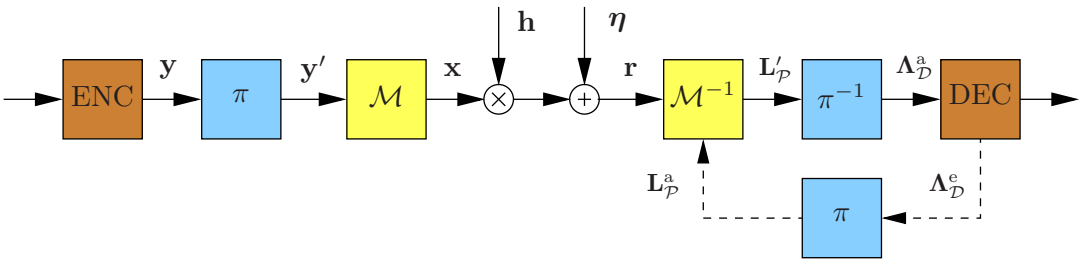
CHALMERS



On Bit-interleaved Coded Modulation with QAM Constellations

ALEX ALVARADO

Department of Signals and Systems
Communication Systems Group
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2008

Thesis for the degree of Licentiate of Engineering

On Bit-interleaved Coded Modulation with QAM Constellations

Alex Alvarado

CHALMERS

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology

Göteborg, Sweden 2008

Alvarado, Alex
On Bit-interleaved Coded Modulation with QAM Constellations

Department of Signals and Systems
Technical Report No. R009/2008
ISSN 1403-266X

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Göteborg, Sweden
Telephone: + 46 (0)31-772 1000

This thesis has been prepared using LaTeX.

Front Cover:
Block diagram of Bit-interleaved Coded Modulation with iterative decoding.

Printed by Chalmers Reproservice,
Göteborg, Sweden, May 2008.

To Mirta, Luis, Rodrigo
and M. Regina (QEPD)

Abstract

Bit-interleaved coded modulation (BICM) is a flexible modulation/coding scheme which allows the designer to choose a modulation constellation independently of the coding rate. This is because the output of the channel encoder and the input to the modulator are separated by a bit-level interleaver. In order to increase spectral efficiency, BICM can be combined with high-order modulation schemes such as quadrature amplitude modulation (QAM) or phase shift keying. BICM is particularly well suited for fading channels, and it only introduces a small penalty in terms of channel capacity when compared to the coded modulation capacity for both additive white Gaussian noise and fading channels. Additionally, if the so-called BICM with iterative decoding (BICM-ID) is used, the demapper and decoder iteratively exchange information, improving the system performance.

At the receiver's side of BICM, the reliability metrics are calculated for the coded bits under the form of logarithmic likelihood ratios, or simply *L-values*. These metrics are then deinterleaved and further used by the soft-input channel decoder. This thesis deals with the probabilistic characterization of the L-values calculated by the demapper when BICM is used in conjunction with high order QAM schemes. Three contributions are included in this thesis.

In **Paper A** the issue of the probabilistic modelling of the extrinsic L-values for BICM-ID is addressed. Starting with a simple piece-wise linear model of the L-values obtained via the max-log approximation, expressions for the probability density functions (PDFs) for Gray-mapped 16-QAM are found. The developed analytical expressions are then used to efficiently compute the so-called extrinsic information transfer functions of the demapper, and they are also compared with the histograms of the L-values obtained through numerical simulations.

In **Paper B** closed-form expressions for the PDFs of the L-values in BICM with Gray mapped QAM constellations are developed. Based on these expressions, two simple Gaussian mixture approximations that are analytically tractable are also proposed. The developments are used to efficiently calculate the BICM channel capacity and to develop bounds on the coded bit-error rate when a convolutional code is used. The coded performance of an hybrid automatic repeat request based on constellation rearrangement is also evaluated.

In **Paper C** closed-form expressions for the PDFs of the L-values in BICM transmissions with Gray-mapped QAM constellations over fully-interleaved fading channels are derived. The results are particularized for a Rayleigh fading channel, however, developments for the general case of a Nakagami-m case are also included. Using the developed expressions, the performance of BICM transmissions using convolutional and turbo codes is efficiently evaluated. The BICM channel capacity for different fading channels and constellation sizes is also calculated.

Acknowledgements

First of all, I would like to express my deepest gratitude to Professor Arne Svensson for being my supervisor and for giving me the opportunity of becoming a PhD student at the Communication Systems Group. This gratitude goes also to Professor Erik Ström for having faith in me. I really feel proud for being part of this group formed by so many bright people. I am also very grateful to have Associate Professor Erik Agrell as my co-supervisor, always so open to discuss interesting/fun problems, and always so right about everything.

I cannot miss this opportunity to mention my former advisors in the other side of the world. Many thanks to Professor Leszek Szczeciński at INRS-EMT Montreal Canada, and Professor Rodolfo Feick at UTFSM Valparaíso Chile. A big part of the work included in this thesis was carried out while I was working with them, and I very much appreciate the "introduction" they gave me about on how to do good research.

I thank the current and former members of the Communication Systems Group at Chalmers. Special thanks go to Dr. Andreas Wolfgang, Guillermo García, Daniel Persson and Dr. Florent Munier for their friendship and also for their help translating food names from Swedish to English/Spanish.

For helping me with all the administrative tasks, I would like to send my gratitude to Agneta, Madeleine and Cecilia. My gratitude also goes to Lars for arranging all the computer-related stuff.

From outside Chalmers I would like to thank the "Swedish-course" people: Renata, Jessica, Petra and Alice, and the *helyllesvenskar* Tomas, Annica, Daniel and Katarina. I would also like to thank my "Canadian" friends Eric, Annemarie, Sami and Desi for the great time we spent in Montreal during 2006. Last but not least, a big thanks to *mijn lieve meisje* who gave me support during all this time and who always asked me about the meaning of convolutional codes, QAM and BICM. I was never able to give you a proper answer, but I hope this thesis will do it.

Alex Alvarado
Göteborg, May of 2008

This research was partially supported by the Swedish Research Council, Sweden (under research grant #2006-5599), by NSERC, Canada (under research grant #249704-07), by Fondecyt (under Projects No. 1070742, No. 707063, and PBCT-ACT-11/2004), and by Project Semilla-UDP, Chile.

