Analysis and Design of Bit Interleaved Coded Modulation Based Transceivers

by

Muhammad Talha Malik

B.Sc., GIK Institute of Engineering Sciences and Technology, Pakistan, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE COLLEGE OF GRADUATE STUDIES

(Electrical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

July 2013

Abstract

Bit-interleaved coded modulation (BICM) is a very popular approach for spectrally efficient coded transmission. In BICM, the channel encoder is separated from the modulator by a bit level interleaver. The presence of a random interleaver in BICM allows the designer a flexibility to independently choose the code rate and the modulation order. This allows an easy adaptation of the transmission rate to the channel conditions. BICM maximizes the code diversity, and therefore, is a superior alternative to the conventional trellis coded modulation (TCM) in fading channels. In additive white Gaussian noise (AWGN) channels, BICM is suboptimal because it reduces the minimum Euclidean distance. However, its flexibility and ease of implementation makes it an attractive scheme even for transmission over non-fading channels. The objective of this thesis is to investigate and optimize new BICM designs in order to further improve the BICM-based transceivers. First, we develop an analytical framework for performance evaluation of a new generalized BICM (referred to as BICM-T) transmission over AWGN channels that can be used to predict and optimize the error rate performance of such systems. Second, we investigate the performance of BICM in non-Gaussian channels due to its practical relevance. Moreover, because of its various advantages in designing a wireless transceiver, we real-

Abstract

ize that a BICM-based transceiver will be the natural choice for cooperative communication systems. Therefore, we present an innovative BICM design for cooperative communication where various BICM modules can be optimized jointly considering the average signal to noise ratios of the direct and the two-hop Rayleigh fading channels. The results presented in this thesis show that by optimizing different system modules of our proposed BICM designs, significant gains in the bit error rate (BER) performance can be achieved.

Preface

This work has been done under the supervision of Dr. Md. Jahangir Hossain at the School of Engineering in The University of British Columbia. Parts of this thesis have been submitted in two journal articles:

- M. T. Malik, Md. J. Hossain, and M.-S. Alouini, "BICM-based cooperative communication systems with relay selection: constellation and multiplexer Design", *IEEE Wireless Commun. Letters*, Submitted June 2013.

- M. T. Malik, Md. J. Hossain and M.-S. Alouini, "Analysis and design of generalized BICM-T transceivers", *IEEE Trans. Commun.*, Submitted July 2013.

Part of this thesis has been accepted to a conference:

- M. T. Malik, Md. J. Hossain and M.-S. Alouini, "Generalized BICM-T transceivers: constellation and multiplexer Design", *IEEE Intl. Sym Personal, Indoor and Mobile Radio Commun.*, Accepted for Publication, June 2013.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AWGN Additive White Gaussian Noise

BER Bit Error Rate

BICM Bit Interleaved Coded Modulation

BICM-**T** Trivial Bit Interleaved Coded Modulation

BICM-**S** Bit Interleaved Coded Modulation with Single Interleaver

BISO Binary Input Soft Output

BRGC Binary Reflected Gray Code

CM Coded Modulation

DEC Viterbi Decoder

DE-**MUX** De-Multiplexer

ES Equally Spaced

ENC Channel Encoder

EWDS Equivalent Weight Distribution Spectrum

GMN Gaussian Mixture Noise

LIST OF ABBREVIATIONS

HPAM   Hierarchical Pulse Amplitude Modulation

HQAM   Hierarchical Quadrature Amplitude Modulation

MUX   Multiplexer

ODS   Optimum Distance Spectrum

OWC   Optical Wireless Communications

PAM   Pulse Amplitude Modulation

PDF   Probability Density Function

QAM   Quadrature Amplitude Modulation

REC   Receiver

SNR   Signal to Noise Ratio

TCM   Trellis Coded Modulation

UB   Union Bound

UEP   Unequal Error Protection

Acknowledgements

I would like to convey my sincere gratitude to Dr. Md. Jahangir Hossain for introducing the interesting topic of Bit Interleaved Coded Modulation. His insightful guidance throughout my thesis was invaluable. Furthermore, I am especially grateful to Prof. Mohamed-Slim Alouini for his continuous support and suggestions to improve the quality of this work.

Chapter 1

Introduction

1.1 Background

Over the past years, the growing demand for high data rate and reliable services has led to the need for the design of efficient communication systems. The main hurdle in wireless communication is the time varying nature of the communication channels which is referred to as fading. In order to combat this fading effect and to increase the data rate, error control coding [1] in conjunction with signal modulation scheme is used. In almost any communication system, the channel encoder is a very important part of the transmitter. The purpose of the channel encoder is to introduce, in a controlled manner, some redundant symbols in the information sequence. This redundant information can be used at the receiver to overcome the effects of noise and interference encountered during the transmission of the signal through the channel.

Channel coding and signal modulation were considered separately in the design of early communication systems. Coded modulation was brought into focus by the pioneering work of Ungerboeck on TCM [2]. In order to maximize the minimum Euclidean distance of the code [2] in TCM, trellis codes are combined with modulation through set partitioning labeling. In

1

1992, in his landmark paper, Zehavi showed that the diversity of the code is improved by decoupling the channel encoder and the modulator resulting in more reliable communication over fading channels [3]. Zehavi's idea was to make the code diversity equal to the smallest number of distinct bits along any error event which is achieved by using a bit-level interleaver at the output of the encoder. This coded modulation scheme is known as BICM. Since BICM separates the channel encoder from the modulator by a bit level interleaver, it also offers the designer a flexibility to choose the code rate and signal modulation independently. After its introduction in Zehavi's landmark paper [3], BICM has been widely generalized by Caire *et. al.* in [4]. Because of this simplicity and flexibility, BICM has become a standard for many contemporary wireless standards e.g., HSPA [5], IEEE 802.11a/g [6], IEEE 802.11n [7] and the DVB standards such as DVB-T2 [8].

1.2   Objectives and Contributions

The objective of this thesis is to investigate and optimize new BICM designs in order to further improve the BICM-based transceivers. First, we develop an analytical framework for performance evaluation of a new generalized BICM transmission over AWGN channels that can be used to predict and optimize the error rate performance of such systems. Second, we investigate the performance of BICM in non-Gaussian channels due to its practical relevance. Moreover, because of its various advantages in designing a wireless transceiver, we realize that a BICM-based transceiver will be the natural choice for cooperative communication systems. Therefore, we

present an innovative BICM design for cooperative communication where various BICM modules can be optimized jointly considering the average signal to noise ratios of the direct and the two-hop Rayleigh fading channels.

In the following, we provide a brief review of the above mentioned objectives. The detailed review of previous research on each topic can be found in the "Introduction" section of corresponding chapter(s).

1.2.1 Performance Evaluation of BICM Transmission over AWGN Channels

Recent results [9, 10] have shown that the performance of BICM using convolutional codes in non-fading channels can be improved significantly when interleaver is removed from the BICM design. This design is referred to as BICM trivial (BICM-T) which has been formally analyzed in [9] and is shown to be asymptotically as good as Ungerboeck's one dimensional (1D)-TCM. However, due to the use of equally spaced (ES) signal constellation in the design, where the distance between neighbouring signal points is equal, it leads to a suboptimal BICM-T transceiver as the BER performance can be further improved by providing proper unequal error protection (UEP) to the coded bits [11, 12]. Moreover, design and analysis of a more general setup, where the number of outputs from the encoder does not match with the number of inputs to the modulator, are not considered. For such a generalized setup, the important question is how to connect the channel encoder to the modulator in order to minimize the BER performance.

In [12], the performance of BICM transceiver is optimized by using hierarchical constellations, a bit-level multiplexer and multiple interleavers.

Motivated by the results presented in [12], in this thesis we consider an innovative BICM-T transceiver by using a bit-level multiplexer and hierarchical constellations. Using the considered more generalized design, the designer can not only exploit different degrees of freedom to notably improve the BER performance of BICM-T design but also enjoys the same design flexibility as the traditional BICM to independently choose the code rate and the modulation order. The performance of the considered BICM-T transceiver is analyzed. In particular, we develop the union bounds (UBs) on the BER of the system for AWGN channels, which are then used to optimize the constellation parameters as well as multiplexer.

1.2.2 Performance Evaluation of BICM Transmission over non-Gaussian Channels

In many physical channels, such as power line communication channels [13], the ambient noise is known through experimental measurements to be non-Gaussian due to the impulsive nature of man-made electromagnetic interference as well as a great deal of natural noise. While the traditional BICM and the BICM-T systems have been investigated and optimized for AWGN channels [9–12], the analysis of BICM transmission impaired by non-Gaussian noise has received relatively little attention [14–16]. In particular there is a need to investigate the performance of BICM-T system for Gaussian mixture noise (GMN) channels.

In BICM-T, a number of consecutive coded bits are transmitted using the same symbol (due to the absence of random interleaver) and the corresponding L-values are mutually dependent. Therefore, a different particu-

larization (e.g., [9]) is adopted to calculate the L-values at the receiver. In particular, using a zero-crossing model [9], closed form expressions for the probability density function (PDF) of L-values are developed in terms of constellation distance parameters of the hierarchical quadrature amplitude modulation (HQAM). As in [14] it is assumed that the system employs the standard Euclidean-distance decoder, which is an instance of mismatched decoding [17] in the presence of GMN. We develop the UB on BER performance of the BICM-T design for GMN channels which is then used to optimize different modules of the BICM-T system over such channels.

1.2.3 Optimization of BICM-Based Cooperative Communication System

Cooperative wireless communication has gained much interest in recent years as it improves the overall system performance [18–20]. On the other hand, BICM [3] offers many advantages in designing wireless transceiver and is a de-facto standard for most of the contemporary wireless systems. Therefore, BICM-based transceiver will also be the natural choice for cooperative communication systems. Significant research has been conducted towards designing BICM-based transceivers for single-hop systems. However, to date, a little research has been conducted towards designing BICM-based transceiver for cooperative communication systems. But none of the works considered the possibility of optimizing interleaver and constellation.

In single-hop communication, in order to optimize the multiplexer and constellation, the transmitter considers the average signal-to-noise ratio (SNR) of the point-to-point link only. However, due to the inherent nature

of the cooperative communication system, optimization of the multiplexer and constellation is different from the single-hop systems. For example, when the source transmits information to the destination and the selected relay (which then forwards the detected symbols to the destination), it has to optimize the multiplexer and constellation jointly considering average SNRs from the destination and the relay. Also the relay should optimize its constellation when forwarding the detected symbols to the destination.

In order to exploit the full benefit of BICM in the context of cooperative communication, with the proposed framework different BICM modules can be optimized jointly considering the average SNRs of the source-destination channel and the two-hop relay channels. As such the full benefit of BICM can be exploited in the context of cooperative communication.

1.3 Summary

This thesis addresses several topics in design and performance analysis of BICM-based transmission systems. The main contributions are divided into three chapters. In Chapter 2 we develop an analytical framework for performance evaluation of generalized BICM transmission over AWGN channels that can be used to predict and optimize the performance of such systems. In Chapter 3, we analyze the performance of BICM-T transceiver over non-Gaussian channels. In Chapter 4, we proposed a new BICM-based system for cooperative communication where different BICM modules can be optimized jointly considering the average signal to noise ratios of the direct and the two-hop Rayleigh fading channels. Moreover, in Chapter 5, the

summary of contributions and proposals for further research are offered.

Chapter 2

Generalized and Optimal BICM Transceivers for AWGN Channels[1]

2.1 Introduction

In traditional BICM [3, 4], the channel encoder is connected to the modulator via a bit-level interleaver and at the receiver's side, reliability metrics for the coded bits (L-values) are calculated by the demapper which are then de-interleaved and fed to the binary decoder. The presence of a random interleaver in BICM allows the designer a flexibility to independently choose the code rate and the modulation order. Due to its advantages in fading channels over other CM schemes e.g., TCM [2] and multilevel coding [21], BICM is a de-facto standard for many contemporary wireless systems , e.g.,

HSPA, IEEE 802.11a/g, IEEE 802.16, and DVB-S2. In non-fading channels, compared with TCM, BICM gives a smaller minimum Euclidean distance, and also a smaller constraint capacity [4]. However this capacity loss is small when a constellation with Gray labelling is used. As such BICM is still considered as a valid option for CM over non-fading channels. Moreover, due to the random interleaver, the performance analysis of BICM is simplified.

Recent results [9, 10] have shown that the performance of BICM using convolutional codes in non-fading channels can be improved significantly when interleaver is removed from the BICM design. This design is referred to as BICM-T which has been formally analyzed in [9] and is shown to be asymptotically as good as Ungerboeck's one dimensional (1D)-TCM. The authors in [9] considered the BICM-T design and analysis for a simple case where a code rate of 1/2 with 16-ary QAM constellation is used. For this elementary example, the authors showed that the BER performance of BICM-T can be improved significantly over the traditional BICM design in non-fading channels. However, due to the use of ES signal constellation in the design, where the distance between neighbouring signal points is equal, it leads to a suboptimal BICM-T transceiver as the BER performance can be further improved by providing proper UEP to the coded bits [11, 12]. Moreover, design and analysis of a more general setup, where the number of outputs from the encoder does not match with the number of inputs to the modulator, are not considered. For such a generalized setup, the important question is how to connect the channel encoder to the modulator in order to minimize the BER performance.

In [12], the performance of BICM transceiver is optimized by using hi-

9

erarchical constellations, a bit-level multiplexer and multiple interleavers. Motivated by the results presented in [12], in this research we consider[2] an innovative BICM-T transceiver by using a bit-level multiplexer and hierarchical constellations. Using our considered more generalized design, the designer can not only exploit different degrees of freedom to notably improve the BER performance of BICM-T design but also enjoys the same design flexibility as the traditional BICM to independently choose the code rate and the modulation order. This design flexibility makes the generalized BICM-T design appealing compared to other CM for easy rate adaptation in frequency selective multicarrier non-fading channel where different carriers can have different channel qualities, (see for example [22]) and quasi-static fading channel where the channel remains constant over the length of a data frame, and varies independently from frame to frame (see for example [23]).

The performance of the considered BICM-T transceiver is analyzed. In BICM-T, a number of consecutive coded bits are transmitted using one symbol (due to the absence of random interleaver) and the corresponding likelihood values (L-values) are mutually dependent. Therefore, a different particularization as in [9] is adopted to calculate the L-values at the receiver in terms of constellation distance parameters. In particular, using a zero-crossing model introduced in [24], we develop closed form expressions for the probability density function (PDF) of L-values in terms of constellation distance parameters of the hierarchical quadrature amplitude modulation

---

[2]The presented design combines the well known techniques in [9, 12]. However, to the best of our knowledge, they have not been analyzed and optimized in this combination for a generalized scenario and the non-Gaussian noise considered in this thesis.

(HQAM). These PDFs are used to develop the UBs[3] on the BER of the system for additive white Gaussian noise (AWGN) channels, which are then used to optimize the constellation parameters as well as the multiplexer. Presented results for 64-ary QAM with rate 1/3 code shows that the considered design can offer gains up to 2.5 dB over the optimal BICM design for a target bit error rate (BER) of 10^{-6}.

Notations: All through this chapter, boldface letters $\boldsymbol{b}_t = [b_{1,t}, \ldots, b_{N,t}]$ are used to denote row vectors of length N and capital boldface letters $\boldsymbol{B} = [\boldsymbol{b}_1^{\mathrm{t}}, \ldots, \boldsymbol{b}_M^{\mathrm{t}}]^{\mathrm{t}}$ to denote a matrix of M rows, where $(\cdot)^{\mathrm{t}}$ represents transposition. The total Hamming weight of a binary matrix $\boldsymbol{B}$ is denoted by $d_{\mathrm{H}}(\boldsymbol{B})$. Random variables are represented by capital letters B. Probability is denoted by $\mathrm{Pr}(\cdot)$ and the PDF of a random variable Z by $p_Z(z)$. The convolution between two PDFs is represented by $p_{Z_1}(z) * p_{Z_2}(z)$ and $p_Z(z)^{*m}$ represents the m-fold self convolution of the PDF $p_Z(z)$. The Gaussian function with mean μ and variance σ^2 is defined as $\psi(z; \mu, \sigma) \triangleq \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right)$, and the Q-function as $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-u^2/2\right) du$.

2.2 Proposed BICM-T Transceiver

The model of our considered HQAM-BICM-T system is shown in Fig. 2.1. Next we describe operation of different blocks of this transmission scheme.

---

[3]All the UBs developed in the thesis are approximated UBs because they are developed using the approximated PDFs o L-values.

2.2.1 Encoder and Multiplexer

The k_c vectors of data bits $\boldsymbol{u}_j = [u_{j,1}, \ldots, u_{j,N_c}]$ with $j = 1, \ldots, k_c$ are encoded by a convolutional encoder (ENC) of rate $R = k_c/n$ to yield the vectors of coded bits $\hat{\boldsymbol{c}}_l = [\hat{c}_{l,1}, \ldots, \hat{c}_{l,N_c}]$ with $l = 1, \ldots, n$. A bit level multiplexing unit (MUX) bijectively maps $\hat{\boldsymbol{C}} = [\hat{\boldsymbol{c}}_1^t, \ldots, \hat{\boldsymbol{c}}_n^t]^t$ from the encoder output onto $\boldsymbol{C} = [\boldsymbol{c}_1^t, \ldots, \boldsymbol{c}_q^t]^t$ with $\boldsymbol{c}_k = [c_{k,1}, \ldots, c_{k,N_s}]$ and $k = 1, \ldots, q$. Without losing generality, we suppose that $nN_c = qN_s$. The k-th MUX output is linked to the k-th bit position of a modulator which maps the multiplexed coded bits $\boldsymbol{c}_k$ onto symbols using M-ary HQAM constellations labelled by the binary reflected Gray code (BRGC) [25], where $q = \log_2 M$.

In general, the MUX is defined as a one-to-one mapping between the blocks of nN_c and qN_s bits, i.e., $\{0,1\}^{n \times N_c} \leftrightarrow \{0,1\}^{q \times N_s}$. We define it by using a $n \times N_c$ matrix $\tilde{\boldsymbol{K}}$, as in [12], with $k \in \{1, \ldots, q\}$ and $t \in \{1, \ldots, N_s\}$, indicating that the bit $\hat{c}_{l,t'}$ is assigned to the k-th MUX's output (i.e., k-th bit position in the modulator) at time instant t, i.e., $c_{k,t} = \hat{c}_{l,t'}$. This definition of the MUX is completely general but hard to deal with, and therefore, we only consider MUX configurations that function periodically over blocks of nJ bits. Hence, we write $\tilde{\boldsymbol{K}}$ as a concatenation of N_c/J matrices $\boldsymbol{K}_\tau$, each having dimensions $n \times J$ and $\tau = 0, \ldots, N_c/J - 1$, i.e., $\tilde{\boldsymbol{K}} = [\boldsymbol{K}_0, \ldots, \boldsymbol{K}_{N_c/J-1}]$, where J is the period of MUX. The elements of $\boldsymbol{K}_\tau$ are $(k, t + \tau nJ/q)$ where $t \in \{1, \ldots, nJ/q\}$ and $k \in \{1, \ldots, q\}$. Without losing generality, we suppose that $(N_c \bmod J) = 0$ and that $(nJ \bmod q) = 0$. Consider the following example to clarify these definitions.

Example 2.1. Suppose $k_c = 1$ and $J = 3$. Consider a 8-ary constellation

Figure 2.1: System model of the BICM-T transceiver. A channel encoder is followed by the multiplexer (MUX), hierarchical M-PAM modulator, AWGN channel, hierarchical M-PAM de-modulator, demultiplexer (DEMUX), and decoder at the receiver side.

$(q = 3)$ and rate $R = 1/2$ $(n = 2)$ encoder. For this case, one possible MUX configuration can be

$$\boldsymbol{K}_\tau = \begin{bmatrix} (1, 1 + 2\tau) & (2, 1 + 2\tau) & (3, 1 + 2\tau) \\ (3, 2 + 2\tau) & (2, 2 + 2\tau) & (1, 2 + 2\tau) \end{bmatrix}, \tag{2.1}$$

which results in the following MUX

$$\tilde{\boldsymbol{K}} = \begin{bmatrix} (1, 1) & (2, 1) & (3, 1) & \bigg| & (1, 3) & (2, 3) & (3, 3) & \bigg| & \cdots \\ (3, 2) & (2, 2) & (1, 2) & \bigg| & (3, 4) & (2, 4) & (1, 4) & \bigg| & \cdots \end{bmatrix}.$$

The one-to-one mapping between $\hat{\boldsymbol{C}}$ and $\boldsymbol{C}$ is then

$$\hat{\boldsymbol{C}} = \begin{bmatrix} \hat{c}_{1,1} & \hat{c}_{1,2} & \hat{c}_{1,3} & \bigg| & \hat{c}_{1,4} & \hat{c}_{1,5} & \hat{c}_{1,6} & \bigg| & \cdots \\ \hat{c}_{2,1} & \hat{c}_{2,2} & \hat{c}_{2,3} & \bigg| & \hat{c}_{2,4} & \hat{c}_{2,5} & \hat{c}_{2,6} & \bigg| & \cdots \end{bmatrix}$$

$$\Longleftrightarrow \boldsymbol{C} = \begin{bmatrix} c_{1,1} & c_{2,3} & \bigg| & c_{1,4} & c_{2,6} & \bigg| & \cdots \\ c_{1,2} & c_{2,2} & \bigg| & c_{1,5} & c_{2,5} & \bigg| & \cdots \\ c_{1,3} & c_{2,1} & \bigg| & c_{1,6} & c_{2,4} & \bigg| & \cdots \end{bmatrix}. \tag{2.2}$$

Since $\boldsymbol{K}_\tau$ matrix is a permutation of the set $\{1, \ldots, nJ/q\} \times \{1, \ldots, q\}$, one can create $(nJ)!$ different matrices $\boldsymbol{K}_\tau$. However, this number can be reduced since trivial operations can be applied to $\boldsymbol{K}_\tau$ without affecting the system performance as explained in [12].

We note that while the matrices resulting from trivial operations can be abandoned, the sole way to find the optimal matrix of the available non-trivial matrices $\boldsymbol{K}_\tau$, for a given modulation order and code rate, is through the BER-criterion we develop in Section 2.5. To ease the notation, for the rest of this chapter we will refer to the matrix $\boldsymbol{K}_\tau$ as $\boldsymbol{K}$.

2.2.2 H-PAM Constellations

We consider HQAM constellations with BRGC labelling presented in [26]. In such constellations, each symbol can be represented by superposition of independent real/imaginary parts, so we consider an equivalent M-ary hierarchical pulse amplitude modulation (HPAM) constellation. The multiplexed coded bits $\boldsymbol{c}_{k,t}$ from the MUX output at any time instant t are mapped to a HPAM symbol $x_t \in \mathcal{X} = \{x_{t,0}, \ldots, x_{t,M-1}\}$ using a memoryless mapping $\mathcal{M} : \{0,1\}^q \to \mathcal{X}$.

The HPAM constellations (for example see Fig. 2.2 for 8-HPAM) are defined by the distances d_k with $k = 1, \ldots, q$. In this figure, black circles represent the M signal constellation points and the triangles/squares are "virtual" symbols that assist the understanding of HPAM signal constellation. The bit position of the binary labelling is denoted by $k = 1, \ldots, q$, where $k = 1$ represents the most significant bit. At first, the bit value for the position $(k = 1)$ chooses one of the squares in Fig. 2.2. Then the bit
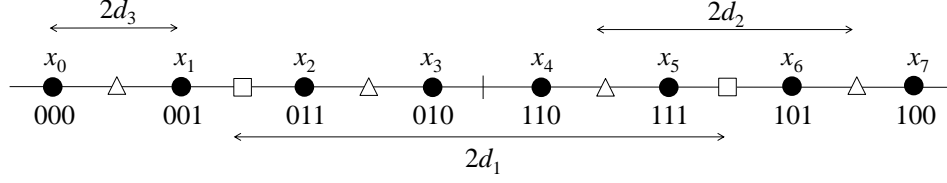
Figure 2.2: HPAM ($M = 8$) constellation with BRGC labelling.

value for the position ($k = 2$) chooses one of the two triangles surrounding the previously selected square. Finally, given the bit values for first two bit positions, the bit value for the least significant bit ($k = 3$) selects one of the two black symbols surrounding the triangle selected previously which is then transmitted to the receiver.

