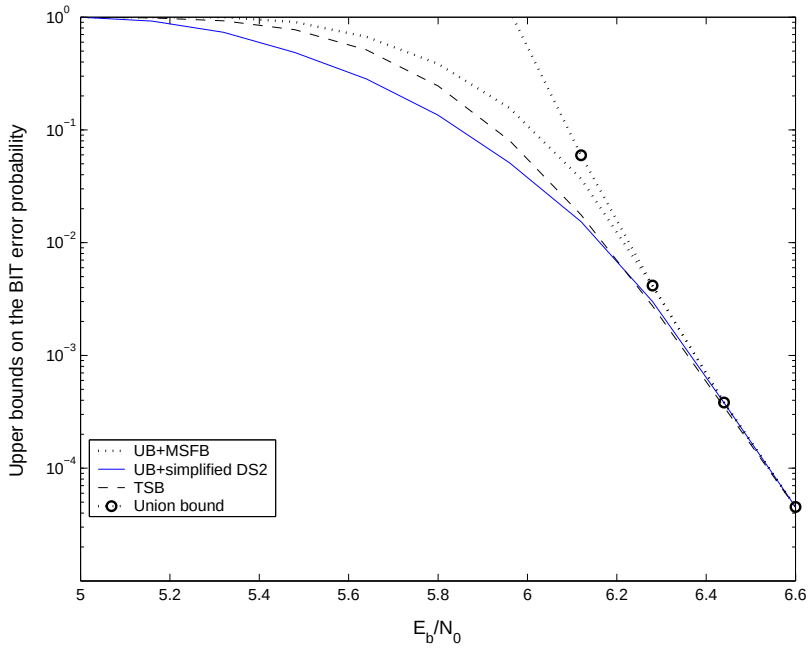


Figure 3.6: Comparison between various upper bounds on the ML decoding bit error probability of the ensemble of (1033,1013) uniformly interleaved turbo-Hamming code. The compared bounds are the union bound (UB), the tangential-sphere bound (TSB), the upper bound from Theorem 3.4 which combines the union bound with the MSFB (UB+MSFB), and the upper bound from Theorem 3.5 which combines the union bound with the simplified DS2 bound (UB+simplified DS2).

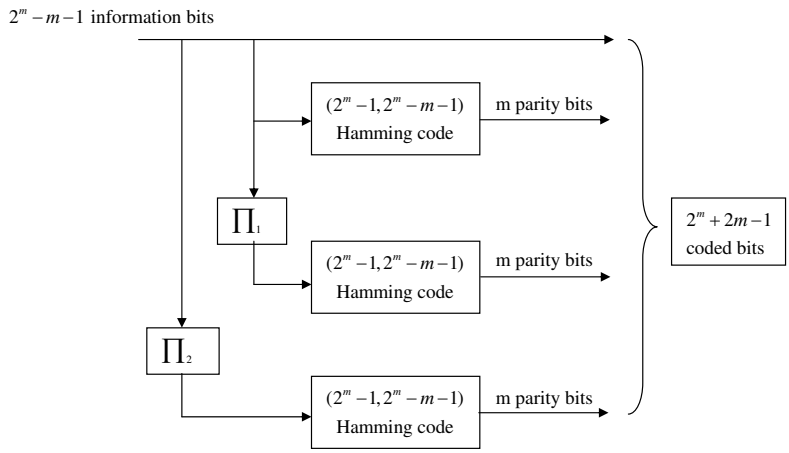


Figure 3.7: A multiple turbo-Hamming encoder. The encoder consists of parallel concatenated Hamming codes with two uniform, statistically independent interleavers. The code length is $2^m + 2m - 1$ and the code rate is $R = \frac{2^m - m - 1}{2^m + 2m - 1}$ bits per channel use.

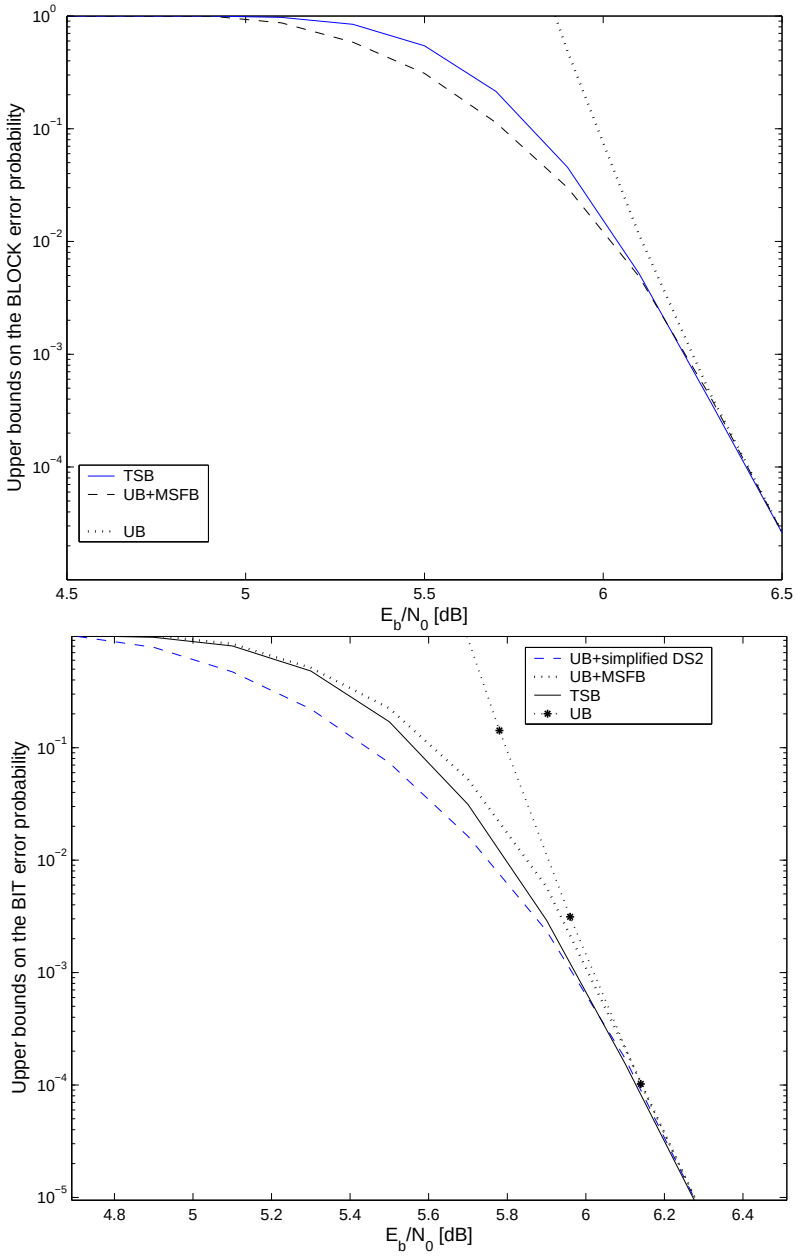


Figure 3.8: Comparison between various upper bounds on the ML decoding error probability, referring to the ensemble of uniformly interleaved multiple turbo-Hamming codes where the three component codes are (1023, 1013) Hamming codes (see Fig. 3.7). The upper plot refers to upper bounds on the block error probability, and the compared bounds are the union bound (UB), the tangential-sphere bound (TSB), and the upper bound of Theorem 3.1 which combines the union bound with the MSFB (UB+modified SFB). The lower plot refers to upper bounds on the bit error probability, and the compared bounds are the union bound (UB), the tangential-sphere bound (TSB), the upper bound of Theorem 3.4 which combines the union bound with the MSFB, and the upper bound of Theorem 3.5 which combines the union bound with the simplified DS2 bound (UB+simplified DS2).