Contents

Acronyms

AMI:	Average Mutual Information
AWGN:	Additive White Gaussian Noise
ARQ:	Automatic Repeat-reQuest
BER:	Bit Error Rate
BCJR:	Bahl Cocke Jelinek and Raviv
BICM:	Bit-interleaved Coded Modulation
BICM-ID:	BICM with Iterative Decoding
BISO:	Binary Input Soft Output
BPSK:	Binary Phase Shift Keying
BRGC:	Binary Reflected Gray Code
CDF:	Cumulative Distribution Function
CM:	Coded Modulation
EDS:	Euclidean Distance Spectrum
EFF:	Error Free Feedback
EXIT:	Extrinsic Information Transfer
HARQ:	Hybrid AQR
LLR:	Logarithmic Likelihood Ratio
MAP:	Maximum A Posteriori
MI:	Mutual Information
ML:	Maximum Likelihood
OFDM:	Orthogonal Frequency-Division Multiplexing
PAM:	Pulse Amplitude Modulation

PEP:	Pairwise Error Probability
PDF:	Probability Density Function
PSK:	Phase Shift Keying
QAM:	Quadrature Amplitude Modulation
QPSK:	Quadrature Phase Shift Keying
RSC:	Recursive and Systematic Convolutional code
SNR:	Signal to Noise Ratio
UB:	Union Bound
UWB:	Ultra Wide-Band
WINNER:	Wireless World Initiative New Radio
3GPP:	Third Generation Partnership Project

Part I

Introduction

Introduction

This first part of the thesis is intended to give a short introduction to its main topic: BICM with QAM constellations. This introduction is written aiming to facilitate the understanding of the included contributions in the second part. This chapter is organized as follows. In Section 1 the system model of BICM is introduced. In Section 2 the definition of the *L-values* is presented, and in Section 3 the concepts of mutual information and channel capacity are reviewed. In Section 4 BICM with iterative decoding is introduced, and in Section 5 previous works related to the PDF of the L-values are reviewed. In Section 6 the purpose and the contributions included in this thesis are presented. Finally in Section 7 some open problems in the field are listed, and in Section 8 a list of related contributions not included in this thesis is also presented.

1 Bit-interleaved Coded Modulation

Bit-interleaved coded modulation (BICM) was first introduced by Zehavi in [1], and later analyzed from an information theory point of view in the landmark paper of Caire *et al.* [2]. BICM owes its popularity to the fact that the channel encoder and the modulator separated by a bit-level interleaver may be chosen independently allowing for a simple and flexible design [2, Sec. V]. BICM is considered the dominant technique for coded modulation in fading channels [3], and it only introduces a small penalty when compared to the coded modulation capacity [2, 4]. BICM schemes have been proposed in the IEEE wireless standards such as IEEE 802.11a/g [5] (wireless local area network) and IEEE 802.16 [6] (broadband wireless access). Other examples include the low complexity receivers proposed by the IEEE for the multiband orthogonal frequency-division multiplexing (OFDM) ultra wideband (UWB) transceivers [7], and the wireless world initiative new radio (WINNER) consortium [8]. BICM-OFDM is also considered as a good candidate for power line communication systems [9]. An additional advantage of BICM compared to other schemes such as trellis coded modulation is that due to the flexibility imposed by the bit-level interleaver, the implementa-

tion of adaptive modulation and coding schemes is straightforward [10] (see for example [11]).

In order to increase the spectral efficiency, BICM can be combined with high-order modulation schemes. The most common modulation schemes used in practice are phase shift keying (PSK) and quadrature amplitude modulation (QAM). This thesis focuses on the latter. Borrowing from the idea of iterative processing, the performance of BICM can be further increased by exchanging information between the demapper and the decoder. This scheme called BICM with iterative decoding (BICM-ID) was proposed in [12] and further studied in [13–18].

In Fig. 1 a general discrete baseband BICM transmission model is shown. The vector of coded bits $\mathbf{y}$ generated by the binary channel encoder (ENC) is interleaved (π) generating $\mathbf{y}' = \pi(\mathbf{y})$. These coded and interleaved bits are then gathered in length–n codewords $\mathbf{c}_t$ such that $\mathbf{y}' = [\mathbf{c}_0, \ldots, \mathbf{c}_{N-1}]$, where N is the symbol block length. At any time instant t, the codeword $\mathbf{c}_t$ is mapped to a complex symbol $x_t \in \mathcal{X}$ using a binary memoryless mapping $\mathcal{M} : \{0,1\}^n \rightarrow \mathcal{X}$, where $\mathcal{X}$ is the constellation alphabet. The symbols $\mathbf{x} = [x_0, \ldots, x_{N-1}] = [\mathcal{M}\{\mathbf{c}_0\}, \ldots, \mathcal{M}\{\mathbf{c}_{N-1}\}]$ are sent through the channel whose output is given by $r_t = h_t \cdot x_t + \eta_t$, where h_t is a complex channel gain and η_t is a zero-mean, real, white Gaussian noise sample with variance N_0. Since the mapping is memoryless and both the noise and the complex channel gain samples are independent and identically distributed, from now on we drop the time index t.

The magnitude of the complex channel gain samples follow a Nakagami-m distribution, which allows us to consider a wide range of channels ranging from a Rayleigh fading channel ($m = 1$) to an additive white Gaussian noise (AWGN) channel ($m \rightarrow \infty$). The instantaneous signal to noise ratio (SNR) is given by $\gamma = |h|^2/N_0$. The probability density function (PDF) of the instantaneous SNR is then given by

$$\mathrm{p}_\gamma(\gamma; \overline{\gamma}, m) = \frac{\gamma^{m-1}}{\Gamma(m)} \left(\frac{m}{\overline{\gamma}}\right)^m \exp\left(-\frac{m\gamma}{\overline{\gamma}}\right), \tag{1}$$

where $\overline{\gamma} = \mathbb{E}[\gamma]$ is the average SNR, $\mathbb{E}[\cdot]$ is the expectation operator, and $\Gamma(v+1) = \int_0^\infty \lambda^v \mathrm{e}^{-\lambda} d\lambda$ is the gamma function. We limit our consideration to the idealized but often adopted context of a perfect channel estimator, i.e., the complex channel gain $\mathbf{h} = [h_0, \ldots, h_{N-1}]$ is known at the receiver.