We use a vector $\boldsymbol{b}(m) = [b_1(m), \ldots, b_q(m)]$ to represent the binary representation of integers $0 \leq m \leq M - 1$, where the most significant bit of m is $b_1(m)$ and the least significant bit is $b_q(m)$. Therefore we can write the HPAM constellation elements $x_m \in \mathcal{X}$ as

$$x_m = \sum_{k=1}^{q} (-1)^{b_k(m)-1} d_k. \tag{2.3}$$

The normalized signal constellation parameters are defined as

$$\alpha_k \triangleq \frac{d_{k+1}}{d_1}, \tag{2.4}$$

with $k = 1, \ldots, q - 1$. The average symbol energy for equiprobable symbol transmission using (2.3) is given by $E_{\mathrm{s}} = d_1^2(1 + \sum_{k=1}^{q-1} \alpha_k^2)$ with α_k is defined by (2.4). All through this thesis, the constellation is considered to have normalized unity energy, which results in the following expression

15

$d_1 = (\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_{q-1}^2 + 1)^{-1/2}$. The following extra constraints on the values of α_k must be added to restrict the HPAM constellation to BRGC [12],

$$\alpha_k \geq \sum_{j=k+1}^{q-1} \alpha_j, \quad \sum_{k=1}^{q-1} \alpha_k \leq 1, \quad \alpha_{q-1} \geq 0, \tag{2.5}$$

with $k = 1, \ldots, q-1$.

An inspection of the BRGC for HPAM (Fig. 2.2) shows that, depending on the bit position, BRGC offers UEP [11] to the transmitted bits. In particular, the bit at the most significant bit position $(k = 1)$ receives the highest level of protection whereas the bit at the least significant bit position $(k = q)$ receives the lowest level of protection. Moreover, a bit labeled by one for bit position $(k = 2)$ in Fig. 2.2 (inner constellation points) will receive less protection than the bit labeled by zero transmitted in the same bit position (outer constellation points). Therefore, for $k = 2$, the binary-input soft-output (BISO) channel is asymmetric. To simplify the analysis, we use a similar approach to "symmetrize" the channel as used in [9]. The bits $\boldsymbol{c}_k$ from the MUX output are randomly inverted before mapping to the HPAM symbol, i.e., $\tilde{\boldsymbol{c}} = \boldsymbol{c} \oplus \boldsymbol{s} = [\tilde{\boldsymbol{c}}_1, \ldots, \tilde{\boldsymbol{c}}_{N_s}]$, where the entries of the matrix $\boldsymbol{s} = [\boldsymbol{s}_1^t, \ldots, \boldsymbol{s}_{N_s}^t] \in \{0,1\}^{q \times N_s}$, with $\boldsymbol{s}_t = [s_{1,t}, \ldots, s_{q,t}]$, are random vectors of bits and $\oplus$ is modulo-2 element-wise addition. Such a scrambling symmetrizes the BISO channel but it does not eliminate the UEP. We introduce this scrambling to facilitate the analysis, hence, it is not included in the BICM-T transceiver (in Fig. 2.1) nor used for simulations. We will see in Section 2.5 that the UB developed by using such symmetrization matches perfectly with

16

the numerical simulations.

The real part of the signal received from the source is $y_t = x_t + n_t$ where n_t is an AWGN noise with variance $N_0/2$, x_t represents the transmitted symbol, and $x_t, y_t, n_t \in \Re$. Since we have unit average transmitted symbol energy, SNR is given by $\gamma \triangleq E_s/N_0 = 1/N_0$. The receiver (REC) computes the L-values using the received signal which for the bit position k, is given by [9],

$$\tilde{l}_{k,t} = \log \frac{p_{Y_t}(y_t|\tilde{C}_{k,t} = 1)}{p_{Y_t}(y_t|\tilde{C}_{k,t} = 0)}, \tag{2.6}$$

from which we have,

$$p_{Y_t}(y_t|\tilde{c}_{k,t} = u) = \frac{\exp(u\tilde{l}_{k,t})}{1 + \exp(\tilde{l}_{k,t})}, \tag{2.7}$$

with $u \in \{0, 1\}$. Since $\tilde{\boldsymbol{c}}_t = \boldsymbol{c}_t \oplus \boldsymbol{s}_t$, we can write

$$l_{k,t} = (-1)^{s_{k,t}} \tilde{l}_{k,t}, \tag{2.8}$$

i.e., to reverse the scrambling, the sign of the L-values is altered by $(-1)^{s_{k,t}}$. These L-values are reorganized by the demultiplexer unit (DEMUX) which performs the reverse operation done by the MUX. The resulting vectors of L-values are sent to a soft input Viterbi decoder [27] which estimates the transmitted information bits from the source.

2.2.3 Decoder and the Decoding Errors

The MUX determines the correspondence between the encoder's output and the modulator's input. A binary codeword from the encoder's output is converted to an equivalent codeword at the MUX's output i.e., the modulator's input. Therefore we define the combination of the ENC and the MUX as an equivalent encoder (cf. Fig. 2.1). Similarly on the receiver side, the DEMUX and the DEC make an equivalent decoder (cf. Fig. 2.1). Using the vector of channel observations $\boldsymbol{y} = [y_1, \ldots, y_{N_s}]$ at the receiver, the equivalent decoder chooses the most likely sequence $\boldsymbol{c'}$ of the multiplexed coded bits $\boldsymbol{c}$ using a bit level metric [9], i.e.,

$$\boldsymbol{c'} = \underset{\bar{\boldsymbol{c}} \in \mathcal{C}_{eq}}{\operatorname{argmax}} \left\{ \log \prod_{t=1}^{N_s} p_{Y_t}(y_t | C_{1,t} = \bar{c}_{1,t}) \cdots p_{Y_t}(y_t | C_{q,t} = \bar{c}_{q,t}) \right\} \qquad (2.9)$$

$$= \underset{\bar{\boldsymbol{c}} \in \mathcal{C}_{eq}}{\operatorname{argmax}} \left\{ \sum_{t=1}^{N_s} \sum_{k=1}^{q} \log p_{Y_t}(y_t | C_{k,t} = \bar{c}_{k,t}) \right\}, \qquad (2.10)$$

where $\mathcal{C}_{eq}$ is the equivalent set of binary codes. Using a similar expression to (2.7), we can write (2.10) as

$$\boldsymbol{c'} = \underset{\bar{\boldsymbol{c}} \in \mathcal{C}_{eq}}{\operatorname{argmax}} \left\{ \sum_{t=1}^{N_s} \sum_{k=1}^{q} \bar{c}_{k,t} l_{k,t} - \sum_{t=1}^{N_s} \sum_{k=1}^{q} \log(1 + \exp(l_{k,t})) \right\} \qquad (2.11)$$

$$= \underset{\bar{\boldsymbol{c}} \in \mathcal{C}_{eq}}{\operatorname{argmax}} \left\{ \sum_{t=1}^{N_s} \sum_{k=1}^{q} \bar{c}_{k,t} l_{k,t} \right\}, \qquad (2.12)$$

where the second term in (2.12) is not relevant to the decoder's decision because it is independent of $\bar{\boldsymbol{c}}$.

The decoder makes a false decision based on (2.12) if, instead of the

transmitted equivalent codeword $\boldsymbol{c}$, it detects a codeword $\boldsymbol{c}'$. The probability of this event is so called pairwise error probability (PEP) and given by

$$
\mathrm{PEP}(\boldsymbol{c} \to \boldsymbol{c}') \triangleq \mathrm{Pr} \left\{ \sum_{t=1}^{N_{\mathrm{s}}} (c'_{1,t} l_{1,t} + \cdots + c'_{q,t} l_{q,t}) \right.
$$

$$
\left. \geq \sum_{t=1}^{N_{\mathrm{s}}} (c_{1,t} l_{1,t} + \cdots + c_{q,t} l_{q,t}) \right\} \tag{2.13}
$$

$$
= \mathrm{Pr} \left\{ \sum_{t=1}^{N_{\mathrm{s}}} (e_{1,t} l_{1,t} + \cdots + e_{q,t} l_{q,t}) \geq 0 \right\} \tag{2.14}
$$

where we represent elements of the "error" codeword $(\boldsymbol{e} = \boldsymbol{c}' - \boldsymbol{c})$ as $e_{k,t}$.

We note that q consecutive multiplexed coded bits for each time t are transmitted using same symbol x_t over same channel and noise realization. Therefore, in the PEP expression (2.14), each sequence $\boldsymbol{l}_1, \boldsymbol{l}_2 \cdots, \boldsymbol{l}_{N_{\mathrm{s}}}$ contains q dependent L-values. As such, a different approach as used in [9] must be adopted to calculate the PEP. In the next section, we describe how to evaluate the performance of our considered BICM-T design.

2.3 Performance Evaluation

As the channel is symmetric due to scrambling, we assume that the source transmitted a codeword with all zeros. Let $\mathcal{E}$ be the set of codewords corresponding to the paths of the code diverging at time instant t from the zero-state, and remerging after T stages. These corresponding equivalent codewords are denoted as $\boldsymbol{e} \triangleq [\boldsymbol{e}_1^{\mathrm{t}}, \ldots, \boldsymbol{e}_T^{\mathrm{t}}]$ where $\boldsymbol{e}_t = [e_{1,t}, \ldots, e_{q,t}]$. Let the metric related to the equivalent codeword $\boldsymbol{e}$ is $\Lambda_{\boldsymbol{e}}$ which is a sum of

independent random variables, i.e.,

$$\Lambda_{\boldsymbol{e}} \triangleq \Lambda^{(t)} + \Lambda^{(t+1)} + \Lambda^{(t+2)} + \dots, \tag{2.15}$$

where $\Lambda^{(t)} = \sum_{k=1}^{q} e_{k,t} l_{k,t}$. Therefore, we can write,

$$\Lambda^{(t)} \equiv \Lambda(\boldsymbol{e}_t, \boldsymbol{s}_t) = \sum_{k=1}^{q} (-1)^{s_{k,t}} e_{k,t} \tilde{l}_{k,t}. \tag{2.16}$$

Since $\tilde{l}_{k,t}$ are random variables, for the relevant cases in (2.16), we need $(M-1)$ PDFs $p_{\lambda_{\pi_i}}(\lambda|\boldsymbol{C}_{\pi_i,t})$ where $\pi_i \in \Pi$, $i = 1, \dots, M-1$ and Π is the set containing all possible subsets of bit indices $1, \dots, q$. The vector of bits whose indices belong to π_i is given by $\boldsymbol{C}_{\pi_i,t}$. We note that $p_{\lambda_{\pi_i}}(\lambda|\boldsymbol{C}_{\pi_i,t})$ is not always conditioned on a single bit as for traditional single interleaver BICM designs [3, 4]. Infact for different values of π_i, $p_{\lambda_{\pi_i}}(\lambda|\boldsymbol{C}_{\pi_i,t})$ is conditioned on $1, 2 \dots$ or q bits. To clarify these notations, consider the following example.

Example 2.2. Assuming 4-ary constellation ($q = 2$) and using (2.16), we can express the metrics as

$$\Lambda^t \equiv \Lambda(\boldsymbol{e}_t, \boldsymbol{s}_t) = \begin{cases} 0, & \text{if } \boldsymbol{e}_t = [0,0] \\ (-1)^{s_{1,t}} \tilde{l}_{1,t}, & \text{if } \boldsymbol{e}_t = [1,0] \\ (-1)^{s_{2,t}} \tilde{l}_{2,t}, & \text{if } \boldsymbol{e}_t = [0,1] \\ \sum_{k=1}^{2} (-1)^{s_{k,t}} \tilde{l}_{k,t}, & \text{if } \boldsymbol{e}_t = [1,1] \end{cases}. \tag{2.17}$$

For this example $\Pi = \{1, 2, (1,2)\}$ and for the cases in (2.17), we need three PDFs i.e., $p_{\lambda_1}(\lambda|C_1)$, $p_{\lambda_2}(\lambda|C_2)$ and $p_{\lambda_{1,2}}(\lambda|\boldsymbol{C}_{1,2})$. We note that the

conditional PDF $p_{\lambda_{1,2}}(\lambda|\boldsymbol{C}_{1,2})$ is conditioned on pair of bits $\boldsymbol{C}_{1,2} = [C_1, C_2]$ rather than a single bit.

From (2.15), and due to the independence of each L-value $\Lambda(\boldsymbol{e}_t, \boldsymbol{s}_t)$ in (2.16), the PEP can be expressed as,[4]

$$\mathrm{PEP}(\boldsymbol{e}) = \mathrm{PEP}(w_{\boldsymbol{e},\Pi}) = \int_0^\infty p_{\Lambda_{\pi_1}}(\lambda|\boldsymbol{C}_{\pi_1} = \boldsymbol{0})^{*w_{\boldsymbol{e},\pi_1}} \dots$$

$$p_{\Lambda_{\pi_{M-1}}}(\lambda|\boldsymbol{C}_{\pi_{M-1}} = \boldsymbol{0})^{*w_{\boldsymbol{e},\pi_{M-1}}} d\lambda, \qquad (2.18)$$

where, $w_{\boldsymbol{e},\pi_i}$ represent the number of columns in $\boldsymbol{e}$ where the bit positions with indices belonging to π_i have bit value equal to 1. Then, the UB can be expressed as [9],

$$\mathrm{UB} = \frac{1}{k_c} \sum_{w_\Pi} \mathrm{PEP}(w_\Pi) \sum_{\boldsymbol{e} \in \mathcal{C}_{w_\Pi}} d_{\mathrm{H}}(\boldsymbol{u_e})$$

$$= \frac{1}{k_c} \sum_{w_\Pi} \mathrm{PEP}(w_\Pi) \beta^{\mathbb{C}}_{w_\Pi, \boldsymbol{K}}, \qquad (2.19)$$

where $\mathcal{C}_{w_\Pi} \triangleq \{\boldsymbol{e} \in \mathcal{E} : w_{\pi_i} = w_{\boldsymbol{e},\pi_i}\}$, $d_{\mathrm{H}}(\boldsymbol{u_e})$ is the Hamming weight of the input sequence $\boldsymbol{u_e}$ associated with the equivalent codeword $\boldsymbol{e}$ and $\beta^{\mathbb{C}}_{w_\Pi, \boldsymbol{K}} = \sum_{\boldsymbol{e} \in \mathcal{C}_{w_\Pi}} d_{\mathrm{H}}(\boldsymbol{u_e})$ is an equivalent weight distribution spectrum (EWDS) of the code $\mathbb{C}$. In what follows, we describe this EWDS of the code.

EWDS of the code $\mathbb{C}$ considers the generalized weight $[w_{\pi_1}, \dots, w_{\pi_{M-1}}]$ of codewords and MUX configuration matrix $\boldsymbol{K}$. The correspondence be-

---

[4]For notational convenience, we used $\boldsymbol{C}_{\pi_i} = \boldsymbol{0}$, where $\boldsymbol{0}$ is a vector of same size as $\boldsymbol{C}_{\pi_i}$

tween the encoder's output $\hat{C}$ and MUX output i.e., modulator's input C is determined by the matrix K, and by the instantaneous time at which the sequence diverge from all-zero state. However, only J time instants must be considered because of the periodic structure of $\tilde{K}$. Therefore, we can express the EWDS as

$$\beta_{w_\Pi,K}^{\mathbb{C}} = \frac{1}{J} \sum_{n=1}^{J} \beta_{w_\Pi,K}^{\mathbb{C}(n)}, \tag{2.20}$$

where the EWDS is given by $\beta_{w_\Pi,K}^{\mathbb{C}(n)}$ when the decoder diverges at time $t+n$ with arbitrary t.

Example 2.3 (EWDS of the code $(5,7)_8$). Consider the constraint length $K = 3$ convolutional code with polynomial generators $(5,7)_8$. The free distance of the code is $d_{\mathrm{H}}^{\mathrm{free}} = 5$, and $\beta_5^{\mathbb{C}} = 1$, i.e., there is one divergent path at Hamming distance five from the all-zero codeword, and the input Hamming weight of the path is one. Consider the MUX in Example 1 with period $J = 3$ and 8-ary $(q = 3)$ constellation. For this case $\Pi = \{1, 2, 3, (1, 2), (2, 3), (1, 3), (1, 2, 3)\}$. The $J = 3$ possible input sequences which have the Hamming weight one are $\boldsymbol{u_1}^{(1)} = [\ldots, 0, 1, 0, 0, 0, 0, \ldots]$, $\boldsymbol{u_1}^{(2)} = [\ldots, 0, 0, 1, 0, 0, 0, \ldots]$ and $\boldsymbol{u_1}^{(3)} = [\ldots, 0, 0, 0, 1, 0, 0, \ldots]$, which results in the following $\hat{C}$ matrices,

$$\hat{C}^{(1)} = \begin{bmatrix} \ldots 1 & 0 & 1 & 0 & 0 \ldots \\ \ldots 1 & 1 & 1 & 0 & 0 \ldots \end{bmatrix},$$

$$\hat{\boldsymbol{C}}^{(2)} = \begin{bmatrix} \dots 0 & 1 & 0 & 1 & 0 \dots \\ \dots 0 & 1 & 1 & 1 & 0 \dots \end{bmatrix},$$

$$\hat{\boldsymbol{C}}^{(3)} = \begin{bmatrix} \dots 0 & 0 & 1 & 0 & 1 \dots \\ \dots 0 & 0 & 1 & 1 & 1 \dots \end{bmatrix}, \tag{2.21}$$

which by using (1) yield

$$\boldsymbol{C}^{(1)} = \begin{bmatrix} \dots 1 & 1 & 0 & 0 \dots \\ \dots 0 & 1 & 0 & 0 \dots \\ \dots 1 & 1 & 0 & 0 \dots \end{bmatrix},$$

$$\boldsymbol{C}^{(2)} = \begin{bmatrix} \dots 0 & 1 & 1 & 1 \dots \\ \dots 1 & 1 & 0 & 0 \dots \\ \dots 0 & 0 & 0 & 0 \dots \end{bmatrix},$$

$$\boldsymbol{C}^{(3)} = \begin{bmatrix} \dots 0 & 1 & 0 & 1 \dots \\ \dots 0 & 0 & 1 & 1 \dots \\ \dots 1 & 0 & 0 & 0 \dots \end{bmatrix}. \tag{2.22}$$

If we consider only this error event at minimum Hamming distance, the final EWDS given by (2.20) is obtained by computing the number of columns of $\boldsymbol{C}^{(n)}$ with $n = 1, 2, 3$, where the bit positions with indices belonging to π_i, with $i = 1, 2, \cdots, 7$, have bit value equal to 1, i.e.,

$$\beta_{w_\Pi, \boldsymbol{K}}^{\mathbb{C}} = \begin{cases} \frac{1}{3}, & \text{if } w_\Pi = [0, 0, 0, 0, 0, 1, 1] \\ \frac{1}{3}, & \text{if } w_\Pi = [0, 1, 2, 1, 0, 0, 0] \\ \frac{1}{3}, & \text{if } w_\Pi = [1, 1, 1, 1, 0, 0, 0] \end{cases} .$$

23

The spectrum $\beta^{\mathbb{C}}_{w_{\Pi},\boldsymbol{K}}$ can be numerically calculated using a breadth first search algorithm [28]. Clearly, the spectrum must be truncated so that only diverging sequences with total Hamming weight $w_1 + \ldots + w_q \leq \hat{w}$ are considered.

2.4 PDF of the L-values

In order to calculate the PEP for BICM-T in (2.18), we need to compute the conditional PDFs $p_{\lambda_{\pi_i}}(\lambda|\boldsymbol{C}_{\pi_i,t})$ for $\pi_i \in \Pi$. In this section we develop the approximations for these PDFs using so called zero crossing model.

We can express the L-values in (2.6) as

$$\tilde{l}_{k,t}(y_t|\boldsymbol{s}_t) = \log \frac{\sum_{x\in\mathcal{X}_{k,1}} p_{Y_t}(y_t|X_t = x)}{\sum_{x\in\mathcal{X}_{k,0}} p_{Y_t}(y_t|X_t = x)}, \tag{2.23}$$

where $\mathcal{X}_{k,b}$ represents the symbols with bit value b at bit position k. This L-value can be approximated using max-log simplification [27] by

$$\tilde{l}_{k,t}(y_t|\boldsymbol{s}_t) \approx \gamma \left[\min_{x\in\mathcal{X}_{k,0}} (y_t - x)^2 - \min_{x\in\mathcal{X}_{k,1}} (y_t - x)^2 \right]. \tag{2.24}$$

The L-values depend on the transmitted symbol x_t, however, since a codeword with all zeros is transmitted ($[c_{1,t}, \ldots, c_{q,t}] = [0, \ldots, 0]$), x_t is completely determined by $\boldsymbol{s}_t$, i.e., $x_t = \mu(\boldsymbol{s}_t)$. As such, we use $\tilde{l}_{k,t}(y_t|\boldsymbol{s}_t)$ to emphasize that the L-values depend on the received signal y_t and the scrambling sequence $\boldsymbol{s}_t$. For a given transmitted symbol x_t, the received signal y_t is a Gaussian random variable with mean x_t and variance $N_0/2$. The L-value in (2.24) is a piece-wise linear function of y_t and is therefore piecewise
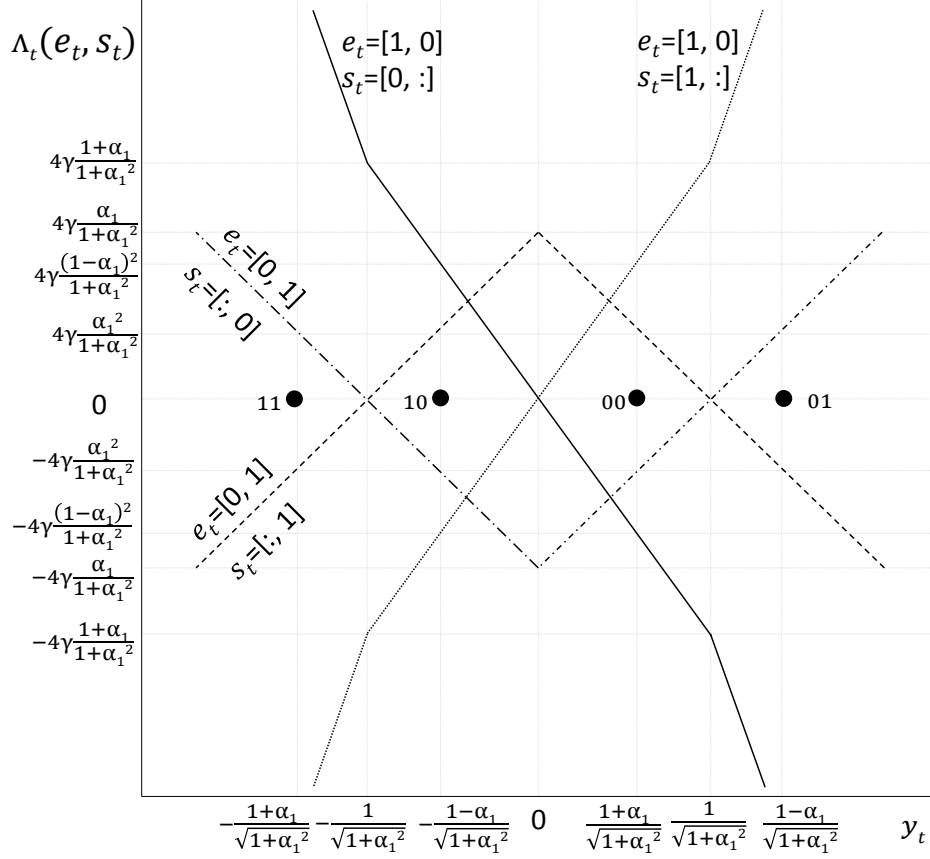
Figure 2.3: Piece-wise linear relation between the received signal y_t and L-values $\Lambda(e_t, s_t)$ in (2.16) for hierarchical 4-PAM and for different possible values of e_t and s_t. Figure shows the relation when $e_t = [1, 0]$ or $e_t = [0, 1]$. The symbols transmitted are shown with circles. For $e_t = [1, 0]$, the notation $s_t = [0/1, :]$ is used to represent that the L-values $\Lambda(e_t, s_t)$ are independent of $s_{2,t}$.