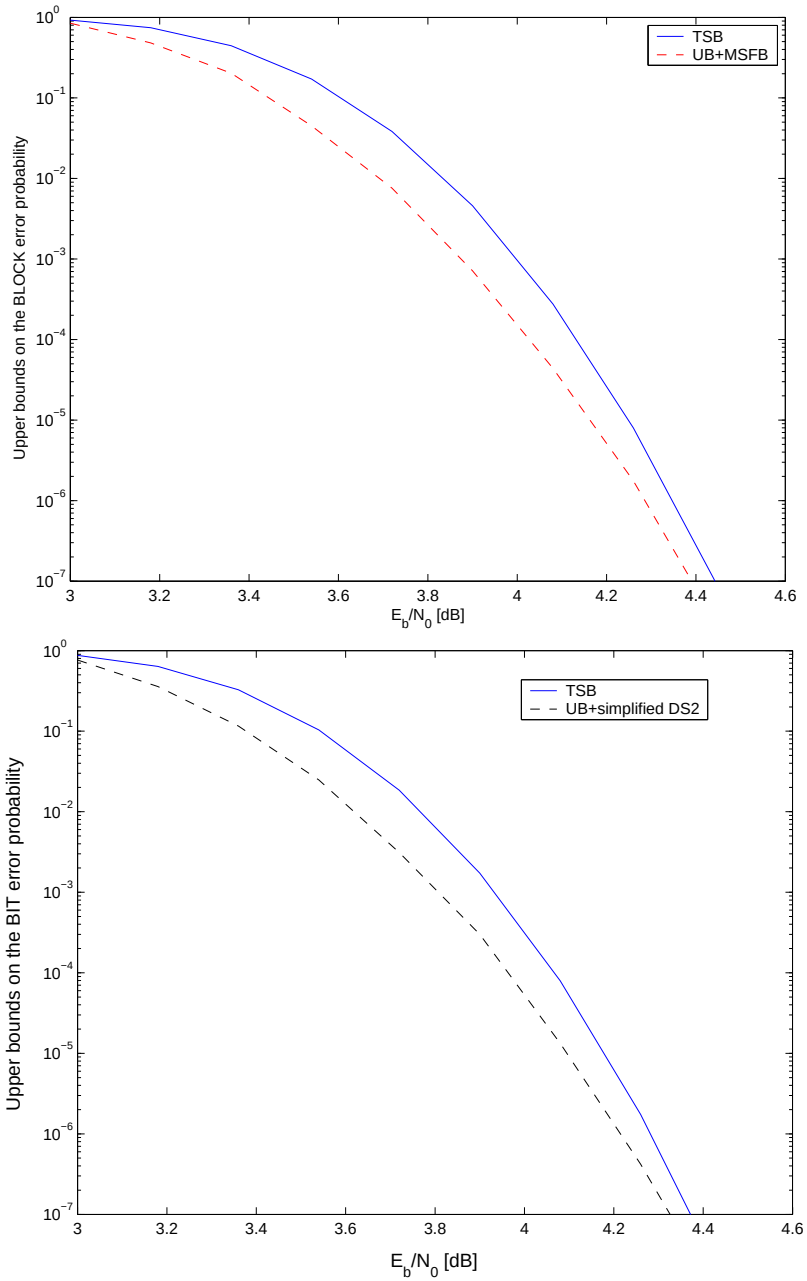


Figure 3.9: Comparison between upper bounds on the block and bit error probabilities for an ensemble of uniformly interleaved turbo codes whose two component codes are chosen uniformly at random from the ensemble of (1072, 1000) binary systematic linear block codes; its overall code rate is 0.8741 bits per channel use. The compared bounds under ML decoding are the tangential-sphere bound (TSB), and the bounds in Theorems 3.1 and 3.5. The upper and lower plots provide upper bounds on the block and bit error probabilities, respectively.

Chapter 4

Summary and Conclusion

4.1 Contribution of the Thesis

The tangential-sphere bound (TSB) of Poltyrev [31] often happens to be the tightest upper bound on the ML decoding error probability of block codes whose transmission takes place over a binary-input AWGN channel. However, in the random coding setting, it fails to reproduce the random coding exponent [19] while the second version of the Duman and Salehi (DS2) bound does [15, 35]. In the first part of this work, we consider some recently introduced performance bounds which suggest an improvement over the TSB. These bounds rely solely on the distance spectrum of the code (or their input-output weight enumerators for the analysis of the bit error probability). In Chapter 2.4, we study the error exponents of these recently introduced bounding techniques. This forms a direct continuation to the derivation of these bounds by Yousefi et al. [45, 46, 47] who also exemplified the superiority of their recently introduced bounds over the TSB for short binary linear block codes. We conclude that all the aforementioned upper bounds posses the same error exponent as the TSB. Moreover, the error exponents of the TSB versions for the bit error probability, as provided in [33, 51], coincide and are equal to the error exponent of the TSB for the block error probability. The explicit expression of this error exponent is given in Theorem 2.4, and is therefore identical to the error exponent of the TSB as was first derived by Poltyrev [31] for random codes, and later simplified by Divsalar and adapted to general ensembles of binary linear block codes [11]. Since the gap between the error exponent of the TSB and the random coding error exponent of Gallager [19]

(see Fig. 2.3 in p. 33) becomes larger as the code rate is increased, tightened upper bounds are especially needed for high-rate linear codes.

In Chapter 3 we derive tightened upper bounds on the decoding error probability of linear block codes, under ML decoding. The bounds derived in the second part of the thesis form an improvement over the Shulman and Feder bound [37], and as particular cases of the generalized of Duman and Salehi bound [15, 35], are more simple for calculation. They reproduce the random coding error exponent as a by product of their superiority over the Shulman and Feder bound. These bounds on the block and bit error probabilities depend respectively on the distance spectrum and input-output weight enumeration function of the codes, so one can easily apply them to various codes and ensembles. The effectiveness of these bounds (which are valid for arbitrary memoryless, binary-input and output-symmetric channels) is exemplified for various ensembles of turbo-like codes when transmission takes place over the binary-input AWGN channel. For some ensembles of turbo-like codes (especially, ensembles of high-rate codes), they provide a better bounding technique than the TSB. in some cases, an expurgation of the distance spectrum of binary linear block codes further tightens the resulting upper bounds.

4.2 Topics for Further Research

In the following, we propose some topics for further research:

- In [46, 47], Yousefi et. al. derive improved versions of tangential-sphere bound, by using the Hunter bound. One may obtain various upper bounds from the bounding technique of [46] (i.e., by applying the Hunter bound on the probability of a union involved in the TSB (see Section 2.2.3)), by means of more delicate treatment of the correlation coefficients. In Chapter 2, we show that the ITSB and AHP upper bounds [46, 47] have the same error exponent as the TSB. In fact, by introducing Lemma 2.3 we prove a stronger argument. Namely, we show that as long as the complementary events $E^c_{0\to\lambda_i}$ correspond to codewords with the same Hamming weight, no improvement is achieved over the error exponent of the TSB. This implies that the potential of using a Bonferroni-type inequality (of order 2 and more) may not have been fully exploited. Hence, a possible research may be a search for an upper bound which is based on a

Bonferroni-type inequality [17], and whose error exponent is at least as large as the error exponent of the TSB; this bound should also depend solely on the distance spectrum of the code.

- In [22], Herzberg and Poltyrev adapt the TSB to upper bound the decoding error probability of M-ary phase-shift keying (PSK) block coded modulation, under *coherent* ML decoding. However, in practical communication system, the detection is rarely coherent, due to oscillator instability in the receiver. The oscillator instability due to noise, which manifest itself as phase noise (PHN), is one of the primary factors that limit the achievable performance in many communication systems [9], [29]. Hence, finding an upper bound on the decoding error probability of M-ary PSK block coded modulation under *none-coherent* ML decoding is of high importance. The PHN is generally modelled as a wide-sense stationary Gaussian process or a Weiner process ([29], [10]). For both types, the PHN obviously does not change the energy of the received vectors (given the value of the PHN). Hence, the receiving signals still have constant energy, and one may apply the TSB as an upper bound on the aforementioned error probability.

- The constant envelope of continuous phase modulations (CPM) and their excellent spectral properties make them attractive in many digital transmission systems [41]. In [8], Brutel and Boutros consider serial concatenation of outer convolutional code and a continuous phase modulation as an inner code separated by a random interleaver. They asses the performance of this ensemble via union bounds. We propose to apply the tightened upper bounds derived in Chapter 3 on the decoding error probability of the above ensemble, as well as other continuous phase modulated turbo-codes. The proposed bounding techniques are expected to provide tighter upper bounds than those introduced in [8].