At the receiver's side, the reliability metrics $\mathbf{L}'_{\mathcal{P}}$ are calculated by the demapper $\mathcal{M}^{-1}$ in the form of logarithmic likelihood ratios (LLRs), or simply L-values. The deinterleaver (π^{-1}) generates the metrics $\mathbf{\Lambda}^{\mathrm{a}}_{\mathcal{D}} = \pi^{-1}(\mathbf{L}'_{\mathcal{P}})$ used by the soft-input channel decoder (DEC) to produce an estimation of the transmitted bits. In Fig. 1 a feedback loop is also included to present the so-called BICM-ID scheme. In this scheme, the decoder generates *extrinsic* information $\mathbf{L}^{\mathrm{e}}_{\mathcal{D}}$, which is transformed by the interleaver into *a-priori*

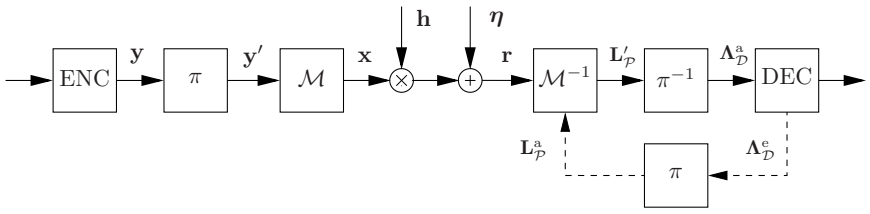


Figure 1: Model of BICM(-ID) transmission.

information for the demapper, $\mathbf{L}^{\mathrm{a}}_{\mathcal{P}} = \pi(\mathbf{\Lambda}^{\mathrm{e}}_{\mathcal{D}})$. The subscripts $\mathcal{P}$ and $\mathcal{D}$ are associated with the demapper and decoder respectively, and the superscripts e and a to extrinsic and a priori respectively. Note that in this case the demapper calculates the L-values $\mathbf{L}'_{\mathcal{P}}$ not only based on the channel information, but also on the a-priori information provided by the decoder.

In this work we analyze the most frequently considered idealized transmission setup, and consequently many aspects of practical receivers (such as for example channel estimation, time/frequency synchronization, or I-Q imbalance) are not taken into account [1, 2, 19]. Moreover, we assume infinitely long sequences and operation with an *ideal* infinite depth interleaver. It is worthwhile to mention that as with all idealized treatments, the results in a practical situation will deviate from theory. However, the theoretical framework provides a basis for a more detailed study, which may require numerical simulations to deal with an otherwise intractable analysis.

Throughout this thesis two different setups for BICM with QAM constellations are analyzed. The first case is when BICM is used together with capacity approaching codes (also called strong codes), i.e., turbo codes [20]. The second one is when the encoder/decoder is implemented using convolutional codes. If no iterations are performed between the demapper and the decoder, the scheme can be used in delay constrained systems, for example in control information blocks as proposed by the WINNER project [8] or in the low complexity receivers proposed by the IEEE for the multiband OFDM-UWB transceivers [7]. If iterations are present at the receiver, either because the decoder is iterative in nature or because the demapper and decoder exchange information in an iterative fashion, the system becomes relevant for applications where a more complex receiver and/or higher latency are affordable.

2 Logarithmic Likelihood Ratios

The logarithmic likelihood ratios (LLRs, or L-values) are used in BICM to represent the available knowledge about the coded bits. The L-values are used when information is exchanged between the demapper and the decoder, or inside the decoder when this is implemented using and iterative algorithm. The sign of an L-value corresponds to a hard decision on the bit, and its magnitude represents the reliability of the hard decision. In this section the L-values calculated by the demapper are presented and some approximations are reviewed.

2.1 Demapper

In this section we discuss L-values' calculations for a bit positions k given a received symbol r. Consequently, we do not use use the vectorial notation presented in the previous section. Moreover, to alleviate the notation, we will omit the subscripts $\mathcal{P}$ and $\mathcal{D}$. Here, the subindex k represents the bit position in the complex symbol. Note that to harmonize the notation used here with the one used in the included papers, an L-value calculated using its exact expression will be denoted by $\tilde{L}_k$, while an L-value calculated using an approximation will be denoted by L_k.

The a posteriori L-values calculated by the demapper $\mathcal{M}^{-1}$ are defined as

$$\tilde{L}'_k = \log\left(\frac{\Pr\{c_k = 1|r\}}{\Pr\{c_k = 0|r\}}\right), \tag{2}$$

where r is the received signal, $\Pr\{\cdot\}$ denotes probability, $k = 0, \ldots, n-1$ is the bit position, and c_k is the k-th bit in the transmitted codeword $\mathbf{c} = [c_0, \ldots, c_{n-1}]$.

Using Bayes' rule it is possible to express (2) as a sum of extrinsic and a priori L-values:

$$\begin{aligned}\tilde{L}'_k &= \log\left(\frac{\Pr\{r|c_k = 1\}}{\Pr\{r|c_k = 0\}}\right) + \log\left(\frac{\Pr\{c_k = 1\}}{\Pr\{c_k = 0\}}\right) \\ &= \tilde{L}^{\mathrm{e}}_k + L^{\mathrm{a}}_k,\end{aligned} \tag{3}$$

where we have used L^{a}_k instead of $\tilde{L}^{\mathrm{a}}_k$ to emphasize the fact that these a priori L-values are obtained from the decoder, and we do not make any assumption on how they were calculated.

Using the definitions in Section 1, the conditional PDF of the received signal is given by

$$\mathrm{p}(r|x, h, \gamma) = \frac{1}{\sqrt{2\pi N_0}} \exp(-\gamma |r/h - x|^2), \tag{4}$$

where the division by h corresponds to the signal scaling due to channel tracking.

Additionally, we know that

$$\Pr\{c_j = b\} = \frac{\exp(b \cdot L_j^{\mathrm{a}})}{1 + \exp(L_j^{\mathrm{a}})}. \tag{5}$$

Assuming independent bits c_k, and using (4) and (5), it can be demonstrated that the extrinsic L-values can be calculated as

$$\begin{aligned}\tilde{L}_k^{\mathrm{e}} &= \log\left\{\frac{\displaystyle\sum_{a\in\mathcal{X}_{k,1}} \exp(-\gamma|r/h-a|^2)\cdot \prod_{j=0,j\neq k}^{n-1} \Pr\{c_j=\beta_j(a)\}}{\displaystyle\sum_{a\in\mathcal{X}_{k,0}} \exp(-\gamma|r/h-a|^2)\cdot \prod_{j=0,j\neq k}^{n-1} \Pr\{c_j=\beta_j(a)\}}\right\} \\ &= \log\left\{\frac{\displaystyle\sum_{a\in\mathcal{X}_{k,1}} \exp\left(-\gamma|r/h-a|^2 + \sum_{j=0,j\neq k}^{n-1} \beta_j(a)\cdot L_j^{\mathrm{a}}\right)}{\displaystyle\sum_{a\in\mathcal{X}_{k,0}} \exp\left(-\gamma|r/h-a|^2 + \sum_{j=0,j\neq k}^{n-1} \beta_j(a)\cdot L_j^{\mathrm{a}}\right)}\right\},\end{aligned} \tag{6}$$

where $\mathcal{X}_{k,b}$ is the set of symbols from $\mathcal{X}$ having the k-th bit equal to b, and $\beta_j(a)$ is the j-th bit of the codeword labeling the symbol a.