Gaussian. Hence, each L-value $\Lambda(e_t, s_t)$ in (2.16) is a sum of piece-wise Gaussian functions. The piecewise linear relationships are shown in Figures 2.3 and 2.4 for 4-HPAM. To facilitate the analysis, we use the zero-crossing approximation [9] of the L-values which replaces all the Gaussian pieces by one
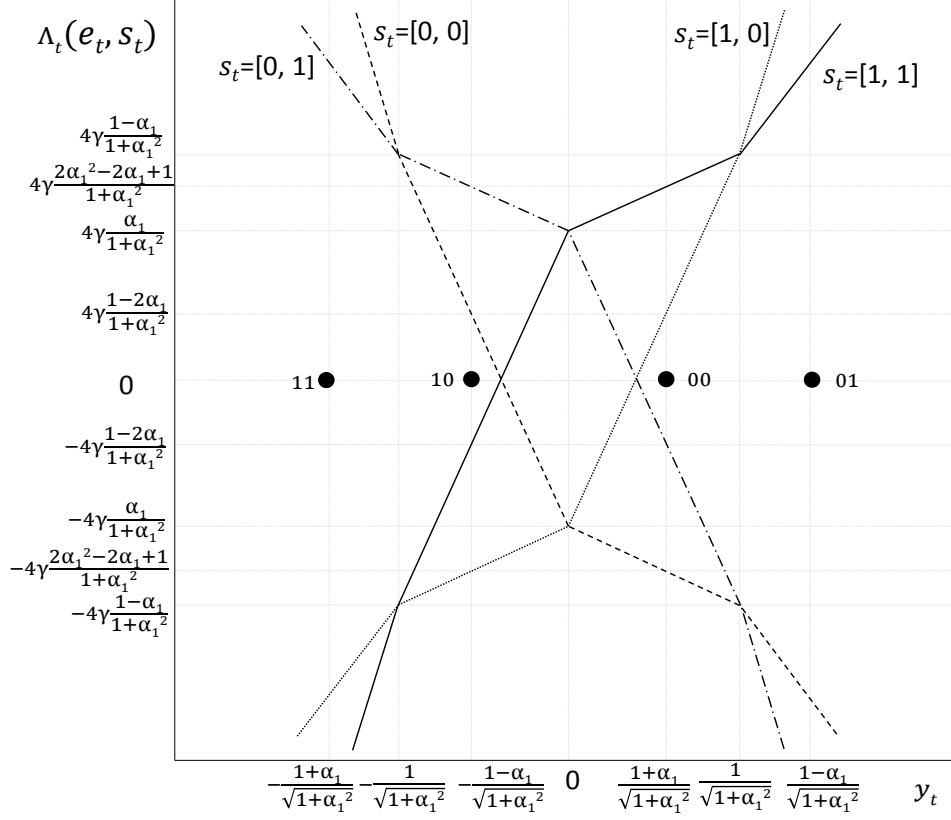
Figure 2.4: Piece-wise linear relation between the received signal y_t and L-values $\Lambda(\boldsymbol{e}_t, \boldsymbol{s}_t)$ in (2.16) for hierarchical 4-PAM and for different possible values of $\boldsymbol{e}_t$ and $\boldsymbol{s}_t$. Figure shows the relation when $\boldsymbol{e}_t = [1, 1]$. The symbols transmitted are shown with circles. For $\boldsymbol{e}_t = [1, 0]$, the notation $\boldsymbol{s}_t = [0/1, :]$ is used to represent that the L-values $\Lambda(\boldsymbol{e}_t, \boldsymbol{s}_t)$ are independent of $s_{2,t}$.

Gaussian function. According to this approximation,

$$\Lambda(y_t | \boldsymbol{e}_t, \boldsymbol{s}_t) \approx \hat{a}(\boldsymbol{e}_t, \boldsymbol{s}_t) y_t + \hat{b}(\boldsymbol{e}_t, \boldsymbol{s}_t), \qquad (2.25)$$

where $\hat{a}$ and $\hat{b}$ are determined by the linear piece crossing x-axis and closest to the symbol x_t. Using (2.25), we can model the conditional L-values as

Gaussian random variables whose mean and variance depend on γ, $\boldsymbol{e}_t$ and $\boldsymbol{s}_t$, i.e.,

$$p_\Lambda(\lambda | \boldsymbol{S}_t = \boldsymbol{s}_t) = \psi(\lambda; \hat{\mu}(\boldsymbol{e}_t, \boldsymbol{s}_t), \hat{\sigma}^2(\boldsymbol{e}_t, \boldsymbol{s}_t)), \tag{2.26}$$

where the mean value and variance are given by

$$\hat{\mu}(\boldsymbol{s}_t) = x_t \hat{a}(\boldsymbol{e}_t, \boldsymbol{s}_t) + \hat{b}(\boldsymbol{e}_t, \boldsymbol{s}_t) \tag{2.27}$$

$$\hat{\sigma}^2(\boldsymbol{e}_t, \boldsymbol{s}_t) = [\hat{a}(\boldsymbol{e}_t, \boldsymbol{s}_t)]^2 \frac{N_0}{2}. \tag{2.28}$$

This approximation strategy follows the recommendation of [29], where it is shown that the behavior of the PDF of L-values around $\Lambda = 0$ is important to accurately calculate the BER. To illustrate that the PDF of Λ in (2.16) is accurately approximated by the zero-crossing approximation around $\Lambda = 0$, we consider the following example,

Example 2.4. Let $\boldsymbol{e}_t = [1, 0]$ and $\boldsymbol{s}_t = [0, 0]$ in Example 2. Then, according to the zero-crossing approximation,

$$\Lambda(y_t | \boldsymbol{e}_t, \boldsymbol{s}_t) \approx -\frac{4\gamma(1 - \alpha_1)}{\sqrt{1 + \alpha_1^2}} y_t. \tag{2.29}$$

Using (2.29), we can find the mean and variance in (2.27) and (2.28) respectively. The PDF of the conditional L-values can then be approximated using (2.26) as,

$$p_\Lambda(\lambda | \boldsymbol{S}_t = [0, 0]) = \psi\left(\lambda; -\frac{4\gamma(1 - \alpha_1)^2}{1 + \alpha_1^2}, \frac{8\gamma(1 - \alpha_1)^2}{1 + \alpha_1^2}\right), \tag{2.30}$$

In Fig. 2.5, we plot the approximated PDF (2.30) as well as the exact PDF (obtained via simulations). We can note from this figure that around $\Lambda = 0$ the approximated PDF matches with the exact PDF.

The closed form expression for PDF of L-values in (2.16) for a general M-ary signal modulation is difficult to develop. Therefore, the analysis of the considered BICM-T design is done on case by case basis. Next, we develop the UB on BER of BICM-T system in terms of constellation distance parameters and MUX configuration for two particular modulation orders: H-16QAM and H-64QAM. The UB for other cases can be developed using
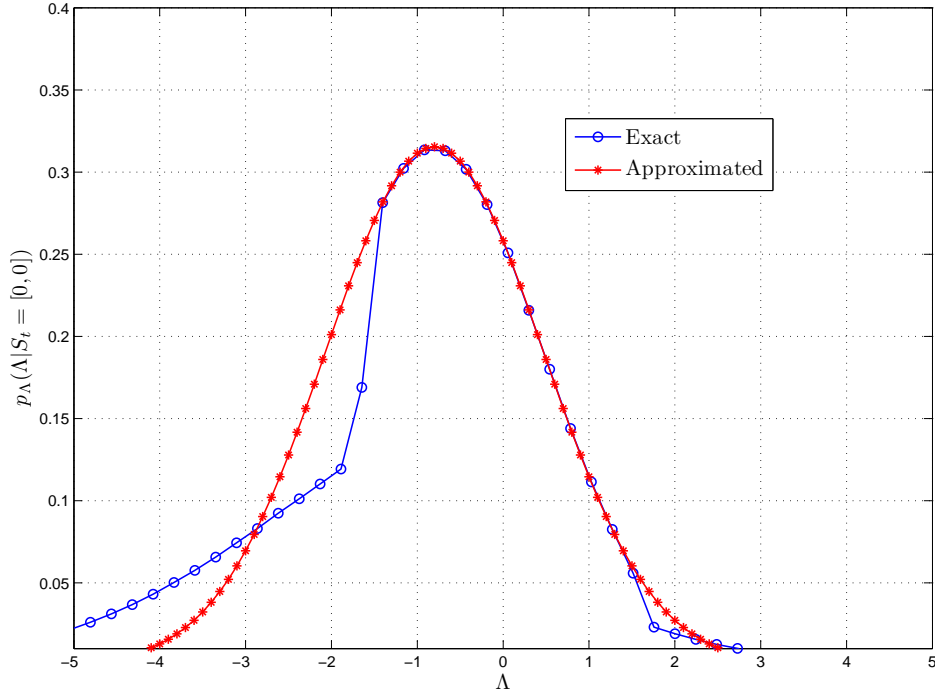
Figure 2.5: The exact and the approximated PDF of Λ given $\boldsymbol{e}_t = [1, 0]$ and $\boldsymbol{s}_t = [0, 0]$.

the same approach.

2.4.1 H-16QAM

Assume $k_c = 1$, $n = 2$, $J = 2$ and $q = 2$. For this particular case, there is a single signal constellation parameter, i.e., α_1. We need to compute three PDFs i.e., $p_{\Lambda_1}(\lambda|C_1)$, $p_{\Lambda_2}(\lambda|C_2)$ and $p_{\Lambda_{1,2}}(\lambda|\boldsymbol{C}_{1,2})$, as explained in Example 2.2, to calculate the PEP in (2.18). In Table 2.1, we present the values of $\hat{a}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ and $\hat{b}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ for this particular case. The means and variances are presented in Table 2.2 in terms of constellation parameter for the same cases of Table 2.1.

By averaging (2.26) over the scrambling sequence $\boldsymbol{s}_t$, we obtain the unconditional PDF of the L-values in (2.16) given by the following equation

$$
p_\Lambda(\lambda) = \begin{cases} \frac{1}{2}[\psi(\lambda; -\frac{1+\alpha_1}{1-\alpha_1}\rho, 2\rho) + \psi(\lambda; -\rho, 2\rho)], & \text{if } \boldsymbol{e}_t = [1, 0] \\ \psi(\lambda; -\frac{\alpha_1^2}{(1-\alpha_1)^2}\rho, 2\frac{\alpha_1^2}{(1-\alpha_1)^2}\rho), & \text{if } \boldsymbol{e}_t = [0, 1] \\ \psi(\lambda; -\frac{1}{(1-\alpha_1)^2}\rho, 2\frac{1}{(1-\alpha_1)^2}\rho), & \text{if } \boldsymbol{e}_t = [1, 1] \end{cases} \cdot \quad (2.31)
$$

where for notation simplicity we have defined,

$$
\rho = 4\gamma \frac{(1 - \alpha_1)^2}{1 + \alpha_1^2}. \quad (2.32)
$$

Using (2.31), elements in the integration in (2.18) can be written as,

$$p_{\Lambda_1}(\lambda|C_1 = 0)^{*w_1} = \left(\frac{1}{2}\right)^{w_1} \sum_{j=0}^{w_1} \binom{w_1}{j} \psi\left(\lambda; -\frac{2\alpha_1\rho}{1-\alpha_1}j + \rho w_1, 2\rho w_1\right)$$

(2.33)

$$p_{\Lambda_2}(\lambda|C_2 = 0)^{*w_2} = \psi\left(\lambda; -\frac{\alpha_1^2\rho}{(1-\alpha_1)^2}w_2, \frac{\alpha_1^2\rho}{(1-\alpha_1)^2}w_2\right) \tag{2.34}$$

$$p_{\Lambda_{1,2}}(\lambda|\mathbf{C_{1,2}} = \mathbf{0})^{*w_{1,2}} = \psi\left(\lambda; \frac{-\rho}{(1-\alpha_1)^2}w_{1,2}, \frac{\rho}{(1-\alpha_1)^2}w_{1,2}\right) \tag{2.35}$$

where we used $\psi(\lambda; \mu_1, \sigma_1^2)*\ldots*\psi(\lambda; \mu_N, \sigma_N^2) = \psi(\lambda; \sum_{n=1}^{N}\mu_n, \sum_{n=1}^{N}\sigma_n^2)$.

Using the relations (2.33)-(2.35) in (2.18) gives

$$\text{PEP}(w_1, w_2, w_{1,2}) = \left(\frac{1}{2}\right)^{w_1} \sum_{j=0}^{w_1} \binom{w_1}{j}.$$
$$\int_0^\infty \psi(\lambda; \mu_{1,2,(1,2),j}, \sigma_{1,2,(1,2)}^2)\mathrm{d}\lambda, \tag{2.36}$$

Table 2.1: Values of $\hat{a}(\mathbf{e}_t, \mathbf{s}_t)$ and $\hat{b}(\mathbf{e}_t, \mathbf{s}_t)$ for HPAM ($M = 4$), $\beta = (1 + \alpha_1^2)^{-1/2}$

	$\mathbf{s}_t = [1,1]$		$\mathbf{s}_t = [1,0]$		$\mathbf{s}_t = [0,0]$		$\mathbf{s}_t = [0,1]$	
	$\hat{a}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{b}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{a}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{b}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{a}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{b}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{a}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{b}(\mathbf{e}_t, \mathbf{s}_t)$
$\mathbf{e}_t = [1,0]$	$4\gamma\beta(1-\alpha_1)$	0	$4\gamma\beta(1-\alpha_1)$	0	$-4\gamma\beta(1-\alpha_1)$	0	$-4\gamma\beta(1-\alpha_1)$	0
$\mathbf{e}_t = [0,1]$	$4\gamma\beta\alpha_1$	$4\gamma\beta^2\alpha_1$	$-4\gamma\beta\alpha_1$	$-4\gamma\beta^2\alpha_1$	$4\gamma\beta\alpha_1$	$-4\gamma\beta^2\alpha_1$	$-4\gamma\beta\alpha_1$	$4\gamma\beta^2\alpha_1$
$\mathbf{e}_t = [1,1]$	$4\gamma\beta$	$4\gamma\beta^2\alpha_1$	$4\gamma\beta$	$-4\gamma\beta^2$	$-4\gamma\beta$	$-4\gamma\beta^2\alpha_1$	$-4\gamma\beta$	$4\gamma\beta^2$

Table 2.2: Values of $\hat{\mu}(\mathbf{e}_t, \mathbf{s}_t)$ and $\hat{\sigma}^2(\mathbf{e}_t, \mathbf{s}_t)$ for HPAM ($M = 4$), $\beta = (1 + \alpha_1^2)^{-1/2}$

	$\mathbf{s}_t = [1,1]$		$\mathbf{s}_t = [1,0]$		$\mathbf{s}_t = [0,0]$		$\mathbf{s}_t = [0,1]$	
	$\hat{\mu}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{\sigma}^2(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{\mu}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{\sigma}^2(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{\mu}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{\sigma}^2(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{\mu}(\mathbf{e}_t, \mathbf{s}_t)$	$\hat{\sigma}^2(\mathbf{e}_t, \mathbf{s}_t)$
$\mathbf{e}_t = [1,0]$	$-4\gamma\beta^2(1-\alpha_1^2)$	$8\gamma\beta^2(1-\alpha_1)^2$	$-4\gamma\beta^2(1-\alpha_1)^2$	$8\gamma\beta^2(1-\alpha_1)^2$	$-4\gamma\beta^2(1-\alpha_1)^2$	$8\gamma\beta^2(1-\alpha_1)^2$	$-4\gamma\beta^2(1-\alpha_1^2)$	$8\gamma\beta^2(1-\alpha_1)^2$
$\mathbf{e}_t = [0,1]$	$-4\gamma\beta^2\alpha_1^2$	$8\gamma\beta^2\alpha_1^2$	$-4\gamma\beta^2\alpha_1^2$	$8\gamma\beta^2\alpha_1^2$	$-4\gamma\beta^2\alpha_1^2$	$8\gamma\beta^2\alpha_1^2$	$-4\gamma\beta^2\alpha_1^2$	$8\gamma\beta^2\alpha_1^2$
$\mathbf{e}_t = [1,1]$	$-4\gamma\beta^2$	$8\gamma\beta^2$	$-4\gamma\beta^2$	$8\gamma\beta^2$	$-4\gamma\beta^2$	$8\gamma\beta^2$	$-4\gamma\beta^2$	$8\gamma\beta^2$

where

$$\mu_{1,2,(1,2),j} = - \left(w_1 + \frac{\alpha_1^2 w_2}{(1-\alpha_1)^2} + \frac{w_{1,2}}{(1-\alpha_1)^2} + \frac{2\alpha_1 j}{1-\alpha_1} \right) \rho, \tag{2.37}$$

$$\sigma^2_{1,2,(1,2)} = 2 \left(w_1 + \frac{\alpha_1^2}{(1-\alpha_1)^2} w_2 + \frac{1}{(1-\alpha_1)^2} w_{1,2} \right) \rho. \tag{2.38}$$

Using (2.37)-(2.38) in (2.36), and (2.19), we get the expression for UB as follows:

$$\mathrm{UB} = \frac{1}{k_c} \sum_{w_1,w_2,w_{1,2}} \beta^{\mathbb{C}}_{w_1,w_2,w_{1,2},\boldsymbol{K}} \left(\frac{1}{2} \right)^{w_1} \sum_{j=0}^{w_1} \binom{w_1}{j}.$$

$$Q \left(\sqrt{ \frac{(w_1 + \frac{\alpha_1^2}{(1-\alpha_1)^2} w_2 + \frac{w_{1,2}}{(1-\alpha_1)^2} + \frac{2j\alpha_1}{1-\alpha_1})^2}{w_1 + \frac{\alpha_1^2}{(1-\alpha_1)^2} w_2 + \frac{1}{(1-\alpha_1)^2} w_{1,2}} \frac{2\gamma(1-\alpha_1)^2)}{1+\alpha_1^2} } \right). \tag{2.39}$$

Note that for $\alpha_1 = 1/2$, (2.39) is identical to the UB expression developed

Table 2.3: Values of $\hat{a}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ and $\hat{b}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ for HPAM ($M = 8$) where $\varrho = \frac{1}{\sqrt{1+\alpha_1^2+\alpha_2^2}}$

		$\boldsymbol{e}_t=[1,0,0]$	$\boldsymbol{e}_t=[0,1,0]$	$\boldsymbol{e}_t=[0,0,1]$	$\boldsymbol{e}_t=[1,1,0]$	$\boldsymbol{e}_t=[0,1,1]$	$\boldsymbol{e}_t=[1,0,1]$	$\boldsymbol{e}_t=[1,1,1]$
$\boldsymbol{s}_t=[1,1,1]$	$\hat{a}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	$-4\gamma\varrho(1-\alpha_1-\alpha_2)$	$-4\gamma\varrho(\alpha_1-\alpha_2)$	$-4\gamma\varrho\alpha_2$	$-4\gamma\varrho(1-\alpha_2)$	$-4\gamma\varrho\alpha_1$	$-4\gamma\varrho(1-\alpha_1)$	$-4\gamma\varrho(1-2\alpha_2)$
	$\hat{b}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	0	$-4\gamma\varrho^2(\alpha_1-\alpha_2)$	$-4\gamma\varrho^2\alpha_2(1+\alpha_1)$	$-4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$-4\gamma\varrho^2\alpha_1(1+\alpha_2)$	$4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$-4\gamma\varrho^2(\alpha_1-\alpha_2)$
$\boldsymbol{s}_t=[1,1,0]$	$\hat{a}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	$-4\gamma\varrho(1-\alpha_1-\alpha_2)$	$-4\gamma\varrho(\alpha_1-\alpha_2)$	$4\gamma\varrho\alpha_2$	$-4\gamma\varrho(1-\alpha_2)$	$-4\gamma\varrho\alpha_1$	$-4\gamma\varrho(1-\alpha_1)$	$-4\gamma\varrho$
	$\hat{b}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	0	$-4\gamma\varrho^2(\alpha_1-\alpha_2)$	$4\gamma\varrho^2\alpha_2(1+\alpha_1)$	$-4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$-4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$-4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$-4\gamma\varrho^2(\alpha_1-2\alpha_1\alpha_2+\alpha_2)$
$\boldsymbol{s}_t=[1,0,0]$	$\hat{a}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	$-4\gamma\varrho(1-\alpha_1-\alpha_2)$	$4\gamma\varrho(\alpha_1-\alpha_2)$	$-4\gamma\varrho\alpha_2$	$-4\gamma\varrho(1-\alpha_2)$	$4\gamma\varrho\alpha_1$	$-4\gamma\varrho(1-\alpha_1)$	$-4\gamma\varrho(1-2\alpha_2)$
	$\hat{b}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	0	$4\gamma\varrho^2(\alpha_1-\alpha_2)$	$-4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$4\gamma\varrho^2\alpha_1(1+\alpha_2)$	$-4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$4\gamma\varrho^2(\alpha_1-\alpha_2)$
$\boldsymbol{s}_t=[1,0,1]$	$\hat{a}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	$-4\gamma\varrho(1-\alpha_1-\alpha_2)$	$4\gamma\varrho(\alpha_1-\alpha_2)$	$4\gamma\varrho\alpha_2$	$-4\gamma\varrho(1-\alpha_2)$	$4\gamma\varrho\alpha_1$	$-4\gamma\varrho(1-\alpha_1)$	$-4\gamma\varrho$
	$\hat{b}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	0	$4\gamma\varrho^2(\alpha_1-\alpha_2)$	$4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$4\gamma\varrho^2(\alpha_1-2\alpha_1\alpha_2+\alpha_2)$
$\boldsymbol{s}_t=[0,0,1]$	$\hat{a}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	$4\gamma\varrho(1-\alpha_1-\alpha_2)$	$-4\gamma\varrho(\alpha_1-\alpha_2)$	$-4\gamma\varrho\alpha_2$	$4\gamma\varrho(1-\alpha_2)$	$-4\gamma\varrho\alpha_1$	$4\gamma\varrho(1-\alpha_1)$	$4\gamma\varrho$
	$\hat{b}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	0	$4\gamma\varrho^2(\alpha_1-\alpha_2)$	$4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$4\gamma\varrho^2(\alpha_1-2\alpha_1\alpha_2+\alpha_2)$
$\boldsymbol{s}_t=[0,0,0]$	$\hat{a}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	$4\gamma\varrho(1-\alpha_1-\alpha_2)$	$-4\gamma\varrho(\alpha_1-\alpha_2)$	$4\gamma\varrho\alpha_2$	$4\gamma\varrho(1-\alpha_2)$	$-4\gamma\varrho\alpha_1$	$4\gamma\varrho(1-\alpha_1)$	$4\gamma\varrho(1-2\alpha_2)$
	$\hat{b}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	0	$4\gamma\varrho^2(\alpha_1-\alpha_2)$	$-4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$4\gamma\varrho^2\alpha_1(1+\alpha_2)$	$-4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$4\gamma\varrho^2(\alpha_1-\alpha_2)$
$\boldsymbol{s}_t=[0,1,0]$	$\hat{a}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	$4\gamma\varrho(1-\alpha_1-\alpha_2)$	$4\gamma\varrho(\alpha_1-\alpha_2)$	$-4\gamma\varrho\alpha_2$	$4\gamma\varrho(1-\alpha_2)$	$4\gamma\varrho\alpha_1$	$4\gamma\varrho(1-\alpha_1)$	$4\gamma\varrho$
	$\hat{b}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	0	$-4\gamma\varrho^2(\alpha_1-\alpha_2)$	$4\gamma\varrho^2\alpha_2(1+\alpha_1)$	$-4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$-4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$-4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$-4\gamma\varrho^2(\alpha_1-2\alpha_1\alpha_2+\alpha_2)$
$\boldsymbol{s}_t=[0,1,1]$	$\hat{a}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	$4\gamma\varrho(1-\alpha_1-\alpha_2)$	$4\gamma\varrho(\alpha_1-\alpha_2)$	$4\gamma\varrho\alpha_2$	$4\gamma\varrho(1-\alpha_2)$	$4\gamma\varrho\alpha_1$	$4\gamma\varrho(1-\alpha_1)$	$4\gamma\varrho(1-2\alpha_2)$
	$\hat{b}(\boldsymbol{e}_t,\boldsymbol{s}_t)$	0	$-4\gamma\varrho^2(\alpha_1-\alpha_2)$	$-4\gamma\varrho^2\alpha_2(1+\alpha_1)$	$-4\gamma\varrho^2\alpha_1(1-\alpha_2)$	$-4\gamma\varrho^2\alpha_1(1+\alpha_2)$	$4\gamma\varrho^2\alpha_2(1-\alpha_1)$	$-4\gamma\varrho^2(\alpha_1-\alpha_2)$