- In [5], Bennatan and Burshtein generalized the Shulman and Feder bound to arbitrary discreet-memoryless channels (DMC). They also combine the SFB with the union-Bhattacharrya bound for further tightening the resulting upper bound. A possible research in this direction is the generalization of the upper

bounds from Chapter 3 to an arbitrary DMC channel. Likewise, one can apply the generalized versions of the bounds from Chapter 3 on the ensemble of modulo-q quantized coset LDPC codes, and compare the results with the upper bound used by Bennatan and Burshtein in [5].

Appendix A

The exponent of $\psi(\mathcal{C})$

In the following, the exponential behavior of the RHS of (2.48) is obtained by using the Chernoff bounding technique for $\psi(\mathcal{C})$.

Note that the geometrical region of the TSB corresponds to a double sided circular cone. For the derivation of the bound for the single cone, we have put the further restriction $z_1 \leq \sqrt{NE_{\mathrm{s}}}$, but since $z_1 \sim N(0, \frac{N_0}{2})$, then this boundary effect does not have any implication on the exponential behavior of the function $\psi(\mathcal{C})$ for large values of N (as also noted in [11, p. 23]). To simplify the analysis, we therefore do not take into consideration of this boundary effect for large values of N. Let $\widetilde{\psi}(\mathcal{C})$ designate the function which is obtained by removing the event $z_1 \leq \sqrt{NE_{\mathrm{s}}}$ from the expression for $\psi(\mathcal{C})$ (see the RHS of (2.48)).

Let us designate the normalized Gaussian noise vector by ν, i.e., $(\nu_1, \ldots, \nu_N) = \sqrt{\frac{2}{N_0}}(z_1, \ldots, z_N)$, and define $\eta \triangleq \tan^2\theta$. The Gaussian random vector has N orthogonal components which are therefore statistically independent. From (2.4) and (2.42), the following equalities hold for BPSK modulated signals:

$$\begin{aligned} r &= \sqrt{2Nc\eta} \\ r_{\nu_1} &= \sqrt{\eta}\left(\sqrt{2Nc} - \nu_1\right) \\ \beta_h(\nu_1) &= \left(\sqrt{2Nc} - \nu_1\right)\sqrt{\frac{h}{N-h}} \\ l_{w,h}(\nu_1, \nu_2) &= \frac{\beta_w(\nu_1) - \rho_{w,h}\,\nu_2}{\sqrt{1-\rho_{w,h}^2}}. \end{aligned} \tag{A.1}$$

Hence, we obtain from (2.48) and the above discussion

$$\widetilde{\psi}(\mathcal{C}) = \min_{w} \Bigg\{ \Pr\left(\sum_{i=2}^{N} \nu_i^2 \le r_{\nu_1}^2, \nu_2 \ge \beta_w(\nu_1)\right) + \sum_{h=1}^{N} A_h \Pr\left(\sum_{i=2}^{N} \nu_i^2 \le r_{\nu_1}^2, \nu_2 \ge \beta_h(\nu_1), \nu_3 \ge -l_{w,h}(\nu_1,\nu_2)\right) + \Pr\left(\sum_{i=2}^{N} \nu_i^2 \ge r_{\nu_1}^2\right) \Bigg\}. \tag{A.2}$$

At this point, we upper bound the RHS of (A.2) by the Chernoff bounds, namely, for three random variables V, W and Z

$$\Pr(V \ge 0) \le \mathbb{E}\left[e^{pV}\right], \quad p \ge 0 \tag{A.3}$$

$$\Pr(W \le 0, V \ge 0) \le \mathbb{E}\left[e^{qW+uV}\right], \quad q \le 0, u \ge 0 \tag{A.4}$$

$$\Pr(W \le 0, V \ge 0, Z \ge 0) \le \mathbb{E}\left[e^{tW+sV+kZ}\right], \quad t \le 0, s \ge 0, k \ge 0. \tag{A.5}$$

The Chernoff versions of the first and last terms in the RHS of (A.2) are introduced in [11, Eqs.(134)–(137)], and are given by

$$\Pr\left(\sum_{i=2}^{N} \nu_i^2 \ge r_{\nu_1}^2\right) \le \sqrt{\frac{1-2p}{1+2p\eta}} e^{-nE_1(c,p,\eta)}, \quad p \ge 0 \tag{A.6}$$

$$\Pr\left(\sum_{i=2}^{N} \nu_i^2 \le r_{\nu_1}^2, \nu_2 \ge \beta_w(\nu_1)\right) \le \sqrt{\frac{1-2q}{1+2q\eta}} e^{-nE_2(c,q,\frac{w}{N},\eta)}, \quad -\frac{1}{2\eta} \le q \le 0 \tag{A.7}$$

where

$$E_1(c,p,\eta) = \frac{2p\eta c}{1+2p\eta} + \frac{1}{2}\ln(1-2p). \tag{A.8}$$

and

$$E_2(c,q,\delta,\eta) = c\left(\frac{2q\eta + (1-2q)\sqrt{\frac{\delta}{1-\delta}}}{1+2q\eta+(1-2q)\sqrt{\frac{\delta}{1-\delta}}}\right) + \frac{1}{2}\ln(1-2q). \tag{A.9}$$

Next, by invoking the Chernoff bound (A.5), we get an exponential upper bound on the second term in the RHS of (2.48). Using the notation

$$\zeta_{w,h} \triangleq \sqrt{\frac{w(N-h)}{h(N-w)}} \tag{A.10}$$

we get (see Appendix B for details)

$$A_h \Pr\left(\sum_{i=2}^{N}\nu_i^2 \le r_{\nu_1}^2, \nu_2 \ge \beta_h(\nu_1), \nu_3 \ge -l_{w,h}(\nu_1,\nu_2)\right)$$
$$\le \sqrt{\frac{1-2t}{1+2t\eta}}e^{-g(c,t,k,s,\eta,h,N)}, \qquad -\frac{1}{2\eta} \le t \le 0,\ k \ge 0,\ s \ge 0 \tag{A.11}$$

where

$$g(c,t,k,s,\eta,h,N) \triangleq \frac{4t\eta nc + 2\sqrt{2Nc}\left(s - \frac{k\zeta_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)\Delta_h - \Delta_h^2\left(s - \frac{k\zeta_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)^2}{2(1+2t\eta)}$$
$$-\frac{\left(s - \frac{k\rho_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)^2}{2(1-2t)} - \frac{k^2}{2(1-2t)} + \frac{N}{2}\ln(1-2t) - nr(\frac{h}{N}) \tag{A.12}$$

and

$$\Delta_h \triangleq \sqrt{\frac{h}{N-h}}.$$

The next step is to find optimal values for K and s in order to maximize the function g. If $K^* = 0$ then the exponent of $\psi(\mathcal{C})$ is identical to that of the TSB. In order to find the optimal $K \ge 0$ and $s \ge 0$ which maximize g, we consider the aforementioned probabilities by discussing separately the three cases where $h < w,\ h > w$ and $h = w$.

Case 1: $h = w$. In this case $\zeta_{w,h} = \zeta_{w,w} = 1$, and we get

$$A_w \Pr\left(\sum_{i=2}^{N}\nu_i^2 \le r_{\nu_1}^2, \nu_2 \ge \beta_w(\nu_1), \nu_3 \ge -l_{w,w}(\nu_1,\nu_2)\right) \le \sqrt{\frac{1-2t}{1+2t\eta}}e^{-g(c,t,k,s,\eta,w,N)}$$
$$-\frac{1}{2\eta} \le t \le 0,\ k \ge 0,\ s \ge 0 \tag{A.13}$$

where

$$g(c,t,k,s,\eta,w,N) = \frac{4t\eta nc + 2\sqrt{2Nc}\left(s - \frac{k}{\sqrt{1-\rho_{w,w}^2}}\right)\Delta_w - \Delta_w^2\left(s - \frac{k}{\sqrt{1-\rho_{w,w}^2}}\right)^2}{2(1+2t\eta)}$$
$$-\frac{\left(s - \frac{k\rho_{w,w}}{\sqrt{1-\rho_{w,w}^2}}\right)^2}{2(1-2t)} - \frac{k^2}{2(1-2t)} + \frac{N}{2}\ln(1-2t) - \ln(A_w). \tag{A.14}$$