Equation (6) tells us how the extrinsic L-values are calculated in BICM-ID for the k-th bit in the symbol, based on both the received symbol r, and the available a priori information L_j^{a} for $j \neq k$.

If there is no exchange of information between the demapper and the decoder (BICM), i.e., there is no a-priori information available ($L_j^{\mathrm{a}} = 0$), (6) is reduced to

$$\tilde{L}_k^{\mathrm{e}} = \log\left\{\frac{\displaystyle\sum_{a\in\mathcal{X}_{k,1}} \exp\big(-\gamma|r/h-a|^2\big)}{\displaystyle\sum_{a\in\mathcal{X}_{k,0}} \exp\big(-\gamma|r/h-a|^2\big)}\right\}, \tag{7}$$

where clearly from (3) it follows that $\tilde{L}'_k = \tilde{L}_k^{\mathrm{e}}$.

Equation (7) tells us how the extrinsic L-values are calculated in BICM for the k-th coded bit based on the received symbol r. Note that this calculation involves the computation of Euclidean distances to all the points in the constellation, applying exponential functions for each of them, and calculating the logarithm of the resulting sum.

2.2 Approximations

Exact metrics' calculation given by (6) and (7) are based on computing the logarithm of the sum of exponential functions. However, it is clear from these equations that the number of exponentials involved increases linearly with the number of constellation symbols. This is the main motivation to seek for simplifications. The most common simplification is the so-called max-log approximation [21], i.e.,

$$\log\left\{\sum_i \exp(\lambda_i)\right\} \approx \max_i\{\lambda_i\}. \tag{8}$$

Using (8), the extrinsic L-values for BICM-ID in (6) become

$$L_k^{\mathrm{e}} = \min_{a\in\mathcal{X}_{k,0}}\left(\gamma|r/h-a|^2 - \sum_{j=0,j\neq k}^{n-1}\beta_j(a)\cdot L_j^{\mathrm{a}}\right) - \min_{a\in\mathcal{X}_{k,1}}\left(\gamma|r/h-a|^2 - \sum_{j=0,j\neq k}^{n-1}\beta_j(a)\cdot L_j^{\mathrm{a}}\right), \tag{9}$$

and for BICM in (7)

$$L_k^{\mathrm{e}} = \min_{a\in\mathcal{X}_{k,0}}\left(\gamma|r/h-a|^2\right) - \min_{a\in\mathcal{X}_{k,1}}\left(\gamma|r/h-a|^2\right), \tag{10}$$

where again $L'_k = L_k^{\mathrm{e}}$.

Although the L-values in (9) and (10) are suboptimal with respect to the ones in (6) and (7), it is known to have small—most often negligible—impact on the receiver's performance when Gray-mapped constellations are used [22–24]. This simplification, proposed already in [1, 2], is frequently adopted for ease of the resulting implementation, e.g., by the 3rd generation partnership project (3GPP) working groups [25]. It is important to mention that the use of the max-log approximation also transforms the nonlinear relationship between r and the L-values into a piece-wise linear function. This fact has been already noted in the literature, cf. for example [26, 27], and it greatly simplifies the analytical treatment.

The max-log approximation given by (8) can also be used in the decoding algorithm. When the maximum a posteriori probability (MAP) algorithm is implemented in the logarithmic domain (Log-MAP) using the Bahl, Cocke, Jelinek and Raviv (BCJR) algorithm, computations involving logarithm of sum of exponentials appear very often. Consequently, the max-log approximation can be used to reduce the decoding complexity originating the so-called Max-Log-MAP algorithm [28, 29].

Different approaches for simplified metrics calculation—most of them based on the Jacobi logarithm [21, 30]—have been investigated in the literature for both the decoding algorithm, and for the metrics calculation in

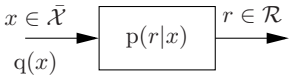


Figure 2: A discrete-time memoryless channel.

the demapper. Discussion about decoding algorithms however, is beyond the scope of this thesis. For more details about this topic, the reader is referred to [23, 31, 32].

3 BICM Capacity

Like any other communication system, BICM has fundamental limits on its transmission rate. Using the standard information theory concepts of mutual information and channel capacity, these bounds can be calculated. In this section these concepts are reviewed, the coded modulation and the BICM capacity are presented, and the idea of a BICM channel is introduced.

3.1 Mutual Information and Channel Capacity

The mutual information (MI) between two events x and r is defined as [33]

$$I(x;r) = \log_2 \left[\frac{\mathrm{p}(r|x)}{\mathrm{p}(r)} \right], \tag{11}$$

and it has a meaning of the amount of information about r given the occurrence of x.

Since we are interested in the MI between inputs and outputs of a communication channel, we define here a discrete-time memoryless channel as a communication channel with inputs $x \in \bar{\mathcal{X}}$ and outputs $r \in \mathcal{R}$, where x and r are outcomes of the random variables $\bar{X}$ and R respectively. Suppose in addition that the input symbols are selected with probability $\mathrm{q}(x)$ and the channel transition probability is given by $\mathrm{p}(r|x)$. A block diagram of this channel is presented in Fig. 2.

Since we are interested in average amounts of information that the output of the channel provides about the input, we define the average MI (AMI) between the random variables $\bar{X}$ and R as

$$I(\bar{X};R) = \mathbb{E}[I(x;r)] \tag{12}$$

$$= \int_{x\in\bar{\mathcal{X}}} \int_{r\in\mathcal{R}} \mathrm{p}(r|x) \cdot \mathrm{q}(x) \cdot \log_2 \left[\frac{\mathrm{p}(r|x)}{\mathrm{p}(r)} \right] dr\, dx, \tag{13}$$

where $\bar{\mathcal{X}}$ and $\mathcal{R}$ are respectively the supports of the random variables $\bar{X}$ and Z.