Table 2.4: Values of $\hat{\mu}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ and $\hat{\sigma}^2(\boldsymbol{e}_t, \boldsymbol{s}_t)$ for HPAM ($M = 8$) where $\varrho = \dfrac{1}{\sqrt{1+\alpha_1^2+\alpha_2^2}}$

		$e_t=[1,0,0]$	$e_t=[0,1,0]$	$e_t=[0,0,1]$	$e_t=[1,1,0]$	$e_t=[0,1,1]$	$e_t=[1,0,1]$	$e_t=[1,1,1]$
$s_t=[1,1,1]$	$\hat{\mu}(e_t,s_t)$	$4\gamma\varrho^2(1-2\alpha_1\alpha_2-\alpha_1^2-\alpha_2^2)$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_2^2$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_1^2$	$4\gamma\varrho^2(1+2\alpha_2-2\alpha_1\alpha_2-\alpha_1^2)$	$4\gamma\varrho^2(1-2\alpha_1\alpha_2-2\alpha_2^2)$
	$\hat{\sigma}^2(e_t,s_t)$	$8\gamma\varrho^2(1-\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2(\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2(1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2(1-\alpha_1)^2$	$8\gamma\varrho^2(1-2\alpha_2^2)$
$s_t=[1,1,0]$	$\hat{\mu}(e_t,s_t)$	$4\gamma\varrho^2(1-2\alpha_2-\alpha_1^2+\alpha_2^2)$	$4\gamma\varrho^2(\alpha_2-\alpha_2)^2$	$4\gamma\varrho^2\alpha_2^2$	$4\gamma\varrho^2(1-\alpha_2)^2$	$4\gamma\varrho^2\alpha_1^2$	$4\gamma\varrho^2(1-2\alpha_2+2\alpha_1\alpha_2-\alpha_1^2)$	$4\gamma\varrho^2(-2\alpha_2+2\alpha_1\alpha_2+\alpha_1^2)$
	$\hat{\sigma}^2(e_t,s_t)$	$8\gamma\varrho^2(1-\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2(\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2(1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2(1-\alpha_1)^2$	$8\gamma\varrho^2$
$s_t=[1,0,0]$	$\hat{\mu}(e_t,s_t)$	$4\gamma\varrho^2(1-2\alpha_1+\alpha_1^2-\alpha_2^2)$	$4\gamma\varrho^2(\alpha_2-\alpha_2)^2$	$4\gamma\varrho^2\alpha_2^2$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_1^2$	$4\gamma\varrho^2(1-\alpha_1)^2$	$4\gamma\varrho^2(1-2\alpha_2+2\alpha_1\alpha_2-2\alpha_2^2)$
	$\hat{\sigma}^2(e_t,s_t)$	$8\gamma\varrho^2(1-\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2(\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2(1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2(1-\alpha_1)^2$	$8\gamma\varrho^2(1-2\alpha_2^2)$
$s_t=[1,0,1]$	$\hat{\mu}(e_t,s_t)$	$4\gamma\varrho^2(1-2\alpha_1-2\alpha_2+2\alpha_1\alpha_2+\alpha_1^2+\alpha_2^2)$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_2^2$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_1^2$	$4\gamma\varrho^2(1-\alpha_1)^2$	$4\gamma\varrho^2(1-2\alpha_1\alpha_2)$
	$\hat{\sigma}^2(e_t,s_t)$	$8\gamma\varrho^2(1-\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2(\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2(1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2(1-\alpha_1)^2$	$8\gamma\varrho^2$
$s_t=[0,0,1]$	$\hat{\mu}(e_t,s_t)$	$4\gamma\varrho^2(1-2\alpha_1-2\alpha_2+2\alpha_1\alpha_2+\alpha_1^2+\alpha_2^2)$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_2^2$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_1^2$	$4\gamma\varrho^2(1-\alpha_1)^2$	$4\gamma\varrho^2(1-2\alpha_1\alpha_2)$
	$\hat{\sigma}^2(e_t,s_t)$	$8\gamma\varrho^2(1-\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2(\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2(1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2(1-\alpha_1)^2$	$8\gamma\varrho^2$
$s_t=[0,0,0]$	$\hat{\mu}(e_t,s_t)$	$4\gamma\varrho^2(1-2\alpha_1+\alpha_1^2-\alpha_2^2)$	$4\gamma\varrho^2(\alpha_2-\alpha_2)^2$	$4\gamma\varrho^2\alpha_2^2$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_1^2$	$4\gamma\varrho^2(1-\alpha_1)^2$	$4\gamma\varrho^2(1-2\alpha_2+2\alpha_1\alpha_2-2\alpha_2^2)$
	$\hat{\sigma}^2(e_t,s_t)$	$8\gamma\varrho^2(1-\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2(\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2(1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2(1-\alpha_1)^2$	$8\gamma\varrho^2(1-2\alpha_2^2)$
$s_t=[0,1,0]$	$\hat{\mu}(e_t,s_t)$	$4\gamma\varrho^2(1-2\alpha_2-\alpha_1^2+\alpha_2^2)$	$4\gamma\varrho^2(\alpha_2-\alpha_2)^2$	$4\gamma\varrho^2\alpha_2^2$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_1^2$	$4\gamma\varrho^2(1-2\alpha_2+2\alpha_1\alpha_2-\alpha_1^2)$	$4\gamma\varrho^2(-2\alpha_2+2\alpha_1\alpha_2+1)$
	$\hat{\sigma}^2(e_t,s_t)$	$8\gamma\varrho^2(1-\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2(\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2(1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2(1-\alpha_1)^2$	$8\gamma\varrho^2$
$s_t=[0,1,1]$	$\hat{\mu}(e_t,s_t)$	$4\gamma\varrho^2(1-2\alpha_1\alpha_2-\alpha_1^2-\alpha_2^2)$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_2^2$	$4\gamma\varrho^2(1-\alpha_2^2)$	$4\gamma\varrho^2\alpha_1^2$	$4\gamma\varrho^2(1+2\alpha_2-2\alpha_1\alpha_2-\alpha_1^2)$	$4\gamma\varrho^2(1-2\alpha_1\alpha_2-2\alpha_2^2)$
	$\hat{\sigma}^2(e_t,s_t)$	$8\gamma\varrho^2(1-\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2(\alpha_1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2(1-\alpha_2)^2$	$8\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2(1-\alpha_1)^2$	$8\gamma\varrho^2(1-2\alpha_2^2)$

in [9] for regular 4-PAM constellation.

2.4.2 H-64QAM

In this example, we consider a 8-ary constellation ($q = 3$) which together with a rate $R = 1/3$ code gives a spectral efficiency of 1 bits/dimension and with $R = 1/2$ code gives a spectral efficiency of 1.5 bits/dimension. In this case there are two constellation parameters i.e., α_1 and α_2. Because there are three bit positions, $\Pi = \{1, 2, 3, (1, 2), (2, 3), (1, 3), (1, 2, 3)\}$ and we need to compute seven conditional PDFs i.e., $p_{\Lambda_1}(\lambda|C_1)$, $p_{\Lambda_2}(\lambda|C_2)$, $p_{\Lambda_3}(\lambda|C_3)$, $p_{\Lambda_{1,2}}(\lambda|\boldsymbol{C}_{1,2})$, $p_{\Lambda_{2,3}}(\lambda|\boldsymbol{C}_{2,3})$, $p_{\Lambda_{1,3}}(\lambda|\boldsymbol{C}_{1,3})$ and $p_{\Lambda_{1,2,3}}(\lambda|\boldsymbol{C}_{1,2,3})$. Since the L-values $\Lambda(\boldsymbol{e}_t, \boldsymbol{s}_t)$ are piecewise linear, we present the values of slope $\hat{a}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ and intercept $\hat{b}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ in terms of constellation parameters required for the zero-crossing approximation (2.25) in Table 2.3 for all possible values of $\boldsymbol{e}_t$ and $\boldsymbol{s}_t$. In Table 2.4 we present the means and variances in (2.27)-(2.28)

for the same cases presented in Table 2.3. By averaging (2.26) over the scrambling sequence $\boldsymbol{s}_t$, we obtain the unconditional PDF of the L-values in (2.16) given by the following equation

$$
p_\Lambda(\lambda) =
\begin{cases}
\frac{1}{4}\sum_{p=1}^{4}\psi(\lambda;\mu_{p,\gamma,\alpha_1,\alpha_2},\sigma^2_{p,\gamma,\alpha_1,\alpha_2}), & \boldsymbol{e}_t = [1,0,0] \\[2ex]
\frac{1}{2}\sum_{p=5}^{6}\psi(\lambda;\mu_{p,\gamma,\alpha_1,\alpha_2},\sigma^2_{p,\gamma,\alpha_1,\alpha_2}), & \boldsymbol{e}_t = [0,1,0] \\[2ex]
\psi(\lambda;\mu_{7,\gamma,\alpha_1,\alpha_2},\sigma^2_{7,\gamma,\alpha_1,\alpha_2}), & \boldsymbol{e}_t = [0,0,1] \\[2ex]
\frac{1}{2}\sum_{p=8}^{9}\psi(\lambda;\mu_{p,\gamma,\alpha_1,\alpha_2},\sigma^2_{p,\gamma,\alpha_1,\alpha_2}), & \boldsymbol{e}_t = [1,1,0] \\[2ex]
\psi(\lambda;\mu_{10,\gamma,\alpha_1,\alpha_2},\sigma^2_{10,\gamma,\alpha_1,\alpha_2}), & \boldsymbol{e}_t = [0,1,1] \\[2ex]
\frac{1}{4}(\sum_{p=11}^{12}\psi(\lambda;\mu_{p,\gamma,\alpha_1,\alpha_2},\sigma^2_{p,\gamma,\alpha_1,\alpha_2}) \\[1ex]
\qquad +2\psi(\lambda;\mu_{13,\gamma,\alpha_1,\alpha_2},\sigma^2_{13,\gamma,\alpha_1,\alpha_2})), & \boldsymbol{e}_t = [1,0,1] \\[2ex]
\frac{1}{4}\sum_{p=14}^{17}\psi(\lambda;\mu_{p,\gamma,\alpha_1,\alpha_2},\sigma^2_{p,\gamma,\alpha_1,\alpha_2}), & \boldsymbol{e}_t = [1,1,1]
\end{cases}
\tag{2.40}
$$

where means and variances in $\psi(\lambda;\mu_{p,\gamma,\alpha_1,\alpha_2},\sigma^2_{p,\gamma,\alpha_1,\alpha_2})$ for different values of p are summarized in Table 2.5.

Using (2.40), we can derive the expressions for elements in the integration defining PEP in (2.18) as,

$$
p_{\Lambda_1}(\lambda|C_1 = 0)^{*w_1} = \left(\frac{1}{4}\right)^{w_1} \sum_{\substack{\sum_{l=1}^{4} j_l = w_1 \\ j_l \geq 0}} \binom{w_1}{j_1 \ldots j_4} \cdot \psi\left(\lambda; \sum_{i=1}^{4}\mu_{i,\gamma,\alpha_1,\alpha_2} \cdot j_i, \right.
$$

$$
\left. \sigma^2_{1,\gamma,\alpha_1,\alpha_2} \cdot w_1 \right)
$$

$$
\tag{2.41}
$$

Table 2.5: Values of $(\mu_{p,\gamma,\alpha_1,\alpha_2})$ and $(\sigma^2_{p,\gamma,\alpha_1,\alpha_2})$ in (2.40) for different values of p

	$\mu_{p,\gamma,\alpha_1,\alpha_2}$	$\sigma^2_{p,\gamma,\alpha_1,\alpha_2}$
$p = 1$	$4\gamma\varrho^2(1 - 2\alpha_1\alpha_2 - \alpha_1^2 - \alpha_2^2)$	$8\gamma\varrho^2(1 - \alpha_1 - \alpha_2)^2$
$p = 2$	$4\gamma\varrho^2(1 - 2\alpha_2 - \alpha_1^2 + \alpha_2^2)$	$8\gamma\varrho^2(1 - \alpha_1 - \alpha_2)^2$
$p = 3$	$4\gamma\varrho^2(1 - 2\alpha_1 + \alpha_1^2 - \alpha_2^2)$	$8\gamma\varrho^2(1 - \alpha_1 - \alpha_2)^2$
$p = 4$	$4\gamma\varrho^2(1 - 2\alpha_1 - 2\alpha_2 + 2\alpha_1\alpha_2 + \alpha_1^2 + \alpha_2^2)$	$8\gamma\varrho^2(1 - \alpha_1 - \alpha_2)^2$
$p = 5$	$4\gamma\varrho^2(\alpha_1^2 - \alpha_2^2)$	$8\gamma\varrho^2(\alpha_1 - \alpha_2)^2$
$p = 6$	$4\gamma\varrho^2(\alpha_1 - \alpha_2)^2$	$8\gamma\varrho^2(\alpha_1 - \alpha_2)^2$
$p = 7$	$4\gamma\varrho^2\alpha_2^2$	$8\gamma\varrho^2\alpha_2^2$
$p = 8$	$4\gamma\varrho^2(1 - \alpha_2^2)$	$8\gamma\varrho^2(1 - \alpha_2)^2$
$p = 9$	$4\gamma\varrho^2(1 - \alpha_2)^2$	$8\gamma\varrho^2(1 - \alpha_2)^2$
$p = 10$	$4\gamma\varrho^2\alpha_1^2$	$8\gamma\varrho^2\alpha_1^2$
$p = 11$	$4\gamma\varrho^2(1 + 2\alpha_2 - 2\alpha_1\alpha_2 - \alpha_1^2)$	$8\gamma\varrho^2(1 - \alpha_1)^2$
$p = 12$	$4\gamma\varrho^2(1 - 2\alpha_2 + 2\alpha_1\alpha_2 - \alpha_1^2)$	$8\gamma\varrho^2(1 - \alpha_1)^2$
$p = 13$	$4\gamma\varrho^2(1 - \alpha_1)^2$	$8\gamma\varrho^2(1 - \alpha_1)^2$
$p = 14$	$4\gamma\varrho^2(1 - 2\alpha_1\alpha_2 - 2\alpha_2^2)$	$8\gamma\varrho^2(1 - 2\alpha_2)^2$
$p = 15$	$4\gamma\varrho^2(1 - 2\alpha_2 + 2\alpha_1\alpha_2)$	$8\gamma\varrho^2$
$p = 16$	$4\gamma\varrho^2(1 - 2\alpha_2 + 2\alpha_1\alpha_2 - 2\alpha_2^2)$	$8\gamma\varrho^2(1 - 2\alpha_2)^2$
$p = 17$	$4\gamma\varrho^2(1 - 2\alpha_1\alpha_2)$	$8\gamma\varrho^2$

$$p_{\Lambda_2}(\lambda|C_2 = 0)^{*w_2} = \left(\frac{1}{2}\right)^{w_2} \sum_{\substack{\sum_{\bar{l}=1}^{2} \bar{j}_{\bar{l}} = w_2 \\ \bar{j}_{\bar{l}} \geq 0}} \binom{w_2}{\bar{j}_1, \bar{j}_2} \cdot \psi\left(\lambda; \sum_{\bar{i}=5}^{6} \mu_{\bar{i},\gamma,\alpha_1,\alpha_2} \cdot \bar{j}_{\bar{i}-4},\right.$$

$$\left. \sigma^2_{5,\gamma,\alpha_1,\alpha_2} \cdot w_2 \right)$$

$$(2.42)$$

$$p_{\Lambda_3}(\lambda|C_3 = 0)^{*w_3} = \psi\left(\lambda; \mu_{7,\gamma,\alpha_1,\alpha_2,\alpha_2} \cdot w_3, \sigma^2_{7,\gamma,\alpha_1,\alpha_2} \cdot w_3\right) \qquad (2.43)$$

$$p_{\Lambda_{1,2}}(\lambda|\boldsymbol{C}_{1,2}=0)^{*w_{1,2}} = \left(\frac{1}{2}\right)^{w_{1,2}} \sum_{\substack{\sum_{\tilde{l}=1}^{2}\tilde{j}_{\tilde{l}}=w_{1,2}\\ \tilde{j}_{\tilde{l}}\geq 0}} \binom{w_{1,2}}{\tilde{j}_1,\tilde{j}_2}\cdot$$

$$\psi\left(\lambda;\sum_{\tilde{i}=8}^{9}\mu_{\tilde{i},\gamma,\alpha_1,\alpha_2}\cdot\tilde{j}_{\tilde{i}-7},\sigma^2_{8,\gamma,\alpha_1,\alpha_2}\cdot w_{1,2}\right)$$

$$(2.44)$$

$$p_{\Lambda_{2,3}}(\lambda|\boldsymbol{C}_{2,3}=0)^{*w_{2,3}} = \psi\left(\lambda;\mu_{10,\gamma,\alpha_1,\alpha_2,\alpha_2}\cdot w_{2,3},\sigma^2_{10,\gamma,\alpha_1,\alpha_2}\cdot w_{2,3}\right) \quad (2.45)$$

$$p_{\Lambda_{1,3}}(\lambda|\boldsymbol{C}_{1,3}=0)^{*w_{1,3}} = \left(\frac{1}{4}\right)^{w_{1,3}} \sum_{\substack{\sum_{\hat{l}=1}^{3}\hat{j}_{\hat{l}}=w_{1,3}\\ \hat{j}_{\hat{l}}\geq 0}} \binom{w_{1,3}}{\hat{j}_1\ldots\hat{j}_3}\cdot$$

$$\psi\left(\lambda;\sum_{\hat{i}=11}^{12}\mu_{\hat{i},\gamma,\alpha_1,\alpha_2}\cdot\hat{j}_{\hat{i}-10}+2\mu_{13,\gamma,\alpha_1,\alpha_2}\cdot\hat{j}_3,\sigma^2_{11,\gamma,\alpha_1,\alpha_2}\cdot w_{1,3}\right)$$

$$(2.46)$$

$$p_{\Lambda_{1,2,3}}(\lambda|\boldsymbol{C}_{1,2,3}=0)^{*w_{1,2,3}} = \left(\frac{1}{4}\right)^{w_{1,2,3}} \sum_{\substack{\sum_{\check{l}=1}^{4}\check{j}_{\check{l}}=w_{1,2,3}\\ \check{j}_{\check{l}}\geq 0}} \binom{w_{1,2,3}}{\check{j}_1\ldots\check{j}_4}\cdot$$

$$\psi\left(\lambda;\sum_{\check{i}=14}^{17}\mu_{\check{i},\gamma,\alpha_1,\alpha_2}\cdot\check{j}_{\check{i}-13},\sum_{m=14}^{15}\sigma^2_{m,\gamma,\alpha_1,\alpha_2}\cdot(\check{j}_{m-13}+\check{j}_{m-12})\right)$$

$$(2.47)$$

where the multinomial coefficients come from the convolution of the sum of Gaussian functions in (2.40). Using (2.41)-(2.47) in (2.18), we get the expressions for PEP and UB presented in (2.48) and (2.49) respectively,

$$\text{PEP}(w_\Pi) = \left(\frac{1}{2}\right)^{w_\text{T}} \sum_{\substack{\sum_{l=1}^{4} j_l = w_1 \\ j_l \geq 0}} \sum_{\substack{\sum_{\bar{l}=1}^{2} \bar{j}_{\bar{l}} = w_2 \\ \bar{j}_{\bar{l}} \geq 0}} \sum_{\substack{\sum_{\tilde{l}=1}^{2} \tilde{j}_{\tilde{l}} = w_{1,2} \\ \tilde{j}_{\tilde{l}} \geq 0}} \sum_{\substack{\sum_{\hat{l}=1}^{3} \hat{j}_{\hat{l}} = w_{1,3} \\ \hat{j}_{\hat{l}} \geq 0}} \sum_{\substack{\sum_{\check{l}=1}^{4} \check{j}_{\check{l}} = w_{1,2,3} \\ \check{j}_{\check{l}} \geq 0}}$$

$$\binom{w_1}{j_1 \dots j_4} \binom{w_2}{\bar{j}_1, \bar{j}_2} \binom{w_{1,2}}{\tilde{j}_1, \tilde{j}_2} \binom{w_{1,3}}{\hat{j}_1 \dots \hat{j}_3} \binom{w_{1,2,3}}{\check{j}_1 \dots \check{j}_4} \int_0^\infty \psi(\lambda; \mu_{w_\Sigma}, \sigma^2_{w_\Sigma}) \mathrm{d}\lambda, \quad (2.48)$$

$$\text{UB} = \sum_{w_\Pi} \beta^{\mathbb{C}}_{w_\Pi} \left(\frac{1}{2}\right)^{w_\text{T}} \sum_{\substack{\sum_{l=1}^{4} j_l = w_1 \\ j_l \geq 0}} \sum_{\substack{\sum_{\bar{l}=1}^{2} \bar{j}_{\bar{l}} = w_2 \\ \bar{j}_{\bar{l}} \geq 0}} \sum_{\substack{\sum_{\tilde{l}=1}^{2} \tilde{j}_{\tilde{l}} = w_{1,2} \\ \tilde{j}_{\tilde{l}} \geq 0}} \sum_{\substack{\sum_{\hat{l}=1}^{3} \hat{j}_{\hat{l}} = w_{1,3} \\ \hat{j}_{\hat{l}} \geq 0}} \sum_{\substack{\sum_{\check{l}=1}^{4} \check{j}_{\check{l}} = w_{1,2,3} \\ \check{j}_{\check{l}} \geq 0}}$$

$$\binom{w_1}{j_1 \dots j_4} \binom{w_2}{\bar{j}_1, \bar{j}_2} \binom{w_{1,2}}{\tilde{j}_1, \tilde{j}_2} \cdot \binom{w_{1,3}}{\hat{j}_1 \dots \hat{j}_3} \binom{w_{1,2,3}}{\check{j}_1 \dots \check{j}_4} \cdot Q\left(\sqrt{\frac{\mu^2_{w_\Sigma}}{\sigma^2_{w_\Sigma}}}\right) \cdot$$

$$(2.49)$$

where

$$w_\text{T} = 2(w_1 + w_{1,3} + w_{1,2,3}) + w_2 + w_{1,2}, \quad (2.50)$$

$$\mu_{w_\Sigma} = \left(\sum_{i=1}^{4} \mu_{i,\gamma,\alpha_1,\alpha_2} \cdot j_i + \sum_{\bar{i}=5}^{6} \mu_{\bar{i},\gamma,\alpha_1,\alpha_2} \cdot \bar{j}_{\bar{i}-4} + \mu_{7,\gamma,\alpha_1,\alpha_2} \cdot w_3 \right.$$

$$+ \sum_{\tilde{i}=8}^{9} \mu_{\tilde{i},\gamma,\alpha_1,\alpha_2} \cdot \tilde{j}_{\tilde{i}-7} + \sum_{\hat{i}=11}^{12} \mu_{\hat{i},\gamma,\alpha_1,\alpha_2} \cdot \hat{j}_{\hat{i}-10} + 2\mu_{13,\gamma,\alpha_1,\alpha_2} \cdot$$

$$\left. \hat{j}_3 \sum_{\check{i}=14}^{17} \mu_{\check{i},\gamma,\alpha_1,\alpha_2} \cdot \check{j}_{\check{i}-13} \right), \quad (2.51)$$

$$\sigma_{w_\Sigma}^2 = \big(\sigma_{1,\gamma,\alpha_1,\alpha_2}^2 \cdot w_1 + \sigma_{5,\gamma,\alpha_1,\alpha_2}^2 \cdot w_2 + \sigma_{7,\gamma,\alpha_1,\alpha_2}^2 \cdot w_3 + \sigma_{8,\gamma,\alpha_1,\alpha_2}^2 \cdot w_{1,2}$$

$$+\sigma_{10,\gamma,\alpha_1,\alpha_2}^2 \cdot w_{2,3} + \sigma_{11,\gamma,\alpha_1,\alpha_2}^2 \cdot w_{1,3} + \sigma_{14,\gamma,\alpha_1,\alpha_2}^2 \cdot (\check{j}_1 + \check{j}_3)$$

$$+\sigma_{15,\gamma,\alpha_1,\alpha_2}^2 \cdot (\check{j}_2 + \check{j}_4)\big). \tag{2.52}$$

2.5   Numerical Results

Numerical results are presented in this section to illustrate the gains that can be obtained by using HQAM-BICM-T design with optimized system modules. In particular, the results for two spectral efficiencies: 1 bit/dimension and 1.5 bits/dimension are presented. Convolutional codes with rate R, constraint length K are used for encoding at the source. A soft-input Viterbi algorithm without memory truncation is used for decoding at the receiver.