Let us define the parameters

$$\xi = s - \frac{k}{\sqrt{1-\rho_{w,w}^2}} \tag{A.15}$$

$$\tau = s - \frac{k\rho_{w,w}}{\sqrt{1-\rho_{w,w}^2}}. \tag{A.16}$$

From (A.15) and (A.16), we get

$$k = -(\xi - \tau)\alpha \tag{A.17}$$

where

$$\alpha \triangleq \sqrt{\frac{1+\rho_{w,w}}{1-\rho_{w,w}}}. \tag{A.18}$$

Hence, the Chernoff bounding technique gives

$$\Pr\left(\sum_{i=2}^{N}\nu_i^2 \le r_{\nu_1}^2, \nu_2 \ge \beta_w(\nu_1), \nu_3 \ge -l_{w,w}(\nu_1,\nu_2)\right) \le \sqrt{\frac{1-2t}{1+2t\eta}}e^{-g_1(c,t,\xi,\tau,\eta,w,N)} \tag{A.19}$$

$$-\frac{1}{2\eta} \le t \le 0$$

where

$$\begin{aligned} g_1(c,t,\xi,\tau,\eta,h,N) &= \frac{4t\eta nc + 2\sqrt{2Nc}\xi\Delta_w - \Delta_w^2\xi^2}{2(1+2t\eta)} \\ &- \frac{\tau^2}{2(1-2t)} - \frac{(\xi-\tau)^2\alpha^2}{2(1-2t)} + \frac{N}{2}\ln(1-2t). \end{aligned} \tag{A.20}$$

Maximizing the RHS of (A.19) w.r.t. τ yields

$$\begin{aligned} \frac{\partial g_1}{\partial \tau} &= -\frac{\tau}{1-2t} + \frac{(\xi-\tau)\alpha^2}{1-2t} = 0 \\ \Rightarrow \tau^* &= \frac{\alpha^2\xi^*}{1+\alpha^2}. \end{aligned} \tag{A.21}$$

Notice that $\frac{\partial^2 g_1}{\partial \tau^2} < 0$, hence plugging τ^* in (A.20) maximizes g_1. Substituting τ^* into (A.20) gives

$$\begin{aligned} g_2(c,t,\xi,\eta,w,N) &\triangleq g_1(c,t,\xi,\tau^*,\eta,w,N) \\ &= \frac{4t\eta nc + 2\sqrt{2Nc}\Delta_w\xi - \Delta_w^2\xi^2}{2(1+2t\eta)} - \frac{\frac{\alpha^2}{1+\alpha^2}\xi^2}{2(1-2t)} + \frac{N}{2}\ln(1-2t). \end{aligned} \tag{A.22}$$

A differentiation of g_2 w.r.t. ξ and an introduction of the new parameter $\epsilon \triangleq \frac{\alpha^2}{1+\alpha^2}$ gives

$$\begin{aligned} \frac{\partial g_2}{\partial \xi} &= \frac{\sqrt{2Nc}\Delta_w - \Delta_w^2 \xi}{1+2t\eta} - \frac{\epsilon \xi}{1-2t} = 0 \\ \xi^* &= \frac{\sqrt{2Nc}\Delta_w (1-2t)}{\Delta_w^2(1-2t) + \epsilon(1+2t\eta)}. \end{aligned} \tag{A.23}$$

Again, $\frac{\partial^2 g_2}{\partial \xi^2} < 0$, so ξ^* maximizes g_2. From (A.21), $\xi^* - \tau^* > 0$. Since α is non-negative, we get that K^* in (A.17) is not-positive. But since from (A.11), $K \geq 0$, this yields that the optimal value of K is equal to zero. From the Chernoff bound in (A.5), an optimality of K when it is set to zero implies that asymptotically, as $N \to \infty$

$$\Pr\left(\sum_{i=2}^{N} \nu_i^2 \leq r_{\nu_1}^2, \nu_2 \geq \beta_w(\nu_1), \nu_3 \geq -l_{w,w}(\nu_1,\nu_2)\right) \doteq \Pr\left(\sum_{i=2}^{N} \nu_i^2 \leq r_{\nu_1}^2, \nu_2 \geq \beta_w(\nu_1)\right). \tag{A.24}$$

Case 2: $h > w$. In this case, from (2.40) it is obvious that $\rho_{w,h} = \sqrt{\frac{w(N-h)}{h(N-w)}}$. Hence, for this case, we get that $\rho_{w,h} = \zeta_{w,h}$. From (A.12)

$$\begin{aligned} g(c,t,k,s,\eta,h,N) &= \frac{4t\eta nc + 2\sqrt{2Nc}\left(s - \frac{k\zeta_{w,h}}{\sqrt{1-\zeta_{w,h}^2}}\right)\Delta_h - \Delta_h^2\left(s - \frac{k\zeta_{w,h}}{\sqrt{1-\zeta_{w,h}^2}}\right)^2}{2(1+2t\eta)} \\ &- \frac{\left(s - \frac{k\zeta_{w,h}}{\sqrt{1-\zeta_{w,h}^2}}\right)^2}{2(1-2t)} - \frac{k^2}{2(1-2t)} + \frac{N}{2}\ln(1-2t) - nr(\frac{h}{N}). \end{aligned} \tag{A.25}$$

In the following, we introduce the parameters

$$\xi \triangleq s - \frac{k\zeta_{w,h}}{\sqrt{1-\zeta_{w,h}^2}} \tag{A.26}$$

$$\tau \triangleq k. \tag{A.27}$$

Optimization over τ yields $\tau^* = 0$, so $K^* = 0$, and asymptotically (as we let N tend to infinity), one gets the following equality in terms of the exponential behaviors:

$$\Pr\left(\sum_{i=2}^{N} \nu_i^2 \leq r_{\nu_1}^2, \nu_2 \geq \beta_h(\nu_1), \nu_3 \geq -l_{w,h}(\nu_1,\nu_2)\right) \doteq \Pr\left(\sum_{i=2}^{N} \nu_i^2 \leq r_{\nu_1}^2, \nu_2 \geq \beta_h(\nu_1)\right). \tag{A.28}$$

Case 3: $h < w$. From (2.40), the values of h approve that $\rho_{w,h} = \sqrt{\frac{h(N-w)}{w(N-h)}}$, so we get from (A.10) that $\rho_{w,h} < \zeta_{w,h}$. Define

$$\xi \triangleq s - \frac{k\zeta_{w,h}}{\sqrt{1-\rho_{w,h}^2}} \tag{A.29}$$

$$\tau \triangleq s - \frac{k\rho_{w,h}}{\sqrt{1-\rho_{w,h}^2}}. \tag{A.30}$$

From (A.29) and (A.30)

$$k = -(\xi - \tau)\alpha' \tag{A.31}$$

where

$$\alpha' \triangleq \frac{\sqrt{1-\rho_{w,h}^2}}{\zeta_{w,h} - \rho_{w,h}}. \tag{A.32}$$

Since in this case $\rho_{w,h} < \zeta_{w,h}$, then $\alpha' > 0$. Similarly to the arguments in case 1, we get again that the optimal value for K is $K^* = 0$, which implies (A.28) in the limit where the block length tends to infinity.