The channel capacity—introduced in [34] by Claude Shannon—is a very important concept in communication theory because it determines a fundamental limit on the transmission rate of the system. The channel capacity of the channel in Fig. 2 is defined as the maximum AMI, where the maximization is over all possible input distributions $\mathrm{q}(x)$, i.e.,

$$\mathcal{C} = \max_{\mathrm{q}(x)} \big\{ I(\bar{X}; R) \big\}. \tag{14}$$

It is of our interest in this work to consider two important MI calculations. The first one is related to the fact that in practical systems the channel input alphabet has a finite size and all the transmitted symbols are equiprobable. Let $x \in \mathcal{X}$ be the channel input where x is an outcome of the random variable X, and where $\mathcal{X}$ is a finite-size set. Additionally, let the input symbols be selected with probability $\mathrm{q}(x) = 1/|\mathcal{X}|$, where $|\mathcal{X}|$ is the cardinality of $\mathcal{X}$. In this case the AMI between the discrete input X and the continuous output R is given by

$$I(X; R) = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} \int_{r \in \mathcal{R}} \mathrm{p}(r|x) \cdot \log_2 \left[\frac{\mathrm{p}(r|x)}{\frac{1}{|\mathcal{X}|} \sum_{x' \in \mathcal{X}} \mathrm{p}(r|x')} \right] dr, \tag{15}$$

The second important case to consider is when the inputs are binary and equiprobable, which transforms the channel in Fig. 2 into a so-called binary input soft output (BISO) channel. Let the binary set $\mathcal{B} = \{0, 1\}$ be the set containing all the possible channel inputs, where in this case $b \in \mathcal{B}$ is an outcome of the random variable B. Additionally, let the input symbols be selected with probability $\mathrm{q}(b) = 1/2$. The AMI between the binary input B and the continuous output R is then given by

$$I(B; R) = \frac{1}{2} \sum_{b \in \mathcal{B}} \int_{r \in \mathcal{R}} \mathrm{p}(r|b) \cdot \log_2 \left[\frac{2 \cdot \mathrm{p}(r|b)}{\mathrm{p}(r|0) + \mathrm{p}(r|1)} \right] dr. \tag{16}$$

In the following sections, and using the concepts introduced above, the so-called coded modulation (CM) capacity and the BICM capacity are presented. We feel that it is important to mention that the capacity expressions presented in the next sections (sometimes referred in the literature as "channel capacity") are not channel capacity expressions in the strict sense of the definition (14). However, with a slight abuse of notation and in order to follow the common terminology used in the literature, we will refer to them as capacity expressions and we will denote them using $\mathcal{C}$.

3.2 Coded Modulation Capacity

Let $x \in \mathcal{X}$ be the outcome of a random variable X, where $\mathcal{X}$ is a finite set of points in the complex plane, i.e., the constellation symbols in Fig. 1. Let also

$r \in \mathbb{C}$ be the outcome of a complex random variable R which represents the received signal in Fig. 1. The so-called constrained capacity—also known as CM capacity—was first introduced in [4] and then presented in [2] as

$$\mathcal{C}^{\mathrm{CM}} = I(X;R) = n - \mathbb{E}\left[\log_2 \frac{\sum_{a\in\mathcal{X}} \mathrm{p}(r|a)}{\mathrm{p}(r|x)}\right], \tag{17}$$

where $n = \log_2 |\mathcal{X}|$ is the number of bits per transmitted symbol and $\mathcal{X}$ is the constellation alphabet. It is straightforward to see that (17) is simply a re-written version of (15).

Equation (17) is a fundamental bound for the transmission rate that a system using symbol-based decisions can achieve. This bound is applicable for example to trellis coded modulation systems, where the code and the modulator are jointly designed. This scheme can be obtained removing the bit-level interleaver in Fig. 1 and adding a symbol-level interleaver after the mapper $\mathcal{M}$.

3.3 BICM Capacity

The BICM capacity was first presented in [2, Sec. III], where it was shown that due to the ideal interleaving, the BICM system can be regarded as n parallel memoryless and independent BISO channels, each of them associated with a bit position in the binary mapping $\mathcal{M}$.

Let $b \in \mathcal{B}$ be the outcome of a random variable B where the binary set $\mathcal{B} = \{0, 1\}$ is the same as in Section 3.1. Let $I(B;R)$ denote the AMI between the binary input B and the channel output R, where R is defined as in Section 3.2. The BICM capacity is then defined as the average over the AMIs of each of the n equivalent channels, i.e.,

$$\begin{aligned}\mathcal{C}^{\mathrm{BICM}} &= n \cdot I(B;R) \\ &= n - \sum_{k=0}^{n-1} \mathbb{E}\left[\log_2 \frac{\sum_{a\in\mathcal{X}} \mathrm{p}(r|a)}{\sum_{a\in\mathcal{X}_{k,b}} \mathrm{p}(r|a)}\right],\end{aligned} \tag{18}$$

where $\mathcal{X}_{k,b}$ is the subset of symbols in $\mathcal{X}$ labeled with $b \in \mathcal{B}$ in the bit position k.

It was demonstrated in [2] using the data-processing theorem [35] that the BICM capacity is always less or equal than the CM capacity

$$\mathcal{C}^{\mathrm{BICM}} \leq \mathcal{C}^{\mathrm{CM}}. \tag{19}$$

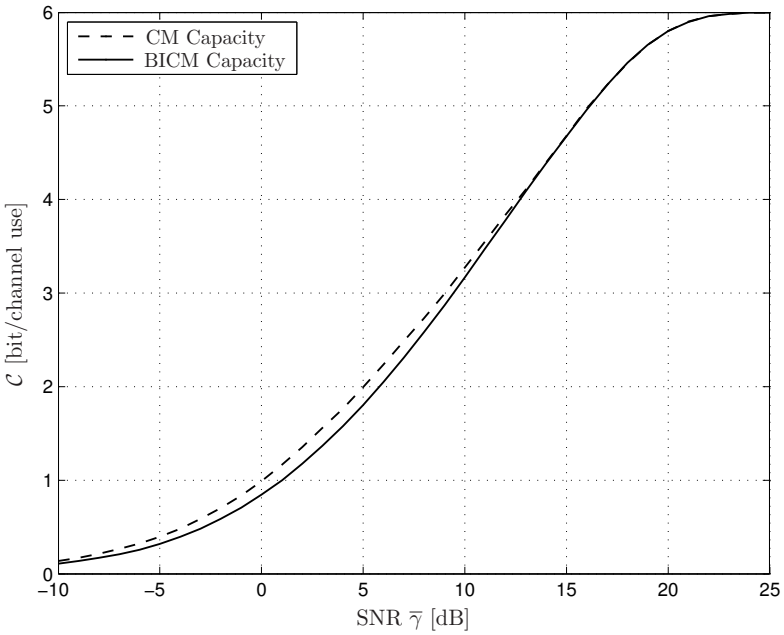


Figure 3: CM and BICM capacity given by (17) and (18) for Gray-mapped 64-QAM in an AWGN channel.

This can be intuitively understood as follows. In BICM the metrics of the coded bits within a received symbol are calculated disregarding the information available on the other $n-1$ bits [36]. The difference between $\mathcal{C}^{\mathrm{BICM}}$ and $\mathcal{C}^{\mathrm{CM}}$ is in general very small, and consequently, the low complexity and high flexibility of BICM make it very attractive from an implementation point of view.