The UB in general depends on the SNR, the signal constellation parameter, and the MUX configuration. Therefore, for an optimal design, $\boldsymbol{K}$ and the signal constellation parameters should be jointly optimized for each value of SNR. The optimum values are defined as $\alpha_1^*(\gamma), \ldots \alpha_q^*(\gamma), \boldsymbol{K}^*(\gamma)$, i.e.,

$$[\alpha_1^*(\gamma), \ldots \alpha_q^*(\gamma), \boldsymbol{K}^*(\gamma)] = \underset{\alpha_1(\gamma),\ldots\alpha_q(\gamma),\boldsymbol{K}}{\operatorname{argmin}} \{\mathrm{UB}(\alpha_1(\gamma), \ldots \alpha_q(\gamma), \boldsymbol{K})\} \tag{2.53}$$

where UB is a function of the signal constellation defined by the parameters $\alpha_1, \ldots, \alpha_q$ and the MUX configuration matrix $\boldsymbol{K}$. The optimization was per-

formed using an exhaustive search over the allowable range of constellation parameters which is a feasible method compared to complex optimization approaches because the UB function is potentially non-convex. Please note that this optimization can be performed offline and the results can be tabulated for practical values of SNRs. Then those results can be used for the practical system. All UB computations were carried out considering a truncated spectrum of the code, i.e., $w_{\pi_i} \leq 40 \ \forall \ \pi_i \in \Pi$. Using more number of terms does not affect the UB as it converges.

2.5.1 Spectral Efficiency 1 Bit/Dimension

H-16QAM with Rate R=1/2 Code

In this example, we consider a rate $R = 1/2$ optimum distance spectrum (ODS) [30] convolutional code with constraint length $K = 3$ and generator polynomials $(5, 7)_8$. This code together with a 4-ary constellation gives a spectral efficiency of 1 bit/dimension. In this case ($n = q = 2$), α_1 is the only constellation parameter. A MUX with shortest possible period $J = 1$ is considered, which results in two different $\boldsymbol{K}$ matrices.

$$\boldsymbol{K}^{(1)} = \begin{bmatrix} (1, 1) \\ (2, 1) \end{bmatrix}, \boldsymbol{K}^{(2)} = \begin{bmatrix} (2, 1) \\ (1, 1) \end{bmatrix}. \tag{2.54}$$

A joint optimization was performed over $\boldsymbol{K}$ and α_1. We found that the optimum MUX is always defined by $\boldsymbol{K}^*(\gamma) = \boldsymbol{K}^{(2)}$. This MUX configuration assigns all the coded bits from the first encoder's output to bit position $k = 2$ and all the coded bits from the second encoder's output to bit position $k = 1$.

Similar configuration was found to be the "Best" configuration in [9] when ES 16-QAM constellation was used. It is interesting to note from the Fig. 2.7 that for higher values of SNR, the optimal constellation tends to be non-uniform. The BER performance of the considered design with optimized constellation parameters and MUX configuration is presented in Fig. 2.8. We also present the BER performance of the optimal HQAM-BICM system of [12], BICM-T ("Best") system of [9] and traditional single interleaver BICM system (BICM-S) [4]. When compared to the BICM-T "Best" design [9] that uses ES 16-QAM, the considered optimized BICM-T design offers gains up to 0.21 dB for a target BER of 10^{-6}.

64-HQAM with Rate $R = 1/3$ Code

In this example, we consider a rate $R = 1/3$ ODS convolutional code with constraint length $K = 3$ and generator polynomials $(5, 7, 7)_8$. This code together with a 8-ary constellation gives a spectral efficiency of 1 bit/dimension. In this case there are two constellation parameters i.e., α_1 and α_2. We consider MUX configurations with the period, i.e., $J = 1$, for which there will be a total of six different MUX configurations as given below:

$$\boldsymbol{K}^{(1)} = \begin{bmatrix} (1,1) \\ (2,1) \\ (3,1) \end{bmatrix}, \boldsymbol{K}^{(2)} = \begin{bmatrix} (1,1) \\ (3,1) \\ (2,1) \end{bmatrix}, \boldsymbol{K}^{(3)} = \begin{bmatrix} (2,1) \\ (1,1) \\ (3,1) \end{bmatrix}$$

$$\boldsymbol{K}^{(4)} = \begin{bmatrix} (2,1) \\ (3,1) \\ (1,1) \end{bmatrix}, \boldsymbol{K}^{(4)} = \begin{bmatrix} (3,1) \\ (1,1) \\ (2,1) \end{bmatrix}, \boldsymbol{K}^{(6)} = \begin{bmatrix} (3,1) \\ (2,1) \\ (1,1) \end{bmatrix}. \qquad (2.55)$$

For SNR γ [dB] $\in \{6, 7, \ldots, 9\}$ (which give a BER below 10^{-3}), we obtained the optimal MUX configuration $\boldsymbol{K}^*(\gamma)$ and constellation parameters $(\alpha_1^*(\gamma), \alpha_2^*(\gamma))$ using the developed UB in(2.49) and an exhaustive search. The optimal matrix for all γ [dB] $\in \{6, 7, \ldots, 9\}$ was found to be $\boldsymbol{K}^*(\gamma) = \boldsymbol{K}^{(1)}$. We note that the optimal constellation parameters do not change significantly for the SNR range of interest. The obtained values of optimal constellation parameters are

$$(\alpha_1^*(\gamma), \alpha_2^*(\gamma)) = (0.47, 0.12). \qquad (2.56)$$

The BER performance of the considered design with optimized constellation parameters and MUX configuration is presented in Fig. 2.9. Here we also present the BER performance of the optimal HQAM-BICM system of [12], the traditional single interleaver BICM system [4], and the BICM-T system that uses an equally spaced signal constellation and no multiplexer. When compared to the optimal HQAM-BICM design [12], the considered system offers gains up to 2.5 dB for a target BER of 10^{-6}.

2.5.2 Spectral Efficiency 1.5 Bits/Dimension

In this section, we consider a rate $R = 1/2$ ODS convolutional code with constraint length $K = 3$ and generator polynomials $(5, 7)_8$. This

code together with a 8-ary constellation gives a spectral efficiency of 1.5 bits/dimension. In this case there are two constellation parameters i.e., α_1 and α_2. We consider MUX configurations with the period, i.e., $J = 3$, for which there will be a total of thirty different MUX configurations [12].

For SNR γ [dB] $\in \{10, 11, \ldots, 16\}$ (which give a UB below 10^{-3}), we obtained the optimal MUX configuration $\boldsymbol{K}^*(\gamma)$ and constellation parameters $(\alpha_1^*(\gamma), \alpha_2^*(\gamma))$ using the developed UB in(2.49) and an exhaustive search. The optimal MUX for all γ [dB] $\in \{10, 11, \ldots, 16\}$ was found to be

$$\boldsymbol{K}^*(\gamma) = \begin{bmatrix} (1,1) & (2,1) & (3,1) \\ (3,2) & (2,2) & (1,2) \end{bmatrix}. \tag{2.57}$$

We note that the optimal constellation parameters do not change significantly for the SNR range of interest. The obtained values of optimal constellation parameters are

$$(\alpha_1^*(\gamma), \alpha_2^*(\gamma)) = (0.46, 0.01). \tag{2.58}$$

The results in (2.58) indicate that the optimal constellation is approximately a 16-QAM constellation ($\alpha_2 \approx 0$) with $\alpha_1 \approx 0.46$ which translates into a system where the third output of the MUX is completely eliminated. Another way of interpreting the results in (2.58) is that for this code, the minimum BER is obtained when the original rate 1/2 code is punctured to a rate 3/4 code and transmitted with a 16-QAM constellation (with $\alpha_1 \approx 0.46$). The BER performance of the considered design with optimized constellation parameters and MUX configuration is presented in Fig. 2.10. Here we

also present the BER performance of the BICM-T system that uses an ES 16-QAM constellation with punctured rate 1/2 code, the optimal HQAM-BICM system of [12], and the traditional single interleaver BICM system [4]. When compared to the optimal HQAM-BICM design [12], the considered system offers gains up to 0.2 dB for a target BER of 10^{-6}. Moreover, the considered system has less complexity as compared to the optimal HQAM-BICM design [12] (due to the absence of interleaver/deinterleaver).

2.5.3   Optimal MUX Configuration for Rate R=1/2 ODS Codes

In this section, we present the MUX configuration $\boldsymbol{K}^*$ that results in optimal BER performance for a target BER of 10^{-6} in Table 2.6 for different rate $R = 1/2$ ODS codes in [30]. The optimal constellation parameters do not change significantly for the codes presented under consideration. We present the results for two particular cases of 16-QAM and 64-QAM which together with rate $R = 1/2$ codes results in spectral efficiencies of 1 bit/dimension and 1.5 bits/dimension respectively.

The results show that the optimal multiplexer configuration can be different for different codes.

2.6   Conclusion

In this Chapter, we presented an innovative BICM-T design that uses HQAM signal constellations in conjunction with a bit-level multiplexer. Due to the use of a bit level multiplexer that connects the channel encoder with

Table 2.6: Optimal multiplexer configuration for different rate $R = 1/2$ ODS codes with 16-QAM and 64-QAM for a target BER of 10^{-6}. All generator polynomials are in octal notation.

K	Generators	$d_{\mathrm{H}}^{\mathrm{free}}$	$\boldsymbol{K^*}(16-\mathrm{QAM})$		$\boldsymbol{K^*}(64-\mathrm{QAM})$		
3	$5, 7$	5	$(2,1)$ $(2,2)$ $(1,1)$ $(1,2)$		$(1,1)$ $(2,1)$ $(3,1)$ $(3,2)$ $(2,2)$ $(1,2)$		
4	$15, 17$	6	$(1,1)$ $(1,2)$ $(2,1)$ $(2,2)$		$(1,1)$ $(1,2)$ $(2,2)$ $(3,1)$ $(3,2)$ $(2,1)$		
5	$23, 35$	7	$(2,1)$ $(2,2)$ $(1,1)$ $(1,2)$		$(2,2)$ $(3,2)$ $(1,2)$ $(1,1)$ $(3,1)$ $(2,1)$		
6	$53, 75$	8	$(2,1)$ $(1,1)$ $(1,2)$ $(2,2)$		$(1,1)$ $(3,1)$ $(2,1)$ $(2,2)$ $(3,2)$ $(1,2)$		
7	$133, 171$	10	$(2,1)$ $(2,2)$ $(1,1)$ $(1,2)$		$(2,1)$ $(1,1)$ $(3,1)$ $(2,2)$ $(3,2)$ $(1,2)$		
8	$247, 371$	10	$(2,1)$ $(2,2)$ $(1,1)$ $(1,2)$		$(1,1)$ $(2,1)$ $(3,1)$ $(3,2)$ $(2,2)$ $(1,2)$		
9	$561, 753$	12	$(1,1)$ $(1,2)$ $(2,1)$ $(2,2)$		$(1,1)$ $(2,1)$ $(3,1)$ $(3,2)$ $(2,2)$ $(1,2)$		
10	$1151, 1753$	12	$(2,1)$ $(2,2)$ $(1,1)$ $(1,2)$		$(1,1)$ $(2,2)$ $(3,2)$ $(3,1)$ $(2,1)$ $(1,2)$		

the modulator, the considered design gives a design flexibility to independently choose the code rate and the modulation order. This can allow easy rate adaption for frequency selective multicarrier non-fading channels and quasi static fading channels. We developed a UB on the BER performance of the considered BICM-T design in terms of signal constellation parameters as well as the multiplexer configuration. Using the developed UB, different

transceiver's modules of the considered BICM-T design can be optimized for performance improvement of such systems. It is shown by the results in this chapter that for higher values of SNR, ES signal constellation becomes non-optimal and to minimize the BER, non-uniform signal constellation and optimum multiplexer configuration should be used. The gains achieved by the system considered in this chapter depend on the code rate and modulation order. The selected numerical result for hierarchical 64-QAM and 1 bit/dimension spectral efficiency showed that by optimizing the modules in our design, gains up to 2.5 dB can be achieved. Moreover these gains are achieved by reducing the complexity of the BICM system.
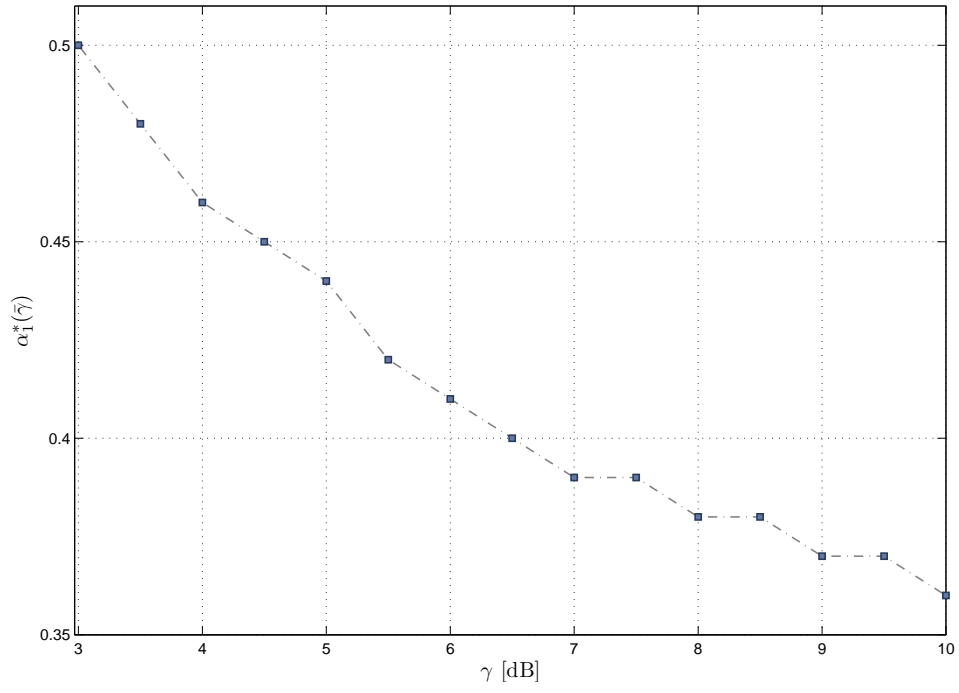
Figure 2.6: Optimal constellation parameter α_1^* for different values of γ [dB] using H-16QAM.
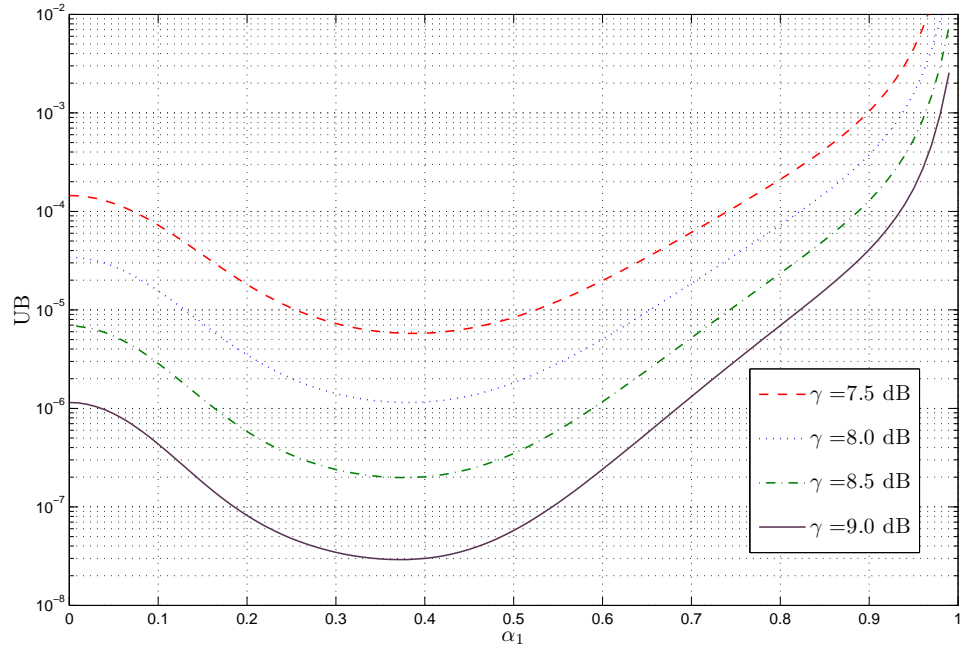
Figure 2.7: UB versus constellation parameter α_1 for different values of γ [dB] using H-16QAM.
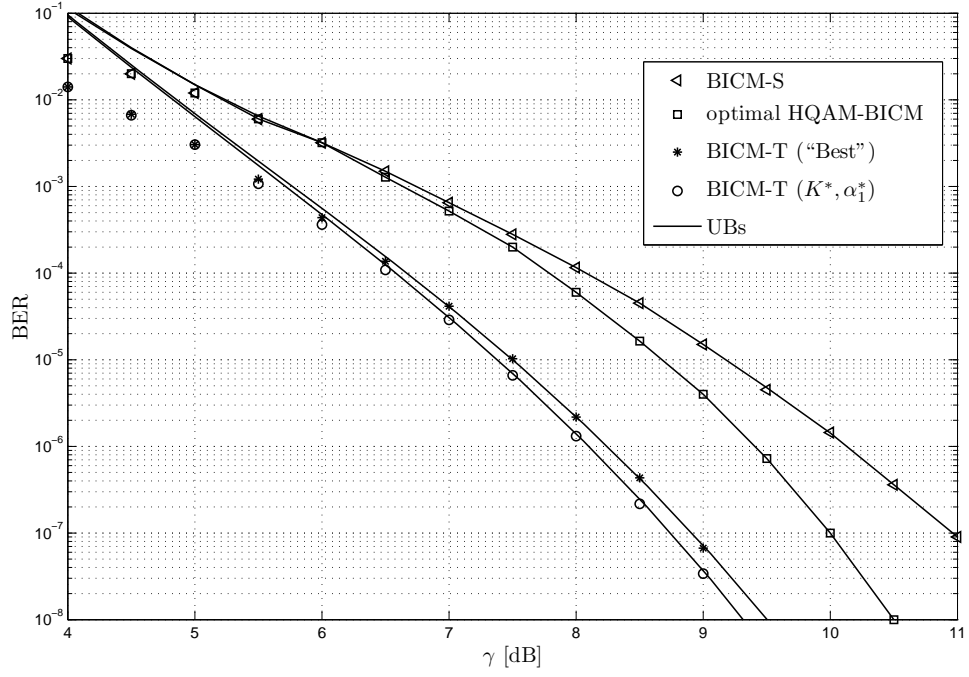
Figure 2.8: BER performance for spectral efficiency 1 bit/dimension using 16-HQAM. The simulations are shown by markers and the UB with solid lines.
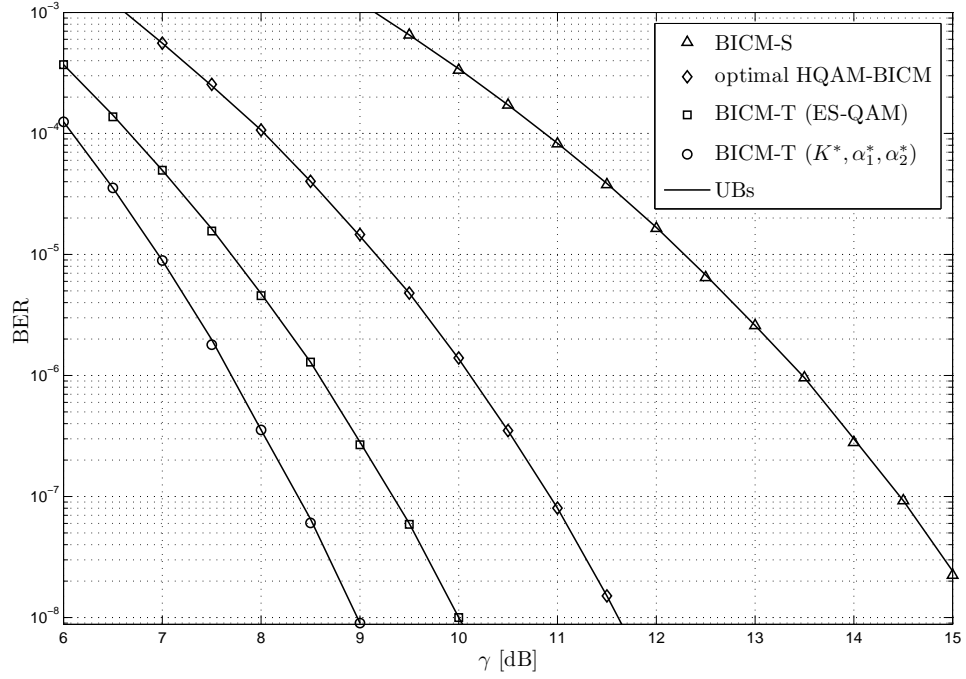
Figure 2.9: BER performance for spectral efficiency 1 bits/dimension using 64-HQAM. The simulations are shown by markers and the UB with solid lines.
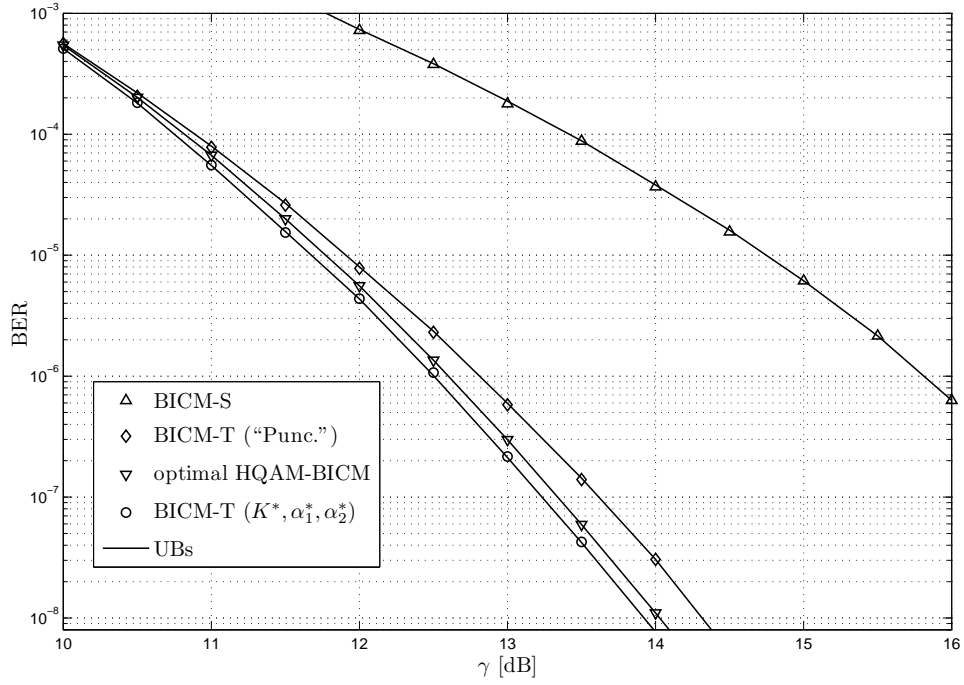
Figure 2.10: BER performance for spectral efficiency 1.5 bits/dimension using 64-HQAM. The simulations are shown by markers and the UB with solid lines.