Appendix B

Derivation of the Chernoff Bound in (A.11) with the Function g in (A.12)

Using the Chernoff bound (A.5) and defining

$$\Delta_w \triangleq \sqrt{\frac{w}{N-w}} \tag{B.1}$$

we get

$$\begin{aligned}
&\Pr\left(\sum_{i=2}^{N}\nu_i^2 \le r_{\nu_1}^2, \nu_2 \ge \beta_h(\nu_1), \nu_3 \ge -l_{w,h}(\nu_1,\nu_2)\right)\\
&\overset{(a)}{\le} \mathbb{E}\left[e^{t\left(\sum_{i=2}^{N}\nu_i^2 - r_{\nu_1}^2\right)+s(\nu_2-\beta_h(\nu_1))+k\left(\nu_3+l_{w,h}(\nu_1,\nu_2)\right)}\right], \qquad t\le 0,\ s\ge 0,\ k\ge 0\\
&\overset{(b)}{=} \mathbb{E}\left[e^{t\left(\sum_{i=2}^{N}\nu_i^2-\eta(\sqrt{2nc}-\nu_1)^2\right)+s\left(\nu_2-\Delta_h(\sqrt{2nc}-\nu_1)\right)+k\left(\nu_3+\frac{\Delta_w(\sqrt{2nc}-\nu_1)-\rho_{w,h}\nu_2}{\sqrt{1-\rho_{w,h}^2}}\right)}\right]\\
&= \mathbb{E}\left[e^{t\sum_{i=2}^{N}\nu_i^2-t\eta\nu_1^2-2tn\eta c+2\eta t\sqrt{2nc}\nu_1+s\nu_2-s\Delta_h\sqrt{2nc}+s\Delta_h\nu_1+k\nu_3+\frac{k\Delta_w\sqrt{2nc}}{\sqrt{1-\rho_{w,h}^2}}-\frac{k\Delta_w\nu_1+\rho_{w,h}\nu_2}{\sqrt{1-\rho_{w,h}^2}}}\right]\\
&\overset{(c)}{=} \mathbb{E}\left[e^{t\sum_{i=4}^{N}\nu_i^2}\right]\mathbb{E}\left[e^{-t\eta\nu_1^2+\left(2\eta t\sqrt{2nc}+s\Delta_h-\frac{k\Delta_w}{\sqrt{1-\rho_{w,h}^2}}\right)\nu_1}\right]\mathbb{E}\left[e^{t\nu_2^2+\left(s-\frac{k\rho_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)\nu_2}\right]\\
&\quad\cdot\mathbb{E}\left[e^{t\nu_3^2+k\nu_3}\right]e^{-2tn\eta c-s\Delta_h\sqrt{2nc}+\frac{k\Delta_w\sqrt{2nc}}{\sqrt{1-\rho_{w,h}^2}}}.
\end{aligned} \tag{B.2}$$

where inequality (a) follows from the Chernoff bound (A.5), equality (b) follows from (A.1), and equality (c) follows from the statistical independence of the components of the normalized noise vector ν. For a zero-mean and unit-variance Gaussian random variable X, the following equality holds:

$$\mathbb{E}\left[e^{aX^2+bX}\right] = \frac{e^{\frac{b^2}{2(1-2a)}}}{\sqrt{1-2a}}, \quad a \leq \frac{1}{2},\ b \in \mathbb{R}. \tag{B.3}$$

Evaluating each term in (B.2) with the equality in (B.3), and substituting

$$\zeta_{w,h} = \frac{\Delta_w}{\Delta_h} \tag{B.4}$$

which follows from (A.10) and (B.1), then gives

$$\mathbb{E}\left[e^{t\sum_{i=4}^{N}\nu_i^2}\right] = \left(\frac{1}{\sqrt{1-2t}}\right)^{N-3}, \quad t \leq 0 \tag{B.5}$$

$$\mathbb{E}\left[e^{-t\eta\nu_1^2+\left(2\eta t\sqrt{2nc}+s\Delta_h-\frac{k\Delta_w}{\sqrt{1-\rho_{w,h}^2}}\right)\nu_1}\right] = \frac{1}{\sqrt{1+2t\eta}}\, e^{\frac{\left(2\eta t\sqrt{2nc}+\Delta_h\left(s-\frac{k\zeta_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)\right)^2}{2(1+2t\eta)}} \tag{B.6}$$

$$\mathbb{E}\left[e^{t\nu_2^2+\left(s-\frac{k\rho_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)\nu_2}\right] = \frac{1}{\sqrt{1-2t}}\, e^{\frac{\left(s-\frac{k\rho_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)^2}{2(1-2t)}}, \quad k \geq 0,\ s \geq 0 \tag{B.7}$$

$$\mathbb{E}\left[e^{t\nu_3^2+k\nu_3}\right] = \frac{1}{\sqrt{1-2t}}\, e^{\frac{k^2}{2(1-2t)}}, \quad t \leq 0,\ k \geq 0. \tag{B.8}$$

From (B.6), straightforward algebra gives

$$\mathbb{E}\left[e^{-t\eta\nu_1^2+\left(2\eta t\sqrt{2nc}+s\Delta_h-\frac{k\Delta_w}{\sqrt{1-\rho_{w,h}^2}}\right)\nu_1}\right] e^{-2tn\eta c-s\Delta_h\sqrt{2nc}+\frac{k\Delta_w\sqrt{2nc}}{\sqrt{1-\rho_{w,h}^2}}}$$

$$= \frac{1}{\sqrt{1+2t\eta}} \exp\left\{\frac{-4t\eta nc - 2\sqrt{2nc}\left(s-\frac{k\zeta_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)\Delta_h + \Delta_h^2\left(s-\frac{k\zeta_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)^2}{2(1+2t\eta)}\right\}. \tag{B.9}$$

Plugging (B.5) and (B.7)–(B.9) into (B.2) finally gives

$$
\begin{aligned}
&A_h \Pr\left(\sum_{i=2}^{N} \nu_i^2 \le r_{\nu_1}^2, \nu_2 \ge \beta_h(\nu_1), \nu_3 \ge -l_{w,h}(\nu_1,\nu_2)\right) \\
&\le \frac{A_h}{\sqrt{1+2t\eta}} \left(\frac{1}{\sqrt{1-2t}}\right)^{N-1} e^{\frac{-4t\eta nc - 2\sqrt{2nc}\left(s-\frac{k\zeta_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)\Delta_h + \Delta_h^2 \left(s-\frac{k\zeta_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)^2}{2(1+2t\eta)} + \frac{\left(s-\frac{k\rho_{w,h}}{\sqrt{1-\rho_{w,h}^2}}\right)^2}{2(1-2t)} + \frac{k^2}{2(1-2t)}} \\
&= \sqrt{\frac{1-2t}{1+2t\eta}}\, e^{-g(c,t,k,s,\eta,h,N)}, \qquad -\frac{1}{2\eta} < t \le 0,\ k \ge 0,\ s \ge 0
\end{aligned} \tag{B.10}
$$

which proves the Chernoff bound in (A.11) with the function g introduced in (A.12).

Appendix C

Monotonicity w.r.t. the Correlation Coefficient

Consider the probabilities $\Pr(E_{0\to i}, E_{0\to j}^{\mathrm{c}}, \mathbf{y} \in C_N(\theta)|z_1)$, and denote the Hamming weights of $\mathbf{c}_i$ and $\mathbf{c}_j$ by d_i and d_j, respectively. In [47], it is shown that as long as $d_i > d_j$, the probabilities $\Pr(E_{0\to i}, E_{0\to j}^{\mathrm{c}}, \mathbf{y} \in C_N(\theta)|z_1)$ are monotonically decreasing functions of the correlation coefficients ρ between the planes $(\mathbf{o}, \mathbf{s}_0, \mathbf{s}_i)$ and $(\mathbf{o}, \mathbf{s}_0, \mathbf{s}_j)$. Hence, the complex optimization problem in (2.24) is simplified by choosing the first error event as well as the complementary error events in the RHS of (2.24) to correspond to a codeword with Hamming weight $d_{\min}$, and (2.25) is obtained. Here we prove, that the aforementioned probabilities are monotonically decreasing functions of the correlation coefficients for *any* choice of i, j. As a consequence, one can obtain a version of the ITSB by setting in (2.24) $\pi_1 = \lambda_i = w$ where $w \in \{d_{\min}, \ldots, d_{\max}\}$, and choosing the optimal w which minimizes the resulting upper bound. In order to prove this, we follow the steps in [47, Appendix I] where it is shown that the above probabilities are monotonically decreasing functions of ρ if

$$\frac{z_2}{\beta_j(z_1)} > \rho. \tag{C.1}$$

Note that the joint event $(E_{0\to i},\ \mathbf{y} \in C_N(\theta))$ implies that the noise component z_2 is in the range between $\beta_i(z_1)$ and r_{z_1} (see Fig. 2.1 in p. 12), so the minimum value of the RHS of (C.1) is

$$\frac{\beta_i(z_1)}{\beta_j(z_1)} = \sqrt{\frac{d_i(N-d_j)}{d_j(N-d_i)}}.$$

Clearly,

$$\sqrt{\frac{d_i(N-d_j)}{d_j(N-d_i)}} > \frac{\min(d_i,d_j)[N-\max(d_i,d_j)]}{\sqrt{d_i d_j(N-d_i)(N-d_j)}} \tag{C.2}$$

but from (2.34), it is evident that the RHS of (C.2) is the maximal value of ρ, thus, condition (C.1) is always satisfied referring to the joint event $(E_{0\to i},\ \mathbf{y} \in C_N(\theta))$.