Both channel capacities (17) and (18) can be evaluated using numerical integration. It is also worth to mention that (17) does not depend on the mapping between the codewords and the constellation symbols. On the other hand, for BICM, the bit-to-symbol mapping does affect the capacity. In Fig. 3 both CM and BICM channel capacities are presented for 64-QAM with Gray mapping in an AWGN channel.

3.4 The BICM Channel

In BICM the channel decoder is provided with deinterleaved L-values. Assuming an ideal interleaver, the ensemble of elements between the output of the encoder and the input to the decoder, i.e., the interleaver, the modula-

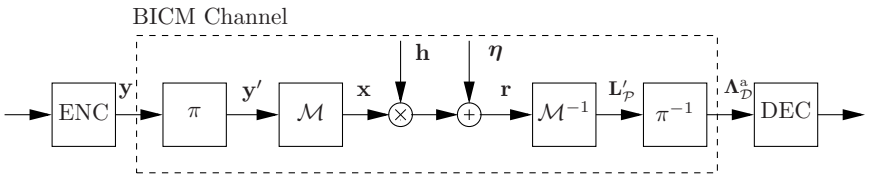


Figure 4: The BICM channel.

tor, the propagation channel, the metrics' calculator, and the de-interleaver, may be seen as a memoryless "BICM channel" with binary inputs $\mathbf{y}$ (coded bits) and real outputs $\mathbf{\Lambda}_{\mathcal{D}}^{\mathrm{a}}$ (deinterleaved L-values) [1, 19], i.e., a BISO channel. This concept is presented in Fig. 4.

Let L_k be a random variable which represents an L-value calculated by the demapper for the k-th bit position given the transmitted symbol a, and let $\mathrm{p}_{L_k}(\lambda|a)$ denote its conditional PDF. Let also L denote an L-value at the output of the BICM channel (cf. Fig. 4), and with a slight abuse of notation, let $\mathrm{p}_L(\lambda|b)$ denote its conditional PDF given a binary input of the BICM channel $b \in \mathcal{B}$. Clearly the PDF of the random variable L is given by

$$\mathrm{p}_L(\lambda|b) = \frac{2}{n \cdot |\mathcal{X}|} \sum_{k=0}^{n-1} \sum_{a \in \mathcal{X}_{k,b}} \mathrm{p}_{L_k}(\lambda|a). \tag{20}$$

In general, when high-order constellations are used, it is well known that the binary mapping makes the BICM channel not symmetric, i.e., $\mathrm{p}_L(\lambda|1) \neq \mathrm{p}_L(-\lambda|0)$ for some $\lambda \in \mathbb{R}$. To clarify this, take for example the case of a Gray-mapped (see Section 4.2) 4-level pulse amplitude modulation (4-PAM) constellation as shown in Fig. 5. If we take a look at the mapping at each bit position, we can see that in average the bits transmitted for $k = 0$ have a higher protection level than for $k = 1$. This can be intuitively understood as follows. For $k = 1$ and independently of the transmitted symbol x, there is always a symbol with the opposite value of the bit at distance 2Δ. However, for $k = 0$, only the two symbols in the center of the constellation ($x = \pm\Delta$) fulfill this property. For the other two symbols in the extremes ($x = \pm 3\Delta$), the closest symbol labeled with the opposite value of the bit is at distance 4Δ yielding a higher protection level.

Following the suggestion of [2], this problem can be overcome using a scrambler that randomly negates the transmitted bit and inverting the sign of the metrics at the receiver. If this symmetrization is used, it is possible

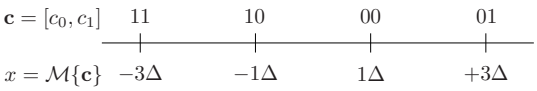


Figure 5: Gray-mapped 4-PAM.

to define an averaged and symmetrized PDF $\mathrm{p}_{L^{\mathrm{s}}}(\lambda|b)$ as

$$\mathrm{p}_{L^{\mathrm{s}}}(\lambda|b) = \frac{1}{n \cdot |\mathcal{X}|} \sum_{k=0}^{n-1} \left[\sum_{a \in \mathcal{X}_{k,b}} \mathrm{p}_{L_k}(\lambda|a) + \sum_{a \in \mathcal{X}_{k,1-b}} \mathrm{p}_{L_k}(-\lambda|a) \right], \qquad (21)$$

where the superscript in L^{s} represents the imposed symmetry condition.

Using the simplified channel model presented in Fig. 4, and the symmetrization procedure described above, the capacity of the system can be calculated directly using (16) as

$$\begin{aligned} \mathcal{C}^{\mathrm{BICM}} &= n \cdot I(B; L^{\mathrm{s}}) \\ &= \frac{n}{2} \sum_{b \in \mathcal{B}} \int_{-\infty}^{\infty} \mathrm{p}_{L^{\mathrm{s}}}(\lambda|b) \cdot \log_2 \left[\frac{2 \cdot \mathrm{p}_{L^{\mathrm{s}}}(\lambda|b)}{\mathrm{p}_{L^{\mathrm{s}}}(\lambda|0) + \mathrm{p}_{L^{\mathrm{s}}}(\lambda|1)} \right] d\lambda, \end{aligned} \qquad (22)$$

where the coefficient n takes into account the fact that there are n bit positions.

Using (22), the BICM capacity can be efficiently calculated via one-dimensional integration. However, the problem that arises here is that the PDF of the metrics must be known analytically. A first approach to solve this problem is to estimate the PDF of the L-values using histograms, however, this does not allow analytical treatment and is usually a numerically complex task, especially when large constellation sized are used. Other options to estimate the PDF of the L-values are the so-called cumulant method [37] or the Gaussian mixture models [38, 39]. As an example, in Fig. 6 the normalized BICM capacity ($I(B; L^{\mathrm{s}})$), calculated using (22), and based on histograms for 16 and 256-QAM and different channel models is presented. The L-values are calculated using exact metrics calculation (7) and using the max-log approximation (10). From this figure it is clear that in terms of capacity the degradation produced by the max-log approximation is very small and only perceptible for low SNR values. Similar degradation must be expected when capacity approaching codes, i.e., turbo or low-density parity-check codes are used [23].