Chapter 3

Performance of BICM-T Transceivers over Gaussian Mixture Noise Channels[5]

3.1 Introduction

In many physical channels, such as power line communication channels [13], the ambient noise is known through experimental measurements to be non-Gaussian due to the impulsive nature of man-made electromagnetic interference as well as a great deal of natural noise. While the traditional BICM-S and the BICM-T systems have been thoroughly investigated and optimized for the channels with AWGN [9–12, 31], the analysis of BICM transmission impaired by non-Gaussian noise has received relatively little attention [14–16]. In particular there is a need to investigate the performance of BICM-T system for GMN channels.

In BICM-T, a number of consecutive coded bits are transmitted using

---

[5]A part of this chapter has been submitted for publication. M. T. Malik, Md. J. Hossain and M.-S. Alouini, "Analysis and design of generalized BICM-T transceivers", *IEEE Trans. Commun.*, Submitted July 2013.

the same symbol (due to the absence of random interleaver) and the corresponding L-values are mutually dependent. Therefore, a different particularization (e.g., [9]) is adopted to calculate the L-values at the receiver. In particular, using a zero-crossing model [9], closed form expressions for the PDF of L-values are developed in terms of constellation distance parameters of the HQAM. As in [14] it is assumed that the system employs the standard Euclidean-distance decoder, which is an instance of mismatched decoding [17] in the presence of GMN. UB[6] on BER performance of the design proposed in [31] is then developed for GMN channels by using these PDFs, which is then used to optimize the constellation parameters as well as the multiplexer in the BICM-T system over GMN channel. Presented results for 16QAM show that the BICM-T design can offer gains up to 1.5 dB over the traditional BICM design for a target BER value of 10^{-6} if the impulsive components in GMN are below a certain threshold level. However if the impulsive noise is above a threshold, traditional BICM performs better than BICM-T.

The rest of this chapter is organized as follows. In Section 3.2, we present the system model for BICM-T system that uses a bit level multiplexer and hierarchical constellations. Section 3.3 presents the performance evaluation of the system model presented in Section 3.2. Numerical results are presented in Section 3.4. Section 3.5 concludes this chapter.

---

[6]The UBs developed in this chapter are approximated UBs because they are developed using the approximated PDFs of L-values.
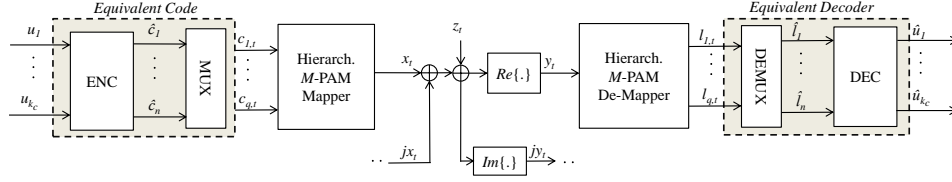
Figure 3.1: System model of the BICM-T transceiver. A channel encoder is followed by the multiplexer (MUX), hierarchical M-PAM modulator, GMN channel, hierarchical M-PAM de-modulator, demultiplexer (DEMUX) and decoder at the receiver side.

3.2 System Model

The system model[7] of HQAM-BICM-T system presented in [31] is shown in Fig. 3.1. The k_c vectors of information bits $\boldsymbol{u}_j = [u_{j,1}, \ldots, u_{j,N_\mathrm{c}}]$ with $j = 1, \ldots, k_\mathrm{c}$ are encoded by a rate $R = k_\mathrm{c}/n$ ENC. The coded bits $\hat{\boldsymbol{c}}_l = [\hat{c}_{l,1}, \ldots, \hat{c}_{l,N_\mathrm{c}}]$ with $l = 1, \ldots, n$ from the encoder are mapped by a multiplexer onto $\boldsymbol{c}_{k,t} = [c_{k,1}, \ldots, c_{k,N_\mathrm{s}}]$ with $k = 1, \ldots, q$, where $nN_\mathrm{c} = qN_\mathrm{s}$. Each output of the MUX is connected to the q-th bit position of a modulator which uses M-ary HQAM constellations labelled by the BRGC [12] to map the bits onto symbols, where $q = \log_2 M$. In BRGC constellation, each symbol can be represented by superposition of independent real/imaginary parts, so we focus on an equivalent HPAM constellation. The bits $\boldsymbol{c}_{k,t}$ from the MUX's output at a time instant t are mapped using a memoryless mapping $\mathcal{M} : \{0,1\}^q \to \mathcal{X}_t$ to a HPAM symbol $x_t \in \mathcal{X}_t = \{x_{t,0}, \ldots, x_{t,M-1}\}$.

The HPAM constellations (e.g., see Fig. 3.2 for 4-HPAM) are defined by the distances d_k with $k = 1, \ldots, q$. The bit value for the bit position

---

[7]This system model is same as that presented in Chapter 2. The only difference is that here we are considering a different noise model.

($k = 1$) selects one of the two squares in Fig. 3.2. For a given value of the first bit position, the bit value for the bit position ($k = 2$) selects one of the two black symbols surrounding the square selected previously which is then transmitted to the receiver. The constellation parameters are defined as $\alpha_k \triangleq d_{k+1}/d_1$, with $k = 1, \ldots, q-1$. The constellation is normalized for unity energy and other constraints on α_k must be added to restrict the constellation to BRGC, as in [12].

The real signal received from the source is $y_t = x_t + z_t$ where z_t is the GMN sample at time instant t. We consider that the GMN samples are uncorrelated over time t. The noise samples are distributed according to the zero-mean Gaussian mixture distribution [14],

$$p_Z(z) = \sum_{n=1}^{N} \frac{\epsilon_n}{2\pi\sigma_n^2} \exp\left(-\frac{|z|^2}{2\sigma_n^2}\right), \tag{3.1}$$

where,

$$\sum_{n=1}^{N} \epsilon_n = 1, \tag{3.2}$$

$$\sum_{n=1}^{N} \epsilon_n \sigma_n^2 = \frac{1}{2\gamma}, \tag{3.3}$$

where γ is the SNR at the receiver. The REC computes the L-values using the received signal which for the bit position k, is given by [9, 31],

$$\tilde{l}_{k,t}(y_t|\boldsymbol{x}_t) = \log \frac{\sum_{a \in \mathcal{X}_{k,1}} p_{y_t}(y_t|x_t = a)}{\sum_{a \in \mathcal{X}_{k,0}} p_{y_t}(y_t|x_t = a)}, \tag{3.4}$$

where $\mathcal{X}_{k,b}$ represents the symbols with bit value b at bit position k. Since
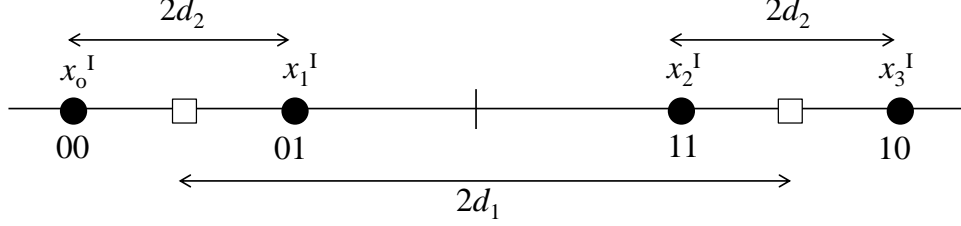
Figure 3.2: HPAM ($M = 4$) constellation with BRGC labelling.

$\tilde{c}_{k,t} = c_{k,t} \oplus s_{k,t}$, we can write $l_{k,t} = (-1)^{s_{k,t}} \tilde{l}_{k,t}$, i.e., to reverse the scrambling, the sign of the L-values is altered by $(-1)^{s_{k,t}}$. This L-value can be approximated using max-log simplification (e.g., see [11]) by,

$$\tilde{l}_{k,t}(y_t|\boldsymbol{x}_t) \approx \gamma \left[\min_{a \in \mathcal{X}_{k,0}} (y_t - a)^2 - \min_{a \in \mathcal{X}_{k,1}} (y_t - a)^2 \right]. \tag{3.5}$$

Although (3.5) does not give the true L-values in the presence of GMN, the use of conventional Euclidean distance metric (3.5) is often considered (for example see [14]) because the optimum maximum likelihood decoding would require the knowledge of the noise PDF or the active mixture component and its variance, which is usually not available at the receiver. A DEMUX then reorders these L-values to invert the MUX's operation. The resulting L-values are then fed to a soft input Viterbi decoder which estimates the transmitted information bits from the source.

3.3 Performance Evaluation

Instead of directly computing the PDFs $p_{\lambda_{\pi_i}}(\lambda|\boldsymbol{c}_{\pi_i,t})$ in (2.18) for GMN channel, we first compute the PDFs conditioned on a given noise compo-

nent. We introduce an auxiliary random variable ξ_t which identifies the noise component n of the PDF (3.1) to which z_t belongs. The distribution of this noise state variable ξ_t is $\Pr\{\xi_t = n\} = \epsilon_n$. The PDF of a component noise random variable Z^{ξ_t} is [14]

$$p_{Z^n}(z) = \frac{1}{2\pi\sigma_n^2}\exp\left(-\frac{|z|^2}{2\sigma_n^2}\right). \tag{3.6}$$

The PDFs $p_{\lambda_{\pi_i}}(\lambda|\boldsymbol{c}_{\pi_i,t})$ in (2.18) can be considered as a weighted sum of PDFs $p_{\lambda_{\pi_i}|n}(\lambda|\boldsymbol{c}_{\pi_i,t})$ conditioned on the state of GMN $\xi_t = n$,

$$p_{\lambda_{\pi_i}}(\lambda|\boldsymbol{c}_{\pi_i,t}) = \sum_{n=1}^{N} \epsilon_n p_{\lambda_{\pi_i}|n}(\lambda|\boldsymbol{c}_{\pi_i,t}). \tag{3.7}$$

In what follows, we develop an UB on the BER of the BICM-T design in terms of constellation distance parameters and MUX configuration for a particular case of rate $1/2$ channel encoder and $(M = 4)$ HPAM. For simplicity of notation and succinct explanation of the procedure to develop PDF of L-values, we only present the analysis for this particular case [31].

Assume $k_c = 1$, $n = 2$, $J = 2$ and $q = 2$ which results in 4 different $\boldsymbol{K}$ matrices [12]. For this particular case, there is a single signal constellation parameter, i.e., α_1. In this case $\Pi = \{1, 2, (1, 2)\}$, so we need to compute three PDFs i.e., $p_{\Lambda_1}(\lambda|\boldsymbol{c}_1)$, $p_{\Lambda_2}(\lambda|\boldsymbol{c}_2)$ and $p_{\Lambda_{1,2}}(\lambda|\boldsymbol{c}_{1,2})$ to calculate the PEP in (2.18). The L-values in (3.5) depend on the transmitted symbol x_t, however, since all zero codeword is transmitted, x_t is totally determined by the scrambling sequence $\boldsymbol{s}_t$. Therefore we can write $\tilde{l}_{k,t}(y_t|\boldsymbol{s}_t)$ to show that the L-values depend on $\boldsymbol{s}_t$. For a given transmitted symbol x_t and noise com-

ponent n, the received signal y_t is a Gaussian random variable with mean x_t and variance σ_n^2. The L-value in (3.5) is a piece-wise linear function of y_t and is therefore piecewise Gaussian. Hence, each L-value $\Lambda(\boldsymbol{e}_t, \boldsymbol{s}_t)$ in (2.16)

$$
p_{\Lambda|n}(\lambda) = \begin{cases} \frac{1}{2}[\psi(\lambda; -\frac{1+\alpha_1}{1-\alpha_1}\rho_n, 2\rho_n) + \psi(\lambda; -\rho_n, 2\rho_n)], & \text{if } \boldsymbol{e}_t = [1,0] \\ \psi(\lambda; -\frac{\alpha_1^2}{(1-\alpha_1)^2}\rho_n, 2\frac{\alpha_1^2}{(1-\alpha_1)^2}\rho_n), & \text{if } \boldsymbol{e}_t = [0,1] \\ \psi(\lambda; -\frac{1}{(1-\alpha_1)^2}\rho_n, 2\frac{1}{(1-\alpha_1)^2}\rho_n), & \text{if } \boldsymbol{e}_t = [1,1] \end{cases}
$$

$$(3.8)$$

$$
[p_{\Lambda_1}(\lambda|c_1 = 0)]^{*w_1} = \left[\frac{1}{2}\sum_{n=1}^{N}\epsilon_n\left[\psi\left(\lambda; -\frac{1+\alpha_1}{1-\alpha_1}\rho_n, 2\rho_n\right) + \psi\left(\lambda; -\rho_n, 2\rho_n\right)\right]\right]^{*w_1}
$$

$$
= \sum_{\mathbf{r}_1 \in \mathcal{W}_{2N}(w_1)} \binom{w_1}{\mathbf{r}_1}\psi\left(\lambda, -\sum_{n=1}^{N}\left(r_{1,2n-1}\frac{1+\alpha_1}{1-\alpha_1} + r_{1,2n}\right)\rho_n, \ldots\right.
$$

$$
\left. \ldots \sum_{n=1}^{N}2(r_{1,2n-1} + r_{1,2n})\rho_n\right)\prod_{n=1}^{N}\left(\frac{\epsilon_n}{2}\right)^{(r_{1,2n-1}+r_{1,2n})}
$$

$$
[p_{\Lambda_2}(\lambda|c_2 = 0)]^{*w_2} = \left[\sum_{n=1}^{N}\epsilon_n\psi\left(\lambda; -\frac{\alpha_1^2}{(1-\alpha_1)^2}\rho_n, 2\frac{\alpha_1^2}{(1-\alpha_1)^2}\rho_n\right)\right]^{*w_2}
$$

$$
= \sum_{\mathbf{r}_2 \in \mathcal{W}_N(w_2)} \binom{w_2}{\mathbf{r}_2}\psi\left(\lambda, -\sum_{n=1}^{N}r_{2,n}\frac{\alpha_1^2}{(1-\alpha_1)^2}\rho_n, \sum_{n=1}^{N}2r_{2,n}\frac{\alpha_1^2}{(1-\alpha_1)^2}\rho_n\right)\prod_{n=1}^{N}\epsilon_n^{r_{2,n}}
$$

$$
[p_{\Lambda_{1,2}}(\lambda|c_{1,2} = 0)]^{*w_{1,2}} = \left[\sum_{n=1}^{N}\epsilon_n\psi\left(\lambda; -\frac{1}{(1-\alpha_1)^2}\rho_n, 2\frac{1}{(1-\alpha_1)^2}\rho_n\right)\right]^{*w_{1,2}}
$$

$$
= \sum_{\mathbf{r}_{1,2} \in \mathcal{W}_N(w_{1,2})} \binom{w_{1,2}}{\mathbf{r}_{1,2}}\psi\left(\lambda, -\sum_{n=1}^{N}r_{(1,2),n}\frac{1}{(1-\alpha_1)^2}\rho_n, \sum_{n=1}^{N}2r_{(1,2),n}\frac{1}{(1-\alpha_1)^2}\rho_n\right)
$$

$$
\cdot\prod_{n=1}^{N}\epsilon_n^{r_{(1,2),n}}.
$$

$$(3.9)$$

is a sum of piece-wise Gaussian functions. The piecewise linear relationships are shown in Fig. 2.3 and 2.4 for 4-HPAM.

To facilitate the analysis, we use the zero-crossing approximation [9, 31] of the L-values which replaces all the Gaussian pieces by one Gaussian function. According to this approximation, $\Lambda(y_t|\boldsymbol{e}_t, \boldsymbol{s}_t, n) \approx \hat{a}(\boldsymbol{e}_t, \boldsymbol{s}_t)y_t + \hat{b}(\boldsymbol{e}_t, \boldsymbol{s}_t)$, where $\hat{a}$ and $\hat{b}$ are determined by the linear piece crossing x-axis and closest to the symbol x_t. In Table 2.1 we present the values of $\hat{a}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ and $\hat{b}(\boldsymbol{e}_t, \boldsymbol{s}_t)$ for 4-HPAM. Using this approximation, we can model the conditional L-values as Gaussian random variables whose mean and variance depend on $\boldsymbol{s}_t$, $\boldsymbol{e}_t$, and γ, i.e., $p_{\Lambda|n}(\lambda|\boldsymbol{s}_t) = \psi(\lambda; \hat{\mu}(\boldsymbol{e}_t, \boldsymbol{s}_t), \hat{\sigma}_n^2(\boldsymbol{e}_t, \boldsymbol{s}_t))$, where $\psi(\lambda; \mu, \sigma^2)$ is a Gaussian PDF with mean μ and variance σ^2. The means and variances are presented in Table 2.2 in terms of constellation parameter. Averaging these conditional Gaussian PDFs over the scrambling sequence $\boldsymbol{s}_t$ results in (3.8)

$$\mathrm{UB} = \frac{1}{k_c} \sum_{l=d_{\mathrm{free}}}^{\infty} \sum_{w_1+w_2+2w_{1,2}=l} \beta_{w_1,w_2,w_{1,2},\boldsymbol{K}}^{\mathbb{C}} \sum_{\mathbf{r}_{\pi_1},\ldots,\mathbf{r}_{\pi_{1,2}}} g(\mathbf{r}_{\pi_1}, \ldots, \mathbf{r}_{\pi_{1,2}})$$

$$Q(h(\mathbf{r}_{\pi_1}, \ldots, \mathbf{r}_{\pi_{1,2}}))$$

where,
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (3.10)

$$g(\mathbf{r}_{\pi_1}, \ldots, \mathbf{r}_{\pi_{1,2}}) = \prod_{\pi_i \in \Pi} \binom{w_{\pi_i}}{\mathbf{r}_{\pi_i}} \prod_{n=1}^{N} \left(\frac{\epsilon_n}{2}\right)^{(r_{1,2n-1}+r_{1,2n})} \prod_{n=1}^{N} \epsilon_n^{(r_{2,n}+r_{(1,2),n})}$$

$$h(\mathbf{r}_{\pi_1}, \ldots, \mathbf{r}_{\pi_{1,2}}) =$$
$$\frac{\sum_{n=1}^{N} -\left(r_{1,2n-1}(1-\alpha_1^2) + r_{1,2n}(1-\alpha_1)^2 + r_{2,n}\alpha_1^2 + r_{(1,2),n}\right)\rho_n}{\left[\sum_{n=1}^{N} 2\left((r_{1,2n-1} + r_{1,2n})(1-\alpha_1)^2 + r_{2,n}\alpha_1^2 + r_{(1,2),n}\right)\rho_n\right]^{\frac{1}{2}}}.$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (3.11)

where for simplicity we have $\rho_n = 2\frac{(1-\alpha_1)^2}{\sigma_n^2(1+\alpha_1^2)}$. Using (3.7) and (3.8), elements in the integration in (2.18) can be derived as (3.9) where we define the set $\mathcal{W}_i(l)$ as all the combinations of i nonnegative integers such that the sum of the elements is l, i.e., $\mathcal{W}_i(l) \triangleq \{(w_1, \ldots, w_i) \in (\mathcal{Z}^+)^i : w_1 + \ldots + w_i = l\}$.

Using (3.9) the expression for UB on BER can be derived as (3.10). The UB is computed considering a truncated weight distribution spectrum of the equivalent code which is calculated using a well known breadth-first search algorithm [12].

3.4   Numerical Results

Numerical results are presented in this section to compare the performance of the traditional BICM and the BICM-T systems for GMN channels. Information bits are encoded using a convolutional code with constraint length $K = 3$, rate $R = 1/2$, and polynomial generators $(5, 7)_8$. A soft-input Viterbi algorithm is used for decoding.

We consider ϵ-mixture noise which is an important instance of the general GMN with two terms, e.g., [32]. The first term represents impulsive noise due to some ambient phenomenon, such as man-made electromagnetic interference, while the second term accounts for the background Gaussian noise. The ϵ-mixture noise parameters can be expressed as [14],

$$\epsilon_1 = \epsilon, \epsilon_2 = 1 - \epsilon,$$

$$\sigma_1 = \frac{\sqrt{\kappa}}{\sqrt{2\gamma(1 + \kappa\epsilon - \epsilon)}}, \tag{3.12}$$

$$\sigma_2 = \frac{1}{\sqrt{2\gamma(1 + \kappa\epsilon - \epsilon)}},$$

where $\kappa = \sigma_1^2/\sigma_2^2$ is a measure for the strength of the impulsive component compared to the thermal noise. In the following, we specify the parameters of ϵ-mixture noise by (ϵ, κ).

Considering $(n = q = 2)$ results in α_1 as the only constellation parameter. A MUX with period $J = 2$ as mentioned in [12] is considered, which results in four different $\boldsymbol{K}$ matrices. A joint optimization was performed over $\boldsymbol{K}$ and α_1. We found that the optimum MUX is always defined by

$$\boldsymbol{K}_\gamma^* = \begin{bmatrix} (2,1) & (2,2) \\ (1,1) & (1,2) \end{bmatrix}, \tag{3.13}$$

as found in [9, 31] as well. This MUX configuration assigns all the coded bits from the first encoder's output to bit position $k = 2$ and all the coded bits from the second encoder's output to bit position $k = 1$.

We compare the performance of the BICM-T with the traditional BICM system for different values of κ and $\epsilon = 0.1$. In Fig. 3.3, we present the UB versus κ for a given value of SNR and using uniform signal constellation. It is interesting to note that depending on the value of κ, BICM-T or BICM-S can be optimal. The results show that if $\kappa < 35$, then the performance of BICM-T is better as compared to the BICM-S design. In Figs. 3.4 and 3.5 we compare the BER performance of the BICM-T with BICM-S system. In Fig. 3.4 we use $\kappa = 10$ and in Fig. 3.5 we use $\kappa = 100$. The Fig 3.4 illustrates that for $\kappa = 10$, the performance of BICM-T using optimal constellation

and multiplexer is much better as compared to BICM-S. On the other hand Fig. 3.5 shows that for $\kappa = 100$, the performance of the traditional BICM-S is better as compared to BICM-T.