Appendix D

The Average Distance Spectrum of the Ensemble of Random Linear Block Codes

In [3, Eq. (2)], Ashikmin and Barg introduce a formula for the average spectrum of a random *linear* code. The formula coincides with the well-known average spectrum of fully random code. Clearly, Eq. (2) there can not be the exact expression for the average distance spectrum of the ensemble of random linear block codes, since for linear codes, the all-zero vector is always a codeword which forces $\mathbb{E}[A_0] = 1$. In order to obtain the exact expression, we follow the arguments in [3], while limiting ourselves to the case of *binary* codes (the generalization for q-ary codes is straightforward). Consider the ensemble $\mathcal{C}$ which contains all the (N, K) linear codes. Let $\mathbb{E}[A_w]$ denote the average number of codewords with Hamming weight w. The probability that the first row in the parity-check matrix of some code from $\mathcal{C}$ satisfies a check equation for a specific codeword of Hamming weight w is $\frac{2^{N-1}-1}{2^N-1}$. The reason for the subtraction of 1 from both the nominator and denominator of the above expression is that we do not allow an all-zero row in the parity check matrix (otherwise the code rate will be *below* $\frac{K}{N}$). The probability that the i^{th} row satisfies the check equation is $\frac{2^{N-1}-2^{i-1}}{2^N-2^{i-1}}$ (we subtract all the rows that are linearly dependent with the first $i-1$ rows); therefore,

the probability that this codeword is contained in a code from $\mathcal{C}$ equals

$$\prod_{i=1}^{N-K} \frac{2^{N-1} - 2^{i-1}}{2^N - 2^{i-1}} = \frac{2^K - 1}{2^N - 1} \quad w \neq 0$$

Thus

$$\mathbb{E}[A_w] = \begin{cases} \binom{N}{w} \frac{2^K - 1}{2^N - 1} & 0 < w \leq n \\ 1 & w = 0 \end{cases} \tag{D.1}$$

It can be easily verified that $\sum_{w=0}^{N} \mathbb{E}[A_w] = 2^K$ as one could expect. Moreover, for the asymptotic case where $N, K \to \infty$, (D.1) converges to Eq. (2) in [3] (for $w \neq 0$). The problem with Ashikmin and Barg derivation is that they assume that the rows of the parity check matrix is statistically independent (which is correct for the asymptotic case). Another way to look at Ashikmin and Barg formula is to consider the $\frac{K}{N}$ as the *design* rate of the code. Anyway, the expression in [3] can be used as an (asymptotically tight) upper bound on the average distance spectrum of the ensemble of linear block codes.

References

[1] E. Agrell, "Voronoi regions for binary linear block codes," *IEEE Trans. on Information Theory*, vol. 42, no. 1 pp. 310–316, January 1996.

[2] E. Agrell, "On the Voronoi neighbors ratio for binary linear block codes," *IEEE Trans. on Information Theory*, vol. 44, no. 7 pp. 3064–3072, November 1998.

[3] A. Ashikmin and A. Barg, "Minimal vectors in linear codes," *IEEE Trans. on Information Theory*, vol. 44, no. 5 pp. 2010–2017, November 1998.

[4] A. Barg and G.D.Forney "Random codes: minimum distances and error exponents," *IEEE Trans. on Information Theory*, vol. 48, no. 9 pp. 2568–2573 , September 2002.

[5] A. Bennatan and D. Burshtein, "On the application of LDPC codes to arbitrary discrete-memoryless channels," *IEEE Trans. on Information Theory*, vol. 50, pp. 417–438, March 2004.

[6] E. R. Berlekamp, "The technology of error correction codes," *Proc. of the IEEE*, vol. 68, no. 5, pp. 564–593, May 1980.

[7] C. Berrou, A. Glavieux and P. Thitimajshima,"Near Shannon limit error-correcting coding and decoding," *Proceedings 1993 IEEE International Conference on Communications (ICC '93)*, pp. 1064–1070, Geneva, Switzerland, May 1993.

[8] C. Brutel and J. Boutros, "Serial concatenation of interleaved convolutional codes and M-ary continuous phase modulations," *Annals of Telecommunications*, vol. 54, no. 3–4, pp. 235–242, March–April 1999.

[9] R. Corvaja and S. Pupolin,"Phase noise effects in QAM systems," *IEEE Int. Symp. Pers., Indoor, Mobile Radio Communication,* vol. 2, p. 452, 1997.

[10] A. Demir, A. Mehrotra and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst. I.* vol. 47 no. 5, pp. 655–674, May 2000.

[11] D. Divsalar, "A simple tight bound on error probability of block codes with application to Turbo codes," *TMO progress Report 42-139* NASA, JPL, Pasadena, CA, USA, 1999.

[12] D. Divsalar and E. Biglieri, "Upper bounds to error probabilities of coded systems beyond the cuttoff rate," *IEEE Trans. on Communications*, vol. 51, pp. 2011–2018, December 2003.

[13] D. Divsalar, H. Jin, and J. McEliece, "Coding theorems for turbo-like codes", *1998 Allerton Conference*, Sep. 23–25,1998.

[14] T. M. Duman and M. Salehi,"New performance bounds for turbo codes," *IEEE trans. on Communications*, vol. 46, no. 6, pp. 717–723, June 1998.

[15] T. M. Duman,"Turbo codes and turbo coded modulation systems: Analysis and performance bounds," Ph.D. dissertation, Elect. Comput. End. Dep., Northeastern University, Boston, MA, USA, May 1998.

[16] R. M. Fano, *Transmission of Information*, jointly published by the MIT Press and John Wiley & Sons, 1961.

[17] J. Galambos and I. Simonelli, *Bonferroni-type inequalities with applications*, Springer Series in Statistics, Probability and its Applications, Springer-Verlag, New-York, 1996.

[18] M. Fossorier, "Critical point for maximum-likelihood decoding of linear block codes," *IEEE Communications Letters*, vol. 9, no. 9, pp. 817–819, September 2005.

[19] R. G. Gallager,"*Low-Density Parity-Check Codes*," Cambridge, MA USA, MIT press, 1963.

[20] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. on Information Theory*, vol. 11, no. 1, pp. 3–18, January 1965.

[21] H. Herzberg and G. Poltyrev, "Techniques of bounding the probability of decoding error for block modulation structures", *IEEE trans. on Information Theory,* vol. 40, no. 3, pp. 903–911, May 1994.

[22] H. Herzberg and G. Poltyrev, "The error probability of M-ary PSK block coded modulation schemes," *IEEE Trans. on Communications*, vol. 44, pp. 427-433, April 1996.

[23] D. Hunter, "An upper bound for the probability of a union," *Journal of Applied Probability*, vol.13, pp. 597–603, 1976.

[24] H. Jin and R. J. McEliece, "Typical pairs decoding on the AWGN channel," *Proceedings 2000 IEEE International Symposium on Information Theory and its Applications*, pp. 180–183, Honolulu, Hawaii, USA, November 5–8, 2000.

[25] H. Jin and R. J. McEliece, "Coding theorems for turbo-like ensembles," *IEEE Trans. on Information Theory*, vol. 48, pp. 1451–1461, June 2002.

[26] F. Kienle, T. Brack and N. Wehn, "A Synthesizable IP COre for DVB-S2 LDPC Code Decoding," *Design, Automation and Test in Europe (DATE'05)*, vol. 3 pp. 100-105, 2005.

[27] D.J.C. Mackay and R.M. Neal, "Near Shannon limit performance of low-density parity-check codes", *IEEE Electronic Letters*, vol. 33, no. 6, pp. 457–458, March 1997.

[28] G. Miller and D. Burshtein, "Bounds on the ML decoding error probability of LDPC codes," *IEEE Trans.on Information Theory*, vol. 47, no. 7, pp. 2696–2710, November 2001.

[29] L. Piazzo and P. Mandarini,"Analysys of phase noise effects in OFDM modems," *IEEE Trans. Commun.,* vol. 50, no. 10, pp. 1696–1705, Oct. 2002.