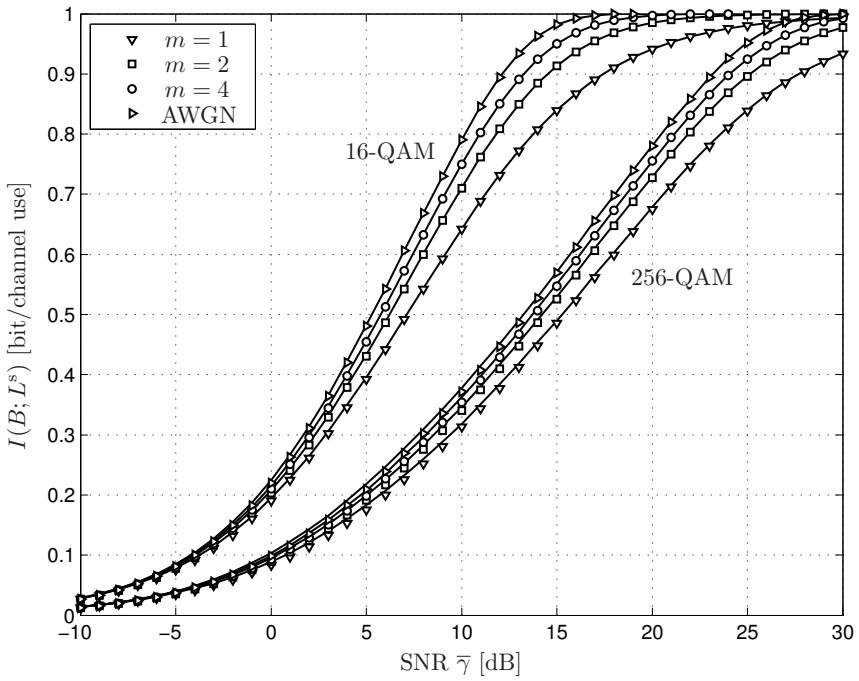


Figure 6: Normalized BICM capacity ($\mathcal{C}^{\mathrm{BICM}}/n$) for 16 and 64-QAM and different channel models based on histograms. Solid lines denote exact metric calculation (7) and markers denote max-log metrics (10).

4 BICM with iterative decoding

Borrowing from the idea of iterative processing, BICM with iterative decoding (BICM-ID) can improve the performance of BICM. In this scheme, the demapper and the channel decoder exchange extrinsic information in an iterative fashion. In this section the concept of an iterative process and the so-called EXIT analysis are introduced, and a discussion about mappings for BICM-ID is also presented.

4.1 EXIT Analysis and Iterative Processing

An iterative process (also referred as a turbo process) is the process when different elements in the receiver exchange information in an iterative fashion in order to improve some parameter estimation. Although when the concept of iterative decoding dates back to 1954 with the work of Elias [29],

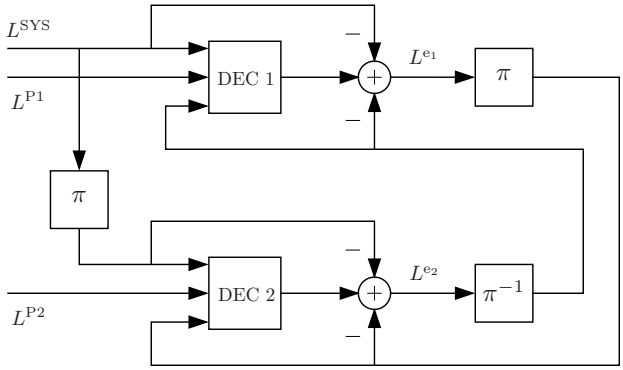


Figure 7: Iterative (turbo) decoder for a parallel concatenated turbo code.

a major breakthrough came when Berrou, Glavieux, and Thitimajshima introduced the so-called turbo codes in [20], which approached the theoretical limits as no other code had done in the past. The original turbo codes were formed by a parallel concatenation of two standard recursive convolutional encoders separated by an interleaver. At the receiver, both decoders exchange information in an iterative fashion continuously improving the decisions about the transmitted bits. After a certain number of iterations the process is stopped since more iterations do not yield a better estimate. An example of this process is shown in Fig. 7, where both decoders (DEC 1 and DEC 2) iteratively exchange extrinsic information. In this figure the superscripts SYS refers to the systematic bits, and P1 and P2 to the parity bits from the first and the second encoder respectively.

After turbo codes were introduced in 1993, the concept of iterative processing has been extended to different parts of the receiver. Nowadays it is common to design iterative receivers with algorithms for turbo synchronization, turbo equalization, turbo channel estimation, turbo multi-user detection, etc.

The resulting iterative process is highly non-linear, so to analyze it, approximations are often applied. The most popular one, known as the density evolution, treats the L-values as realizations of independent variables, whose PDF, evolving throughout the iterations, provides the information necessary to describe the iterative process. Such tracking of the PDFs is non-trivial because, in general, the closed-form expression for the PDF of the L-values are not known and are difficult to obtain [19]. An exception is the simple case of the demapper in binary phase shift keying (BPSK) modulation for which the PDF of the output L-values is known [40]. BPSK, however, does

not lend itself to iterative decoding. A useful simplification was proposed in [40], where only one parameter of the PDF—the AMI between the coded bits and the corresponding extrinsic L-values—is considered. This AMI can be directly calculated using (16), and it has been proved in [41] to produce an accurate description of the iterative process. This approximation produces the so-called extrinsic information transfer (EXIT) charts, which, using a two-dimensional space, provides a good insight into the density evolution.

For a turbo decoding process like the one presented in Fig. 7, the EXIT functions of both decoders (DEC 1 and DEC 2) can be drawn, and consequently the performance of the system can be predicted. Analogously, for a BICM-ID receiver, the EXIT functions of the demapper and the decoder must be considered.

An important property of turbo codes, BICM-ID, and an iterative processes in general, is that the performance in terms of bit-error rate (BER) can be divided into two distinct regions: an *early convergence* region (also called waterfall region) and an *error floor* region. These two regions are illustrated in Fig. 8, which presents the BER obtained using a turbo code[1] in an AWGN channel. The early convergence region is the SNR range where a small increase in the SNR value will produce a substantial decrease in the BER (steep BER curve). The error floor region instead, is the SNR range where no substantial improvements should be expected by increasing the SNR (flat BER curve).

4.2 Mappings for BICM-ID

Soon after BICM-ID was introduced, the key role of the binary mapping was recognized. Abundant literature exists discussing the design of bit mappings for improving the system performance in the waterfall or the error-floor region, cf. for example [15, 16, 40, 42–45].