3.5   Conclusion

In this chapter, we analyzed the performance of a BICM-T system for non-Gaussian channels. The presented numerical results showed that by properly designing the signal constellation and multiplexer in the BICM-T system, gains up to 1.5 dB over the BICM-S system can be achieved if the strength of impulsive noise components is below a certain threshold level. On the other hand, BICM-S is optimal if the strength of the impulsive noise components is above the threshold.
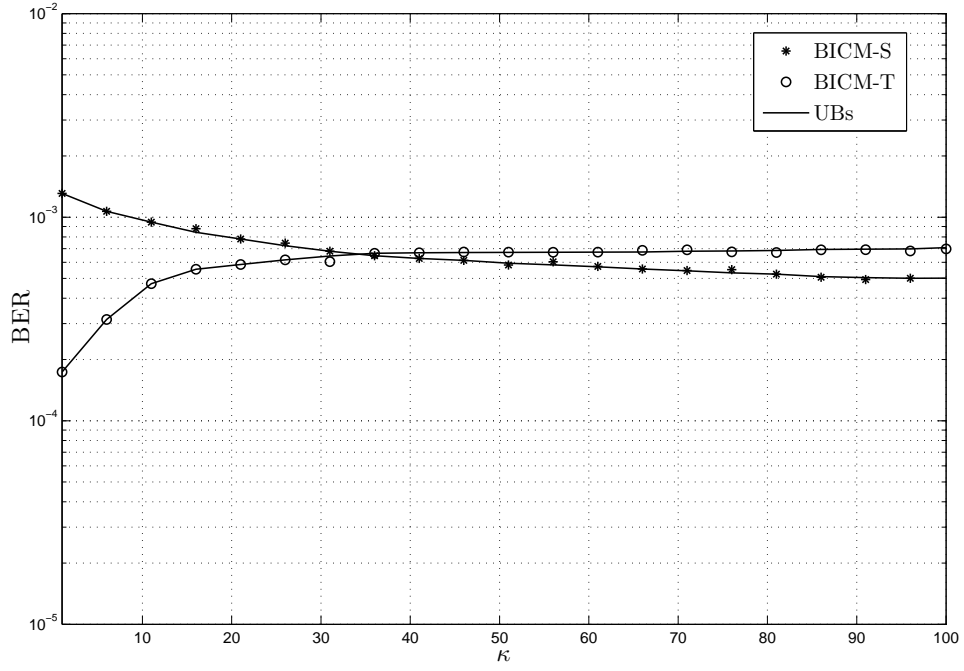
Figure 3.3: Comparison of UB for BICM-T with BICM-S for different values κ, $\epsilon = 0.1$ and $\gamma =8$ dB. The simulations are shown by markers and the UB with solid lines.
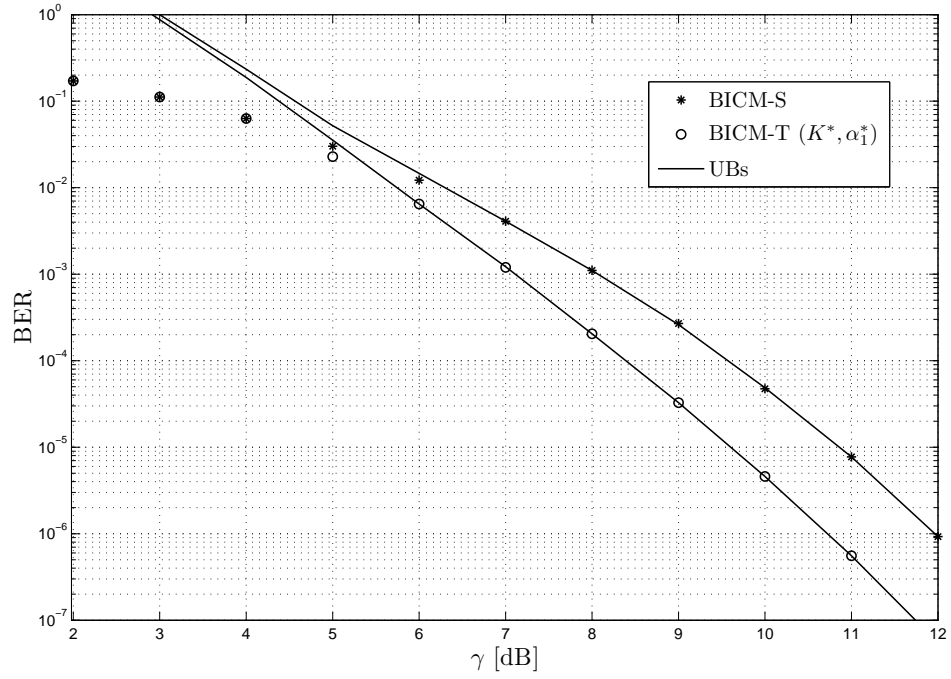
Figure 3.4: BER performance of BICM-T and BICM-S for $\kappa = 10$ and $\epsilon = 0.1$. The simulations are shown by markers and the UB with solid lines.
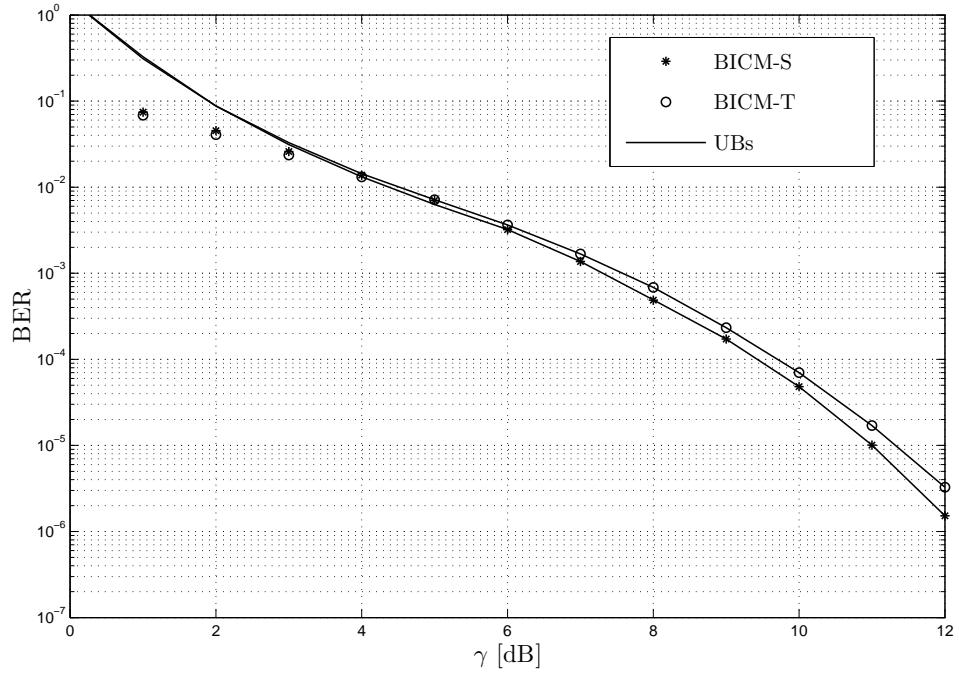
Figure 3.5: BER performance of BICM-T and BICM-S for $\kappa = 100$ and $\epsilon = 0.1$. The simulations are shown by markers and the UB with solid lines.

Chapter 4

BICM-Based Cooperative Communication Systems: Constellation and Interleaver Design[8]

4.1   Introduction

Cooperative communication has gained much interest in recent years as a way to form virtual antenna array among single antenna terminals[9] that improves the overall system performance in wireless environments [18–20]. A smart relay selection scheme proposed in [33] selects one relay from the available relays in the network. This relay selection criterion ensures that the best two-hop relay path is selected between the source and the destina-

---

[8]A part of this chapter has been submitted for publication. M. T. Malik, Md. J. Hossain, and M.-S. Alouini, "BICM-based cooperative communication systems with relay selection: constellation and multiplexer design", *IEEE Wireless Commun. Letters*, Submitted June 2013.

[9]Due to the hardware constraint we consider that communicating terminals have single antenna which is a practical assumption. However our design can be extended for terminals with multiple antennas.

tion, and provides the diversity gain on the order of the number of relays with no loss in performance in terms of diversity-multiplexing tradeoff in orthogonal cooperation. On the other hand, bit-interleaved coded modulation (BICM) [3] offers many advantages in designing a wireless transceiver and is a standard for many contemporary wireless systems. Therefore, a BICM-based transceiver will also be the natural choice for cooperative communication systems. Although a significant research has been conducted towards designing BICM-based transceivers for single-hop systems, however, to date, a little research has been conducted towards designing BICM-based transceiver for cooperative communication systems. A bandwidth efficient BICM-based cooperative communication scheme is proposed in [34]. The performance of BICM-based cooperative communication scheme has also been analyzed [35–37]. None of these works considered the possibility of optimizing various modules of BICM.

Motivated by the results presented in [12] for single-hop communication systems, in this work we propose a new BICM-based cooperative communication system. In [11, 12], it has been shown that the BER performance can be improved if the channel offers an unequal error protection (UEP) to the coded bits and this unequal error protection was achieved by proper design of the multiplexer and signal constellation. In this single-hop communication, in order to optimize the multiplexer and constellation, the transmitter considered the average signal-to-noise ratio (SNR) of the point-to-point link only. However, due to the inherent nature of the cooperative communication system, optimization of the multiplexer and constellation is different from the single-hop systems. For example, when the source transmits information

Figure 4.1: System model of the HQAM-BICM based cooperative transmission: A channel encoder followed by the multiplexer (MUX), the interleavers, the hierarchical M-PAM mapper, detect and forward (DetF) at the selected relay (r_i), and processing blocks at the receiver side.

to the destination and the selected relay (which then forwards the detected symbols to the destination), it has to optimize the multiplexer and constellation jointly considering average SNRs of the direct and the two-hop relay links. Also the relay should optimize its constellation when forwarding the detected symbols to the destination.

In order to exploit the full benefit of BICM in the context of cooperative communication, with the proposed framework different BICM modules can be optimized jointly considering the SNRs of the source-destination channel and the two-hop relay channels. The presented numerical results show that, when the optimized[10] parameters are used in the design, the proposed design can offer gains up to 1.4 dB over the traditional BICM designs for a target bit error rate of 10^{-6}. Moreover the results show that the amount of gain increases with the number of relays available for selection.

---

[10]We obtained the optimal constellations and multiplexer configurations via simulation due to the lack of mathematically tractable probability density functions of L-values for relay channels with higher order signal modulation.

4.2 System Model

The model of our proposed BICM-based cooperative communication system is shown in Fig. 2.1, where a source (s) transmits information to a destination (d) in cooperation with N relay nodes (r_j), with $j = 1, ..., N$. The relay (r_i) is selected from this set of N available relays and works using detect and forward (DetF) strategy. The functionalities of various blocks of the system model are described as follows.

4.2.1 Source

The information bits $\boldsymbol{u}_m = [u_{m,1}, \ldots, u_{m,N_c}]$ with $m = 1, \ldots, k_c$ are encoded by a rate $R = k_c/n$ convolutional encoder (ENC). A bit level multiplexer (MUX) maps the coded bits $\tilde{\boldsymbol{c}}_l = [\tilde{c}_{l,1}, \ldots, \tilde{c}_{l,N_c}]$ with $l = 1, \ldots, n$ from the encoder onto $\boldsymbol{c}_k = [c_{k,1}, \ldots, c_{k,N_s}]$ with $k = 1, \ldots, q$, where $nN_c = qN_s$. The MUX defines a mapping rule and is defined using a $n \times N_c$ matrix $\tilde{\boldsymbol{K}}$, as in [12], whose (l, t')-th entry is a pair (k, t) where $k \in \{1, \ldots, q\}$ and $t \in \{1, \ldots, N_s\}$. The entry (k, t) indicates that the bit $\tilde{c}_{l,t'}$ is assigned to the k-th MUX's output at time instant t. We represent $\tilde{\boldsymbol{K}}$ as a concatenation of N_c/T matrices $\boldsymbol{K}_\tau$, each of dimensions $n \times T$, i.e., $\tilde{\boldsymbol{K}} = [\boldsymbol{K}_0, \ldots, \boldsymbol{K}_{N_c/T-1}]$, where T is the period of MUX. The entries of $\boldsymbol{K}_\tau$ are $(k, t + \tau nT/q)$ where $k \in \{1, \ldots, q\}$ and $t \in \{1, \ldots, nT/q\}$. Without losing generality, $(nT \mod q) = 0$ and $(N_c \mod T) = 0$.

The coded bits $\boldsymbol{c}_k$ are fed to the q parallel interleavers π_k which give independent randomly permuted vectors $\boldsymbol{b}_k = [b_{k,1}, \ldots, b_{k,N_s}]$, i.e., $\boldsymbol{b}_k = \pi_k\{\boldsymbol{c}_k\}$. The q-th interleaver output is connected to the q-th bit position of a modula-

tor which maps the bits onto symbols using hierarchical M-ary quadrature amplitude modulation (HQAM) constellations labelled by the binary reflected Gray code (BRGC) [25], where $q = \log_2 M$. In such constellations, each symbol is a superposition of independently modulated real/imaginary parts, so we focus on an equivalent hierarchical pulse amplitude modulation (HPAM) constellation. At any time instant t, the bits $[b_1, \ldots, b_q]$ are mapped to a HPAM symbol $x_{\mathrm{s}} \in \mathcal{X}_{\mathrm{s}} = \{x_{\mathrm{s},0}, \ldots, x_{\mathrm{s},M-1}\}$ using a binary memoryless mapping $\mathcal{M} : \{0,1\}^q \to \mathcal{X}_{\mathrm{s}}$.

4.2.2 Relay Selection and DetF

The received signal at the j-th relay is given by

$$z_{\mathrm{sr}_j} = a_{\mathrm{sr}_j} \cdot x_{\mathrm{s}} + n_{\mathrm{r}_j} \tag{4.1}$$

where, for a node $\mathrm{w} \in \{\mathrm{r}_j, \mathrm{d}\}$, n_{w} is a zero-mean AWGN with a variance N_0. The channel fading gains between the nodes v and w i.e., $a_{\mathrm{vw}} = \frac{h_{\mathrm{vw}}}{d_{\mathrm{vw}}^{\eta/2}}$ model the path loss and fading, where h_{vw} follow a Rayleigh distribution with a unit average power, η is the path loss exponent and d_{vw} is the distance between v and w. The average SNR between v and w is given by $\overline{\gamma}_{\mathrm{vw}} = \frac{1}{d_{\mathrm{vw}}^{\eta} N_0}$ and the instantaneous SNR $\gamma_{\mathrm{vw}} = \overline{\gamma}_{\mathrm{vw}} |h_{\mathrm{vw}}|^2$ follows an exponential distribution, i.e.,

$$\mathrm{p}(\gamma_{\mathrm{vw}}) = \frac{1}{\overline{\gamma}_{\mathrm{vw}}} \exp\left(-\frac{\gamma_{\mathrm{vw}}}{\overline{\gamma}_{\mathrm{vw}}}\right). \tag{4.2}$$

The two-hop link of relay (r_i) is selected according to the scheme in [33] as,

$$i = \arg \max_{j \in \{1,\ldots,N\}} \min\{\gamma_{\mathrm{sr}_j}, \gamma_{\mathrm{r}_j\mathrm{d}}\}. \tag{4.3}$$

This relay selection criterion provides the diversity gain on the order of the number of relays. From the normalized received signal, $y_{\mathrm{sr}_i} = \{z_{\mathrm{sr}_i}/a_{\mathrm{sr}_i}\}$, selected relay ($r_i$), that works as a detect and forward relay, demodulates the signal (y_{sr_i}) to have estimated coded bits, $[\hat{b}_1, \ldots, \hat{b}_q]$, which are then remodulated using HPAM constellation. The remodulated symbols $x_{\mathrm{r}_i} \in \mathcal{X}_{\mathrm{r}_i} = \{x_{\mathrm{r}_i,0}, \ldots, x_{\mathrm{r}_i,M-1}\}$ are forwarded to the destination. The constellation parameters at the source and the selected relay can be different in general and are denoted by $\boldsymbol{\alpha}_{\mathrm{s}} = [\alpha_{1,\mathrm{s}}, \ldots, \alpha_{(q-1),\mathrm{s}}]$ and $\boldsymbol{\alpha}_{\mathrm{r}_i} = [\alpha_{1,\mathrm{r}_i}, \ldots, \alpha_{(q-1),\mathrm{r}_i}]$, respectively.

4.2.3 Receiver

The destination receives the signal from the source and the selected relay (r_i) as

$$z_{\mathrm{sd}} = a_{\mathrm{sd}} \cdot x_{\mathrm{s}} + n_{\mathrm{d}}$$

$$z_{\mathrm{r}_i\mathrm{d}} = a_{\mathrm{r}_i\mathrm{d}} \cdot x_{\mathrm{r}_i} + n_{\mathrm{d}} \tag{4.4}$$

and the receiver (REC) computes bit-level reliability metrics using the normalized received signals $\underline{\mathbf{y}} = [y_{\mathrm{sd}}, y_{\mathrm{r}_i\mathrm{d}}]$ where $y_{\mathrm{sd}} = \{z_{\mathrm{sd}}/a_{\mathrm{sd}}\}$ and $y_{\mathrm{r}_i\mathrm{d}} =$

$\{z_{\mathrm{r}_i\mathrm{d}}/a_{\mathrm{r}_i\mathrm{d}}\}$. The L-value for the k-th bit position is given by,

$$
\begin{aligned}
\hat{l}_\Sigma^k &= \ln\left\{\frac{\Pr(b_k=1|\underline{\mathbf{y}})}{\Pr(b_k=0|\underline{\mathbf{y}})}\right\} \\
&\approx \ln\left\{\frac{\mathrm{p}(y_{\mathrm{sd}}|b_k=1)\cdot\mathrm{p}(y_{\mathrm{r}_i\mathrm{d}}|b_k=1)}{\mathrm{p}(y_{\mathrm{sd}}|b_k=0)\cdot\mathrm{p}(y_{\mathrm{r}_i\mathrm{d}}|b_k=0)}\right\} \\
&= \hat{l}_{\mathrm{sd}}^k + \hat{l}_i^k,
\end{aligned}
\tag{4.5}
$$

where

$$
\hat{l}_{\mathrm{sd}}^k = \ln\left\{\frac{\mathrm{p}(y_{\mathrm{sd}}|b_k=1)}{\mathrm{p}(y_{\mathrm{sd}}|b_k=0)}\right\},
\tag{4.6}
$$

We assume that $\hat{l}_\Sigma^k$ can be approximated by summation of individual L-values[11].

The L-values for k-th bit position corresponding to selected relay (r_i) is given by,

$$
\begin{aligned}
\hat{l}_i^k &= \ln\left\{\frac{\mathrm{p}(y_{\mathrm{r}_i\mathrm{d}}|b_k=1)}{\mathrm{p}(y_{\mathrm{r}_i\mathrm{d}}|b_k=0)}\right\} \\
&= \ln\left\{\frac{\sum_{c=0}^{1}\mathrm{p}(y_{\mathrm{r}_i\mathrm{d}}|\hat{b}_k=c)\cdot\mathrm{p}(\hat{b}_k=c|b_k=1,\gamma_{\mathrm{sr}_i})}{\sum_{c=0}^{1}\mathrm{p}(y_{\mathrm{r}_i\mathrm{d}}|\hat{b}_k=c)\cdot\mathrm{p}(\hat{b}_k=c|b_k=0,\gamma_{\mathrm{sr}_i})}\right\} \\
&= \ln\left\{\frac{\displaystyle\sum_{x_{\mathrm{r}_i}\in\mathcal{X}_{\mathrm{r}_i}}\sum_{x_{\mathrm{s}}\in\mathcal{X}_{\mathrm{s},k,1}}\mathrm{p}(y_{\mathrm{r}_i\mathrm{d}}|x_{\mathrm{r}_i})\cdot\mathrm{p}(x_{\mathrm{r}_i}|x_{\mathrm{s}},\gamma_{\mathrm{sr}_i})}{\displaystyle\sum_{x_{\mathrm{r}_i}\in\mathcal{X}_{\mathrm{r}_i}}\sum_{x_{\mathrm{s}}\in\mathcal{X}_{\mathrm{s},k,0}}\mathrm{p}(y_{\mathrm{r}_i\mathrm{d}}|x_{\mathrm{r}_i})\cdot\mathrm{p}(x_{\mathrm{r}_i}|x_{\mathrm{s}},\gamma_{\mathrm{sr}_i})}\right\},
\end{aligned}
\tag{4.7}
$$

where $\mathcal{X}_{k,b}$ represents the symbols with bit value b at bit position k. This L-value can be approximated using so called max-log simplification (see for

---

[11]We have also checked that calculation of L-values jointly considering received signals from the source and the selected relay (r_i) i.e., multidimensional observation has a negligible impact on the BER performance.

example [12] [11]) as $\hat{l}_i^k \approx l_i^k$, given by

$$l_i^k = \max_{\substack{x_{\mathrm{s}} \in \mathcal{X}_{\mathrm{s},k,1} \\ x_{\mathrm{r}_i} \in \mathcal{X}_{\mathrm{r}_i}}} \left\{ -\frac{\gamma_{\mathrm{r}_i\mathrm{d}}}{2} (y_{\mathrm{r}_i\mathrm{d}} - x_{\mathrm{r}_i})^2 + \ln\{\mathrm{p}(x_{\mathrm{r}_i}|x_{\mathrm{s}}, \gamma_{\mathrm{sr}_i})\} \right\}$$

$$- \max_{\substack{x_{\mathrm{s}} \in \mathcal{X}_{\mathrm{s},k,0} \\ x_{\mathrm{r}_i} \in \mathcal{X}_{\mathrm{r}_i}}} \left\{ -\frac{\gamma_{\mathrm{r}_i\mathrm{d}}}{2} (y_{\mathrm{r}_i\mathrm{d}} - x_{\mathrm{r}_i})^2 + \ln\{\mathrm{p}(x_{\mathrm{r}_i}|x_{\mathrm{s}}, \gamma_{\mathrm{sr}_i})\} \right\}. \qquad (4.8)$$

The combined metrics are then de-interleaved to obtain $\tilde{l}_\Sigma^k = \pi_k^{-1}\{\hat{l}_\Sigma^k\}$ where $k = 1, \ldots, q$. These L-values are reorganized by the demultiplexer unit (DEMUX) which inverts the operation done by the MUX. The resulting vectors of L-values are fed to a soft input Viterbi decoder which estimates the transmitted information bits from the source.

4.3 Simulation Results

In this section, we present numerical results where information bits are encoded using a rate $R = 1/2$ convolutional code with constraint length $K = 3$ and generator polynomial $(5, 7)_8$. A soft-input Viterbi algorithm is used for decoding.

For an optimal BER performance, one needs to jointly optimize $\boldsymbol{K}$, the constellation parameters at the source i.e., $\boldsymbol{\alpha}_{\mathrm{s}}$, and at the selected relay i.e., $\boldsymbol{\alpha}_{\mathrm{r}_i}$, for a given value of $\bar{\gamma}_{\mathrm{sd}}$ and all relay positions (equivalently average SNRs between the source and the relays as well as average SNRs between the relays and the destination). The optimum values are defined as $\boldsymbol{\alpha}_{\mathrm{s}}^*$, $\boldsymbol{\alpha}_{\mathrm{r}_i}^*$

and $\boldsymbol{K}^*$ i.e.,

$$[\boldsymbol{\alpha}_{\mathrm{s}}^*, \boldsymbol{\alpha}_{\mathrm{r}_i}^*, \boldsymbol{K}^*] = \underset{\boldsymbol{\alpha}_{\mathrm{s}}, \boldsymbol{\alpha}_{\mathrm{r}_i}, \boldsymbol{K}}{\operatorname{argmin}} \left\{ \mathrm{BER}(\boldsymbol{\alpha}_{\mathrm{s}}, \boldsymbol{\alpha}_{\mathrm{r}}, \boldsymbol{K}) \right\}. \tag{4.9}$$

The optimization of BER was performed via simulation and an exhaustive search over the valid range of constellation parameters, with a step size of 0.01 as done in [12] for single-hop channels. Please note that this optimization should be performed at the beginning based on the cooperative network's topology i.e., $\bar{\gamma}_{\mathrm{sd}}$ and all cooperating relays' positions (equivalently average SNRs between the source and the relays as well as average SNRs between the relays and the destination). Then the results can be used for the cooperative system under consideration as long as the topology does not change. This optimization can also be performed offline and the results can be tabulated to be used for the cooperative system under consideration.

4.3.1 Spectral Efficiency 1 Bit/Dimension

For this case ($n = q = 2$), the HPAM constellation has one parameter. Therefore, one constellation parameter at the source $\alpha_{1,\mathrm{s}}$ and at the selected relay node i.e., α_{1,r_i} is to be optimized. We consider a MUX with period $T = 2$, which results in four different $\boldsymbol{K}$ matrices. A numerical optimization was performed over $\boldsymbol{K}, \alpha_{1,\mathrm{s}}$ and α_{1,r_i}, and different positions of the N relays. We found that the optimum MUX is always given by

$$\boldsymbol{K}^* = \begin{bmatrix} (2,1) & (2,2) \\ (1,1) & (1,2) \end{bmatrix}. \tag{4.10}$$

In order to illustrate how relay position affects the choice of constellation parameter, we consider a simple topology with one relay where the relay r_1 lies on the line between the source and the destination and its position is determined by $\delta = d_{sr_i}/d_{sd}$; where $d_{sd} = 1$. The results of joint optimization over $\alpha_{1,s}$ and α_{1,r_1} are presented in Fig. 4.2 and Fig. 4.3, respectively for three different relay locations. It is interesting to observe that the optimum constellation parameters at source and relay nodes not only depend on the average source-destination SNR ($\bar{\gamma}_{sd}$), but also on the location of the cooperating relay. If the relay is placed close to the source, the optimal constellation at the source tends to be non-uniform.

For a given number of relays and their positions, a joint optimization over $\boldsymbol{K}$, $\alpha_{1,s}$ and α_{1,r_i} is performed and the optimal values are then used in our proposed design. The BER performance of the proposed system with optimized parameters is presented in Fig. 4.4. In this, we have also plotted the BER performances of the traditional single interleaver BICM system (BICM-S) [37]. The results show that the proposed design with joint optimization of the MUX and the constellation outperforms the traditional design. They also indicate that the amount of gain increases with the number of relays i.e., the diversity order as the fading becomes less severe.

4.3.2 Spectral Efficiency 1.5 Bits/Dimension

Considering 8-ary constellations ($q = 3$) and rate $R = 1/2$ code ($n = 2$) gives a spectral efficiency of 1.5 bits/dimension. In this case, the HPAM constellation has two parameters, therefore there are total 4 parameters in-

---

[12]We considered a convolutional code with $K = 8$ and generator polynomials $(371, 247)_8$.