[30] H. Pfister and I. Sason, "Capacity-Achieving Ensembles of Accumulate-Repeat-Accumulate Codes for the Erasure Channel with Bounded Complexity," submitted to *IEEE Trans. on Information Theory*, December 2005. See http://arxiv.org/abs/cs.IT/0512006.

[31] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra,"*IEEE Trans. on Information Theory*, vol. IT–40, no. 4, pp. 1284–1292, July 1994.

[32] I. Sason and S. Shamai, "Bounds on the error probability of ML decoding for block and turbo-block codes," *Annals of Telecommunication*, vol. 54, no. 3–4 pp. 183–200, March–April 1999.

[33] I. Sason and S. Shamai, "Improved upper bounds on the ML decoding error probability of parallel and serial concatenated Turbo codes via their ensemble distance spectrum," *IEEE Trans. on Information Theory*, vol. 46, no. 1 pp. 24–47, January 2000.

[34] I. Sason and S. Shamai, "On Improved bounds on decoding error probability of block codes over interleaved fading channels, with applications to Turbo-like codes," *IEEE Trans. on Information Theory*, vol. 47, no. 6, pp. 2275–2299, September 2001.

[35] I. Sason and S. Shamai, "Performance analysis of linear codes under maximum-likelihood decoding: a tutorial," submitted to *Foundations and Trends in Communications and Information Theory*, NOW Publishers, Delft, the Netherlands, December 2005.

[36] I. Sason and R. Urbanke, "Parity-check density versus performance of binary linear block codes over memoryless symmetric channels," *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1611–1635, July 2003.

[37] N. Shulman and M. Feder, "Random coding techniques for nonrandom codes," *IEEE Trans. on Information Theory*, vol. 45, no. 6, pp. 2101–2104, September 1999.

[38] S. Shamai and I. Sason, "Variations on the Gallager bounds, connections, and applications, " *IEEE Trans. on Information Theory*, vol. 48, no. 12, pp. 3029–3051, December 2002.

[39] M. Sipser and D. A. Spielman, "Expander Codes," *IEEE Trans. on Information Theory*, vol. 42, pp. 1710–1722, Nov, 1996.

[40] E. Soljanin and R. Urbanke," On the performance of recursive decoding schemes," Bell laboratories, Lucent technologies, Technical Report, 1996. [online]. Available: http://cm.bell-labs.com/who/ruediger/pub.html.

[41] C-E. Sundberg, "Continuous phase modulations," *IEEE Communication Magazine*, vol. 24, no. 4, pp. 25–38, April 1986.

[42] M. Twitto, I. Sason and S. Shamai, "Tightened upper bounds on the ML decoding error probability of binary linear block codes," submitted to the *IEEE Trans. on Information Theory*, February 2006.

[43] M. Twitto, I. Sason and S. Shamai, "Tightened upper bounds on the ML decoding error probability of binary linear codes," submitted to the *Proceedings 2006 IEEE International Symposium on Information Theory*, 2006.

[44] M. Twitto and I. Sason, "On the Error Exponents of Some Improved Tangential-Sphere Bounds," accepted to the *IEEE Trans. on Information Theory*, August 2006.

[45] S. Yousefi and A. Khandani, "Generelized tangential sphere bound on the ML decoding error probability of linear binary block codes in AWGN interference", *IEEE Trans. on Information Theory*, vol. 50, no. 11, pp. 2810–2815, Nov. 2004.

[46] S. Yousefi and A. K. Khandani, "A new upper bound on the ML decoding error probability of linear binary block codes in AWGN interference ," *IEEE Trans. on Information Theory*, vol. IT–50, no. 12, pp. 3026–3036, December 2004.

[47] S. Yousefi and A. Mehrabian, "Improved tangential sphere bound on the ML decoding error probability of linear binary block codes in AWGN interference," *Proceedings* 37^{th} *Annual Conference on Information Science and Systems (CISS 2005),* John Hopkins University, Baltimor, MD, USA, March 16–18, 2005.

[48] S. Yousefi, "Gallager first bounding technique for the performance evaluation of maximum-likelihood decoded linear binary block codes," *IEE Proceedings on Communications*, vol. 153, no. 3, pp. 317–332, June 2006.

[49] A. M. Viterbi and A. J. Viterbi, "An improved union bound for binary linear codes on the AWGN channel, with application to turbo decoding," *Proceedings of IEEE Information Theory Workshop*, p. 72, San Diego, California, USA, February 1998.

[50] A. M. Viterbi and A. J. Viterbi, "Improved union bound on linear codes for the binary-input AWGN channel, with application to turbo codes," *Proceedings 1998 IEEE International Symposium on Information Theory (ISIT 1998)*, MIT, Cambridge, MA, USA, p. 29, August 16–21, 1998.

[51] J. Zangl and R. Herzog, "Improved tangential-sphere bound on the bit error probability of concatenated codes," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 5, pp. 825–837, May 2001.

[52] A Turbo Encoder-Decoder of the Altera company. [online]. Available: http://www.altera.com/product/ip/altera/m-alt-turbo-dec.html

[53] Texas Instruments, *Implementing a MAP Decoder for CDMA2000 Turbo Codes on a TMS320C62X DSP Device*, Application Report, May 2000.

חסמים עליונים משופרים על הסתברות השגיאה של קודי בלוק ליניאריים ובינאריים ויישומים

משה תואיתו

חסמים עליונים משופרים על הסתברות השגיאה של קודי בלוק ליניאריים ובינאריים ויישומים

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת תואר

מגיסטר למדעים

הנדסת חשמל

משה תואיתו

הוגש לסנט הטכניון — מכון טכנולוגי לישראל

ניסן תשס"ו חיפה אפריל 2006

חיבור על מחקר נעשה בהדרכת דר' יגאל ששון
בפקולטה להנדסת חשמל

הכרת תודה

ברצוני להביע את תודתי העמוקה למנחה שלי, דר' יגאל ששון, על הנחי-
יתו המסורה והסובלנית ועזרתו הרבה במשך כל שלבי המחקר.
כמו כן ברצוני להודות להוריי היקרים, רימון ושרה, על התמיכה והעידוד
שהעניקו לי במהלך לימודי. עבודה זו מוקדשת להם.

אני מודה לטכניון על התמיכה הכספית הנדיבה בהשתלמותי

להורי היקרים,
רימון ושרה.

תוכן עניינים

רשימת איורים

תקציר

עבור רוב מערכות התקשורת המקודדות, לא ניתן לקבל ביטוי מדויק עבור הסתברות השגיאה תחת פענוח סבירות מירבית או פענוח תת-אופטימאלי מעשי. כדי שניתן יהיה בכל זאת להעריך ביצועים של מערכות כאלו, יש צורך בשימוש בחסמים על הביצועים של מערכות אלו, או לחילופין, להשתמש בסימולציות מחשב. החיסרון הבולט בשימוש בסימולציות מחשב ככלי להערכת הביצועים של מערכת, היא משך הזמן הניכר הדרוש להערכת הסתברות שגיאה נמוכה, בעוד ששימוש בחסמים מאפשר קבלת תוצאות בפרק זמן קצר ביותר. יתרה מזאת, ביטוי אנליטי המהווה חסם על הסתברות השגיאה מהווה כלי עזר תיאורטי והנדסי חשוב, ומאפשר לקבל תובנה הנדסית על מידת ההשפעה של פרמטרי המערכת על ביצועי המערכת.

אחד החסמים העליונים הנפוצים ביותר הינו חסם האיחוד. חסם זה מצטיין בפשטותו הרבה וניתן ליישם אותו עבור מגוון רחב של ערוצים ומערכות. חיסרונו העיקרי הוא שעבור קודים ארוכים, הוא חסר תועלת בקצבים הגבוהים מקצב הקיטעון של הערוץ. במשך עשרות שנים היה חסם זה מספק עבור רוב מערכות התקשורת הקיימות, למרות החיסרון שהזכרנו לעיל, מהסיבה הפשוטה שרוב מערכות התקשורת המקודדת עבדו עם קודים בע-לי קצב הנמוך מקצב הקיטעון של הערוץ. הצגת קודי הטורבו לפני כעשור שינתה את פני הדברים. קודים אלו מאפשרים לקבל הסתברות שגיאה נמוכה ביותר בקצבים הקרובים לקיבול הערוץ, וזאת תחת פענוח איטרטיבי מעשי, הניתן למימוש בסיבוכיות מתקבלת על הדעת. תוך שנים ספורות הפכו קודים אלו לקודים סטנדרטיים במערכות תקשורת מודר-ניות רבות (כגון שידורי וידאו לווייני, הדור השלישי של התקשורת התאית, ועוד). עובדה זו היוותה תמריץ חזק מאוד לפיתוח חסמים הדוקים יותר על הסתברות השגיאה, בפרט עבור קודים בעלי קצב הגבוה מקצב הקיטעון של הערוץ. קודי הטורבו שהזכרנו מפוענחים בדרך כלל באמצעות שיטות מעשיות תת-אופטימאליות. למרות זאת, יש עניין בפיתוח חסמים עליונים על הסתברות השגיאה של מפענח הסבירות המרבית האופטימאלי, היות והדבר נותן אינדיקציה לגבי היכולת ברת ההשגה האולטימטיבית של המערכת. לרוב, לא ניתן לאפיין את איזורי ההחלטה של הקוד בצורה מלאה, ולכן חסמים ברי-שימוש יהיו

כאלו שישענו אך ורק על תכונות בסיסיות כגון ספקטרום המרחקים של הקוד, אותו ניתן לקבל אנליטית עבור מגוון רחב של קודים. גם את ספקטרום המרחקים קשה לעיתים לחשב עבור קוד ספציפי. כדי להתגבר על בעיה זו, מגדירים בדרך כלל צביר של קודים שמכיל את הקוד הרצוי, ועבור צביר זה מחשבים את ספקטרום המרחקים הממוצע, אותו ניתן לחשב לעיתים רבות ביתר קלות. כתוצאה מזאת, רצוי שחסם עליון טוב יהיה ניתן להחלה גם על הסתברות השגיאה הממוצעת של צביר קודים, ולא רק על הסתברות השגיאה של קוד ספציפי. אחד החסמים העליונים ההדוקים ביותר הידועים עד כה, הינו החסם המשיקי-כדורי. חסם זה תקף לכל קוד בלוק שמילותיו, לאחר אפנון, הינן בעלי אותה אנרגיה, ואשר משודרות על גבי הערוץ הגאוסי הלבן. חסם זה הדוק לפחות כמו חסם האיחוד בכל ערך של יחס אות לרעש, והוא נותן תוצאות אינפורמטיביות גם עבור קודים בעלי קצב גבוה מקצב הקיטעון של הערוץ. בדומה לחסם האיחוד, ניתן לחשב את החסם המשיקי-כדורי בהתבסס על ספקטרום המרחקים של הקוד בלבד, ולא נדרשת ידיעה של תכונות מורכבות יותר של הקוד. פעמים רבות מתקבל שהחסם המשיקי-כדורי הוא החסם הידוע ההדוק ביותר על הסתברות השגיאה לבלוק ולביט של קודי בלוק המשודרים על פני הערוץ הגאוסי הלבן. לאחרונה הוצגו מספר חסמים עליונים שהינם גרסאות משופרות של החסם המשיקי-כדורי. כמו כן הוצג שיפור קל לעומת החסם המשיקי-כדורי, עבור מספר קודי בלוק קצרים. שאלה חשובה שלא ניתן לה מענה אנליטי עד כה, הינה האם יתרון זה של הגרסאות המשופרות נשמר גם כאשר אורך הבלוק הולך וגדל, ובפרט עבור אורכי בלוק אינסופיים.

בחלקה הראשון של עבודה זו אנו דנים בגרסאות משופרות של החסם המשיקי-כדורי, תוך התמקדות באקספוננטי השגיאה המשויכים לחסמים אלו. אנו מראים שאסימפטוטית, כאשר אורך הקוד שואף לאינסוף, אקספוננטי השגיאה של החסמים הנ"ל זהים לזה של החסם המשיקי-כדורי, ולכן החסמים זהים אקספוננציאלית. מאחר והחסם המשיקי-כדורי פשוט יותר לחישוב מהגרסאות המשופרות שלו, נסיק כי עבור קודים ארוכים, מוטב להשתמש בחסם המשיקי-כדורית מאשר בגרסאות המשופרות שלו שידועות היום.

ידוע שהסתברות השגיאה הממוצעת של צביר קודי הבלוק האקראיים שואפת לאפס כאשר אורך הבלוק שואף לאינסוף, וזאת כל עוד שקצב הקודים קטן מקיבול הערוץ. את אקספוננט השגיאה של צביר זה מצא Gallager באופן מדויק, עבור אורך בלוק אינסופי. בכדי לבחון את הדיקותו של חסם כלשהוא, ניתן להפעיל אותו על צביר זה של קודי בלוק אקראיים, ולהשוות את האקספוננט המתקבל לאקספוננט הצפינה האקראית של גלאגר. מבדיקה כזו נובע שהחסם המשיקי-כדורי לא משחזר את אקספוננט הצפינה האקראית עבור צביר הקודים האקראיים. יתרה מזאת, ניתן להראות שככל שקצב הקוד גדל, כך גדל הפער בין האקספוננט של החסם המשיקי-כדורי לאקספוננט הצפינה האקראית. להבדיל

מהחסם המשיקי-כדורי, חסם אחר, הקרוי "חסם שולמן-פדר", משחזר את חסם הצפינה האקראית עבור צביר קודי הבלוק האקראיים. למרות זאת, עבור קודים מובנים רבים, חסם זה אינו הדוק, ובפרט, נותן תוצאות הדוקות פחות מהחסם המשיקי-כדורי. עובדות אלו נותנות תמריץ לחפש אחר חסם עליון שישחזר את חסם הצפינה האקראית, ויהיה הדוק יותר מחסם שולמן-פדר ומהחסם המשיקי-כדורי, במיוחד עבור קודים בקצב גבוה, היכן שהחולשה של החסם המשיקי-כדורי בולטת ביותר.

בחלק השני של העבודה, אנו מפתחים חסמים עליונים משופרים על הסתברות השגיאה הממוצעת של צברי קודים, תחת פענוח הסבירות המרבית. לצורך כך אנו מניחים שהאותות שודרו על פני ערוצים חסרי-זיכרון, בינאריים במבוא וסימטריים ביציאה. בפיתוח החסמים אנו נעזרים בטכניקת החסימה של הגרסא המוכללת של חסם Duman-Salehi, שאף היא משיגה את אקספוננט הצפינה האקראית עבור צביר קודי הבלוק האקראיים. אנו מראים שהחסמים החדשים נותנים תוצאות טובות יותר מחסם שולמן-פדר, והם פשוטים יותר לחישוב מהגרסא המוכללת של חסם Duman-Salehi. החסמים שאנו מפתחים הינם עבור הסתברות השגיאה לבלוק ולביט של קודי בלוק, וניתן לחשב אותם בהתבסס על ספקטרום המרחקים ופונקצית ספירת משקלי הכניסה והיציאה (IOWEF) של הקוד, בהתאמה. אנו מדגימים את היעילות של החסמים הנ"ל באמצעות מספר צבירים של קודי טורבו בקצב גבוה, המשודרים בערוץ הגאוסי הלבן. מתקבל כי עבור קודים אלו, החסמים החדשים נותנים תוצאות הדוקות יותר מהחסם המשיקי-כדורי, שכאמור, הינו אחד מהחסמים העליונים ההדוקים ביותר הידועים עד כה.

נושא נוסף בו או דנים בקצרה, הינו ההשפעה של טיהור ספקטרום המרחקים של הקוד על הדיקותם של חסמים עליונים שונים, ובכללם החסמים החדשים שפיתחנו. אנו מציגים את השפעת הטיהור באמצעות צביר קודים משורשרים טורית, ומדגימים כיצד ניתן לשפר חסמים עליונים באמצעותו.

התוצאות הטובות המתקבלות משימוש בחסמים העליונים שפיתחנו בעבודה זו, הופכת את טכניקות החסימה האלו לשימושיות לצורך בחינת ביצועי קודי גרף שונים תחת פענוח סבירות מירבית והערכת ההפסד בביצועים כתוצאה משימוש באלגוריתמי פענוח איטרטיביים מעשיים שהינם תת-אופטימאליים.