Table 4.1: Optimal constellation parameters and gains for BER $\approx 10^{-6}$

N	$(d_{\mathrm{sr}_j}, d_{\mathrm{r}_j\mathrm{d}})$	$\boldsymbol{\alpha}_{\mathrm{s}}^*$	$\boldsymbol{\alpha}_{\mathrm{r}_i}^*$	$\boldsymbol{K}^*$			gain [dB]
2	(0.5, 0.5) (0.7, 0.7)	(0.47, 0.18)	(0.46, 0.18)	$\begin{bmatrix} (2,1) & (3,2) & (3,1) \\ (2,2) & (1,1) & (1,2) \end{bmatrix}$			1.0
3	(0.5, 0.5) (0.7, 0.7) (0.4, 0.6)	(0.45, 0.17)	(0.46, 0.18)	$\begin{bmatrix} (2,1) & (3,2) & (3,1) \\ (2,2) & (1,1) & (1,2) \end{bmatrix}$			1.2
5	(0.5, 0.5) (0.7, 0.7) (0.4, 0.6)	(0.45, 0.16)	(0.46, 0.16)	$\begin{bmatrix} (1,2) & (2,1) & (3,1) \\ (3,2) & (2,2) & (1,1) \end{bmatrix}$			1.4
	(0.6, 0.4) (0.7, 0.7)	(0.48, 0.22)	(0.48, 0.21)	$\begin{bmatrix} (3,1) & (1,1) & (2,2) \\ (1,2) & (3,2) & (2,1) \end{bmatrix}$			0.9^{12}

cluding $(\alpha_{1,\mathrm{s}}, \alpha_{2,\mathrm{s}})$ at the source and two constellation parameters at the selected relay node i.e., $(\alpha_{1,\mathrm{r}_i}, \alpha_{2,\mathrm{r}_i})$ that are to be optimized. We consider a MUX with period $T = 3$, which results in thirty different MUX configurations [12].

In Fig. 4.5, we present the BER performance of the proposed system with the optimal constellation parameters and multiplexer configuration. The results show that the proposed system outperforms the traditional BICM-S system [37]. In Table 4.1, we present the optimal constellation parameters, optimal multiplexer and the amount of gain achieved over the BICM-S system for different positions of N relays and for a target BER of 10^{-6}. From this table, it is interesting to observe that the constellation parameters as well as the optimum multiplexer depends on the number and the location of the relays as well as the convolutional code. As shown by the presented results, the amount of gain increases with the number of relays and is 1.4
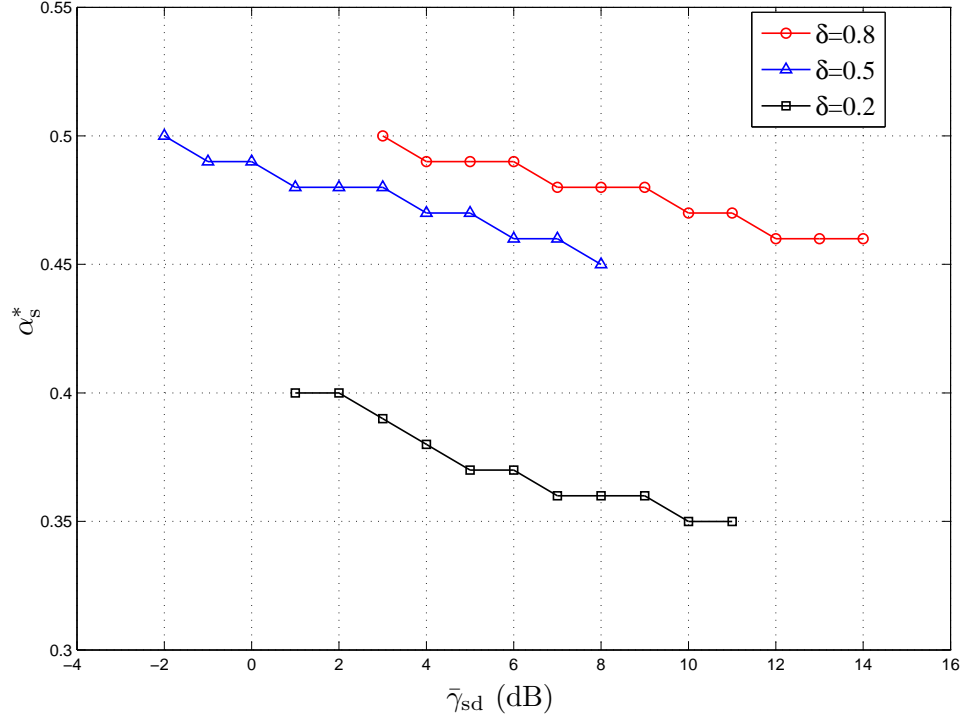
Figure 4.2: Optimal constellation parameter at the source versus average source-destination SNR ($\bar{\gamma}_{sd}$), that gives BER $\in [10^{-8}, 10^{-2}]$, for different relay positions identified by $\delta = d_{sr_1}/d_{sd}$.

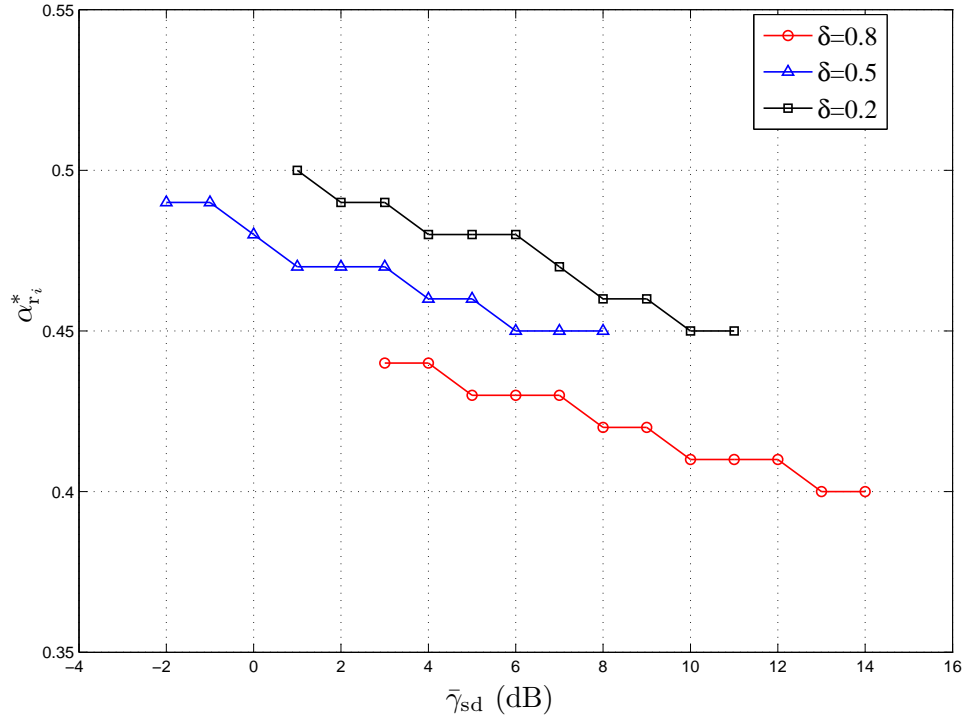dB using $N = 5$ cooperating relays for a target BER of 10^{-6}.

Figure 4.3: Optimal constellation parameter at the relay versus average source-destination SNR ($\bar{\gamma}_{sd}$), that gives BER $\in [10^{-8}, 10^{-2}]$, for different relay positions identified by $\delta = d_{sr_1}/d_{sd}$.

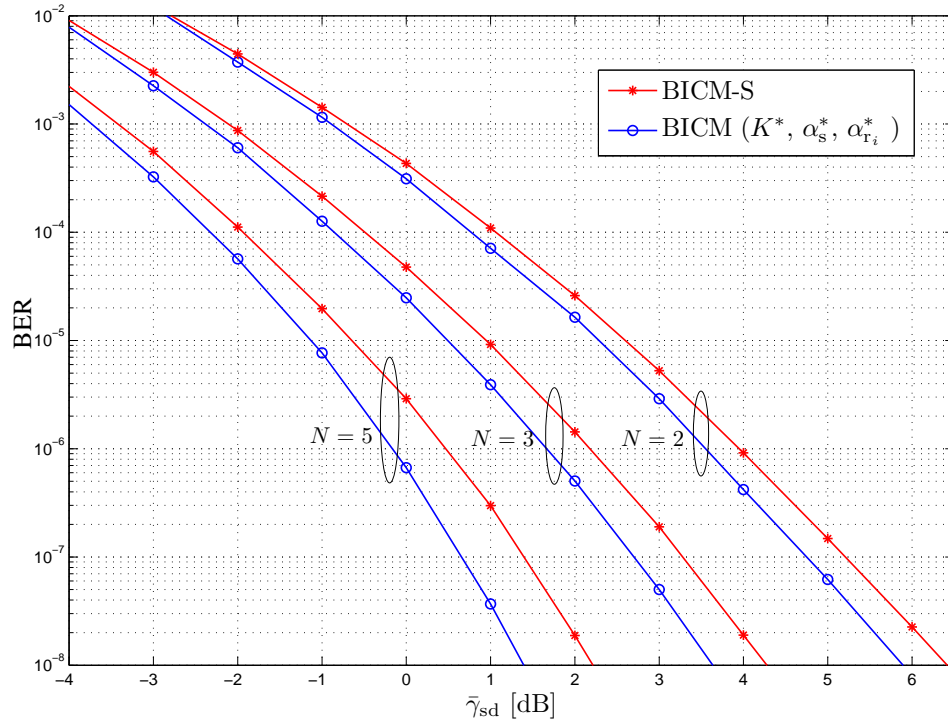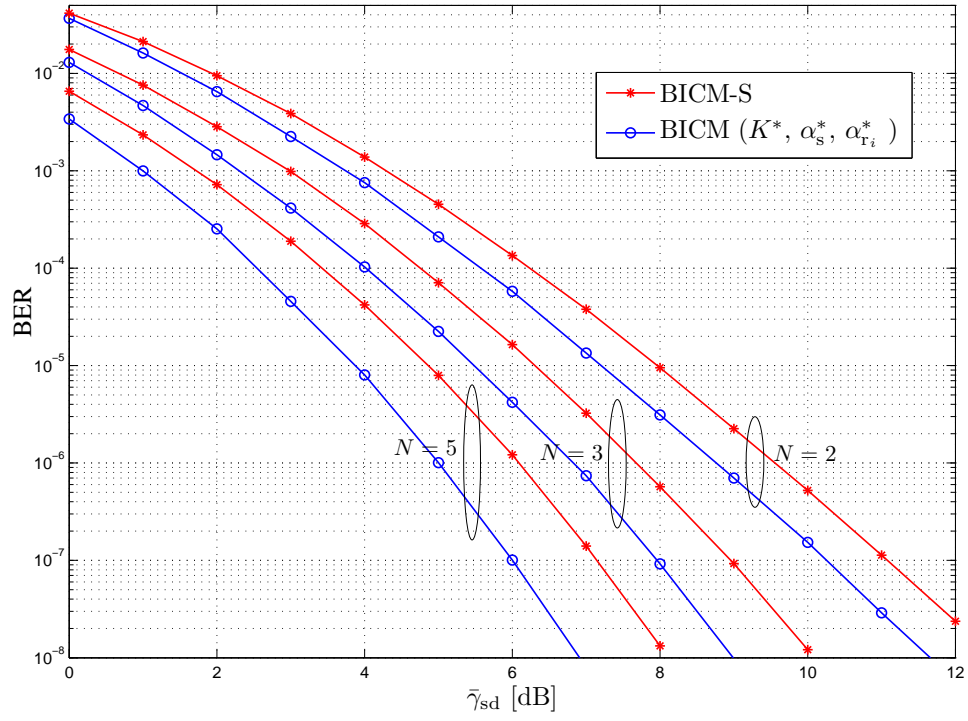Figure 4.4: BER performance for spectral efficiency 1 bit/dimension.

Figure 4.5: BER performance for spectral efficiency 1.5 bits/dimension.

Chapter 5

Summary, Conclusion and Future Works

5.1 Summary and Conclusions

In this thesis, we investigated and optimized new BICM-based transceivers to improve the reliability of the transmission.

In Chapter 2, we presented a generalized BICM-T design that uses HQAM signal constellations in conjunction with a bit-level multiplexer. Due to the use of a bit level multiplexer that connects the channel encoder with the modulator, our considered design gives a design flexibility to independently choose the code rate and the modulation order. This can allow easy rate adaption for frequency selective multicarrier non-fading channels and quasi static fading channels. We developed an UB on the BER performance of our considered BICM-T design in terms of signal constellation parameters as well as the multiplexer configuration. Using the developed UB, different transceiver's modules of the considered BIMC-T design can be optimized for performance improvement of such systems. It is also shown by the results in the chapter that for higher values of SNR, ES signal constel-

lation becomes non-optimal and to minimize the BER, non-uniform signal constellation and optimum multiplexer configuration should be used in the BICM-T design. The selected numerical result for hierarchical 64-QAM and 1.5 bits/dimension spectral efficiency showed that by optimizing the modules in the considered design, gains up to 3.4 dB can be achieved with respect to BICM-S.

In Chapter 3, we investigated the performance of BICM-T in non Gaussian channels due to its practical relevance. The presented numerical results showed that by properly designing the signal constellation and multiplexer in the BICM-T system, gains up to 1.5 dB over the BICM-S system can be achieved if the strength of impulsive noise components is below a certain threshold level. On the other hand, BICM-S is optimal if the strength of the impulsive noise components is above the threshold.

In Chapter 4, we presented an innovative BICM design for cooperative communication where various BICM modules can be optimized depending on the SNRs of different two-hop relay channels. Our proposed a BICM-based transmission framework for cooperative communication systems uses HQAM constellations in conjunction with a deterministic bit-level multiplexer and M-interleavers. Presented results showed that the BER performance is improved with our proposed design as compared to the other designs by optimal selection of the multiplexer and constellation.

5.2   Future Work

There are various open avenues for future research in the topics related to BICM transmission. In the following, an extension of the work in this thesis and a potential research topic is discussed briefly.

5.2.1   Performance Bit-Interleaved Coded Modulation over Gaussian Mixture Noise Channels with Higher Order Modulation

In many physical channels the ambient noise is known through experimental measurements to be non-Gaussian due to the impulsive nature of man-made electromagnetic interference as well as a great deal of natural noise. In Chapter 3 we analyzed and optimized the performance of BICM-T over GMN channels and showed that by properly designing the signal constellation and multiplexer in the BICM-T system, gains up to 1.5dB over the traditional single interleaver BICM systems can be achieved if the strength of impulsive noise components is below a certain threshold level. But this design and analysis considered a simple case of 16-ary QAM. The analysis for a more generalized setup with higher order signal modulation can be done using the similar approach. For such modulation, it will be interesting to see how much gain we can achieve by using an optimal BICM-T design as compared to the traditional BICM-S and under which conditions.

5.2.2 Performance of Bit-Interleaved Coded Modulation over Gamma-Gamma Turbulence Channels

Optical wireless communication (OWC) system has gained increasing attention in recent years because of its ability to offer high speed data transmission over short distances, and to replete the gap between backbone fiber network infrastructure and end users. The main advantage of OWC system is its optical-fiber-like performance with easy and cost effective deployment, being free from spectrum licensing regulation. Atmospheric turbulence induced random irradiance fluctuation is a major source of performance degradation for any outdoor OWC systems. Several fading mitigation techniques have been proposed for OWC systems. In [38], the authors proposed error control coding in conjunction with interleaving to mitigate strong turbulence fading. Due to slow fading nature of OWC channels, such system requires large interleavers. However, recent developments in hardware and signal processing have made such interleaving possible (for example see [39]).

It is very interesting to investigate the performance of BICM over optical wireless turbulence channels. The Gamma-Gamma distribution is a well accepted statistical model for describing turbulence induced irradiance fluctuation over a wide range of turbulence regimes from weak to strong turbulence. Developing an analytical framework for this performance evaluation of different BICM designs will be very useful to optimize various BICM modules to improve the BER performance for such channels.

Bibliography

[1] S. Lin and D. J. Costello, "Error control coding." 2nd ed., Inc. Upper Saddle River, NJ: Prentice-Hall, 2004. → pages 1

[2] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982. → pages 1, 8

[3] E. Zehavi, "8-psk trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 3, pp. 873–884, May 1992. → pages 2, 5, 8, 20, 65

[4] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998. → pages 2, 8, 9, 20, 39, 40, 42

[5] 3GPP, "Universal mobile telecommunications system; multiplexing and channel coding," 3GPP, Tech. Rep. TS 125.212, V7.11.0 Release 7, Sep. 2009. → pages 2

[6] IEEE 802.11, "Part 11: Wireless LAN medium access control and physical layer specifications: High-speed physical layer in the 5GHz band," IEEE Std 802.11a-1999 (R2003), Tech. Rep., July 1999. → pages 2

[7] IEEE 802.11n, "Part 11:wireless LAN medium access control and physical layer specifications. amendment 5: Enhancements for higher throughout," IEEE Std 802.11n-2009, Tech. Rep., Oct. 2009. → pages 2

[8] ETSI, "Digital video broadcasting; frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system," ETSI, Tech. Rep. ETSI EN 302 755 V1.1.1 (2009-09), Sep. 2009. → pages 2

[9] A. Alvarado, L. Szczecinski, and E. Agrell, "On BICM receivers for TCM transmission," *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 2692–2702, Oct. 2011. → pages 3, 4, 5, 9, 10, 16, 17, 18, 19, 21, 25, 32, 39, 50, 51, 53, 57, 59

[10] C. Stierstorfer, R. F. H. Fischer, and J. B. Huber, "Optimizing BICM with convolutional codes for transmission over the AWGN channel," *in Proc. of International Zurich Seminar Commun.*, March 2010. → pages 3, 9

[11] A. Alvarado, E. Agrell, L. Szczecinski, and A. Svensson, "Exploiting UEP in QAM-based BICM: Interleaver and code design," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 500–510, Feb. 2010. → pages 3, 9, 16, 54, 65, 71

[12] Md. J. Hossain, A. Alvarado, and L. Szczecinski, "Towards fully optimized BICM transceivers," *IEEE Trans. Commun.*, vol. 59, no. 11, pp.

3027–3039, Nov. 2011. → pages 3, 4, 9, 10, 12, 14, 16, 39, 40, 41, 42, 50, 52, 53, 55, 58, 59, 65, 67, 71, 72, 74

[13] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broadband powerline communications," *IEEE Trans. Electro-magn. Compat.*, vol. 44, no. 1, pp. 249–258, Feb. 2002. → pages 4, 50

[14] A. K.-Anhari, L. Lampe, "Performance analysis for BICM transmission over gaussian mixture noise fading channels," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 1962–1972, July 2010. → pages 4, 5, 50, 51, 53, 54, 55, 58

[15] T. Li, W. Mow, and M. Siu, "Bit-interleaved coded modulation in the presence of unknown impulsive noise," *IEEE International Conf. Commun. (ICC)*, vol. 7, pp. 3014–3019, June 2006. → pages

[16] A. Nasri and R. Schober, "Performance of BICM-SC and BICM-OFDM systems with diversity reception in non-Gaussian noise and interference," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3316–3327, Nov. 2009. → pages 4, 50

[17] A. Martinez, A. Guillén i Fàbregas, and G. Caire, "Bit-interleaved coded modulation revisited: a mismatched decoding perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2756–2765, June 2009. → pages 5, 51

[18] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diverstiy Part I and Part II," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1948, Nov. 2003. → pages 5, 64

[19] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, "An efficient protocol for realizing cooperative diversity in wireless networks," *in Proc. of IEEE International Symposium on Information Theory (ISIT'01)*, p. 294, Washington, DC, Jun. 2001. → pages

[20] T. E. Hunter and A. Nosratinia, "Cooperative diversity through coding," *in Proc. of IEEE International Symposium on Information Theory (ISIT'02)*, p. 220, Laussane, Switzerland, Jul. 2002. → pages 5, 64

[21] H. Imai and S. Hirakawa, "A new multilevel coding method using error-correcting codes," *IEEE Trans. Inf. Theory*, vol. 23, no. 3, pp. 371–377, May 1997. → pages 8

[22] I. Kalet, "The multitone channel," *IEEE Trans. Commun.*, vol. 37, no. 2, pp. 119–124, Feb. 1989. → pages 10

[23] C. Shen, T. Liu, M. P. Fitz, "On the average rate performance of Hybrid ARQ in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3339–3352, Nov. 2009. → pages 10

[24] A. Alvarado, L. Szczecinski, R. Feick, and L. Ahumada, "Distribution of L-values in Gray-mapped M^2-QAM: Closed-form approximations and applications," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 2071–2079, July 2009. → pages 10

[25] E. Agrell, J. Lassing, E. G. Strom, and T. Ottosson, "On the optimality of the binary reflected Gray code," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3170–3182, Dec. 2004. → pages 12, 68

[26] P. K. Vitthaladevuni and M.-S. Alouini, "A recursive algorithm for the exact BER computation of generalized hierarchical QAM constellations," *IEEE Trans. Inf. Theory*, vol. 49, pp. 297–307, Jan. 2003. → pages 14

[27] A. J. Viterbi, "An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 260–264, Feb. 1998. → pages 17, 24

[28] J. Belzile and D. Haccoun, "Bidirectional breadth-first algorithms for the decoding of convolutional codes," *IEEE Trans. Commun.*, vol. 41, no. 2, pp. 370–380, Feb. 1993. → pages 24

[29] M. Benjillali, L. Szczecinski, S. Aissa, and C. Gonzalez, "Evaluation of bit error rate for packet combining with constellation rearrangement," *Wiley Journal Wireless Comm. and Mob. Comput.*, no. (published online), available at http://doi.wiley.com/10.1002/wcm.528, June 2007. → pages 27

[30] P. Frenger, P. Orten, and T. Ottosson, "Convolutional codes with optimum distance spectrum," *IEEE Commun. Letters*, vol. 3, no. 11, pp. 317–319, Nov. 1999. → pages 38, 42

[31] M. T. Malik, J. Hossain, and M.-S. Alouini, "Generalized BICM-T transceivers: constellation and multiplexer design," *IEEE Intl. Sym.*

Personal, Indoor and Mobile Radio Networks, Accepted for Publication, June 2013. → pages 50, 51, 52, 53, 55, 57, 59

[32] X. Wang and H. V. Poor, "Robust multiuser detection in non-Gaussian channels," *IEEE Trans. signal processing*, vol. 47, no. 2, pp. 3316–3327, Feb. 1999. → pages 58

[33] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006. → pages 64, 69

[34] H. T. Nguyen, H. J. Nguyen, and T. Le-Ngoc, "Bandwidth-efficient bit-intelerleaved coded cooperative communications," *Wireless Personal Communications*, vol. 46, no. 3, pp. 247–248, Aug. 2008. → pages 65

[35] S. A. K. Tanoli, I. Khan, and N. Rajatheva, "Cooperative networks: bit-interleaved code modulation with iterative decoding," *Journal of Communications*, vol. 5, no. 7, pp. 562–570, July 2010. → pages 65

[36] T. Islam, R. Schober, R. K. Mallik, and V. K. Bhargava, "Analysis and design of cooperative BICM-OFDM systems," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1742–1751, June 2011. → pages

[37] M. Benjillali and L. Szczecinski, "Detect-and-forward in two-hop relay channels: A metrics-based analysis," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1729–1736, June 2010. → pages 65, 73, 74

[38] M. Uysal, S. M. Navidpur, and J. Li, "Error rate performance of coded free-space optical links over strong turbulence channels," *IEEE Commun. Letter*, vol. 8, no. 3, pp. 635–637, Oct. 2004. → pages 82

[39] Greco, Joseph A. , "Design of the high-speed framing, FEC, and interleaving hardware used in a 5.4km free-space optical communication experiment." *SPIE–The International Society for Optical Engineering*, Free-Space Laser Communications IX. Ed. Arun K. Majumdar & Christopher C. Davis. San Diego, CA, USA 2009. → pages 82