The mapping block $\mathcal{M}$ in Fig. 1 maps bijectively the codewords $\mathbf{c}$ to the constellation points $\mathcal{M} : \{0,1\}^n \to \mathcal{X}$. Among all the $|\mathcal{X}|!$ possible mappings (including equivalent mappings), one family of mappings is of special interest due to its theoretical and practical implications. A Gray mapping is defined as a binary mapping $\mathcal{M}$ such that the closest neighbor to any constellation symbol is always at Hamming distance one (cf. for example Fig. 5). Within this family of mappings, the so-called binary reflected Gray code (BRGC) is of special importance since it has been proved in [46] to be the optimum mapping in the sense of minimizing the uncoded asymptotic

[1] The numerical results are obtained using a turbo code formed by the parallel concatenation of two recursive systematic convolutional (RSC) codes with polynomial generators $(1, 5/7)_8$. Alternate puncturing of the parity bits yields an overall rate of 1/2. The information block length is 512 bits, the channel is an AWGN channel, and the constellation is 16-QAM with Gray mapping. The decoder is implemented using the Log-MAP algorithm with 10 iterations.

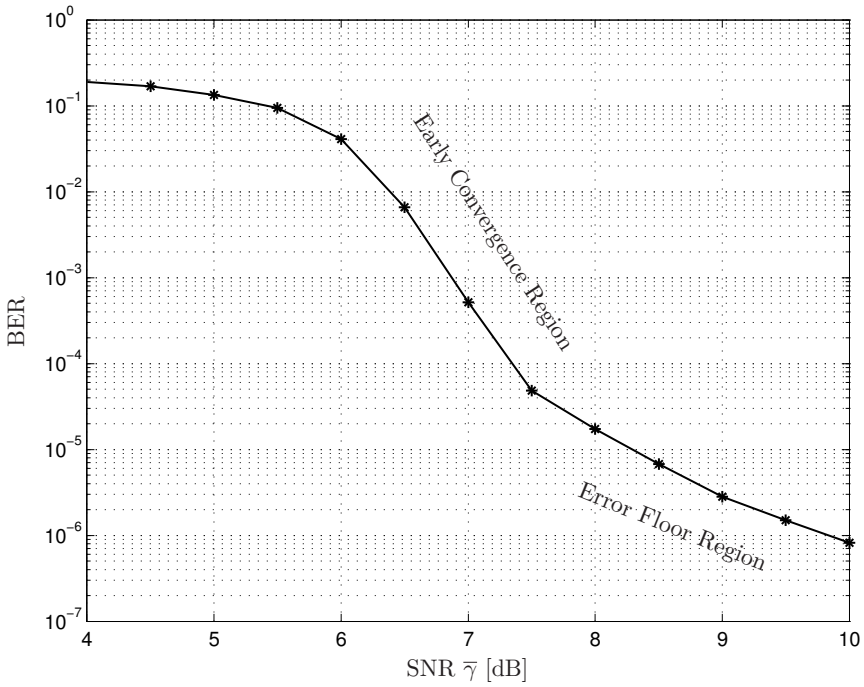


Figure 8: BER for a rate-1/2 turbo code in an AWGN channel using 16-QAM.

BER for M-PSK, M-PAM and M^2-QAM. Moreover, in [47] it has been proved that the BRGC is in fact optimum not only asymptotically, but also for a significant range of SNR. More details about how to construct such BRGC mappings can be found in [46].

It is also worth mentioning that for BICM it was conjectured in [2] that a Gray mapping maximizes the BICM capacity given by (18). It should be mentioned that in [2] it was not specified which specific Gray mapping is the one which maximizes the BICM capacity. This conjecture has recently been disproved in general in [48], however, it seems that the optimality of Gray mappings, and particularly of the BRGC, holds for most of the relevant cases. Additionally, it is worth mentioning that nowadays it is well understood that Gray mappings offer small improvements through the iterations in BICM-ID.

Another interesting concept worth to mention here is the so-called irregular modulation (also called modulation doping), where different signal constellations and/or mappings are applied within the block of transmitted

symbols. This method gives the designer an extra degree of freedom to optimize the iterative process based on some criterion. For a more detailed discussion about this topic, the reader is referred to [17, 49].

In BICM-ID an extreme situation occurs when the iterative process converges and, after certain number of iterations, a perfect knowledge on all the other bit positions can be assumed. This case is usually called error free feedback (EFF), which is important to analyze the BER in the error floor region, and where the detection process of high-order modulation schemes is transformed into a detection of BPSK signals. Here it is worth mentioning that the so-called Euclidean distance spectrum (EDS) (also called Euclidean distance profile) is a very useful concept to analyze the performance of BICM-ID in the error floor region. The EDS for any constellation with EFF is defined as the number of pairs of constellation symbols at a certain Euclidean distance with Hamming distance one. We denote the sorted list of possible Euclidean distances as $\mathcal{L} = (\lambda_1, \ldots, \lambda_P)$, where λ_p is the p-th Euclidean distance and $\lambda_p < \lambda_{p+1}$. We also define here the free Euclidean distance in the EDS, λ_{free}, as the first nonzero element in $\mathcal{L}$.

As an example, in Fig. 9 we present the optimized mapping M16a of [16] and also a mapping denoted by M16c which has the property of having a concave EXIT function[2]. The EDS with EFF for both mappings is given in Table 1 assuming a unitary minimum distance between the constellation points. Note that this analysis can also be done in a per-position basis, which would be useful to analyze the error floor region of BICM-ID when parallel interleavers are used for each bit position (instead of one interleaver) as in the original work of Li *et al.* [12].

We emphasize that the EDS can be used to predict the performance of BICM-ID in the error floor region. If the objective is to minimize the BER in this region, a good mapping would be a mapping where the free Euclidean distance is maximized, and where the multiplicity of λ_{free} is minimized. Based on Table 1 we can see that the mappings M16a and M16c have free Euclidean distances $\lambda_{\text{free}}^{\text{M16a}} = \sqrt{5}$ and $\lambda_{\text{free}}^{\text{M16c}} = 1$ with multiplicity 16 and 4 respectively. Consequently, BICM-ID with M16a will give a lower BER in the error floor region than BICM-ID with M16c. More details about the EDS for BICM can be found in [10].

In Fig. 10 and Fig. 11 a typical EXIT chart for two BICM-ID schemes using 16-QAM are shown. Fig. 10 shows a system formed by an RSC code with polynomial generators $(1, 5/7)_8$ and the optimized mapping M16a [16] (cf. Fig. 9), and Fig. 11 shows a system formed by an RSC code with polynomial generators $(1, 4/7)_8$ and the mapping M16c (cf. Fig. 9). The horizontal and vertical lines represent the decoding process (each segment corresponds to half an iteration). In both figures the EXIT function of

[2]This mapping has not been published and it provides a good tradeoff between early convergence and error floor performance.

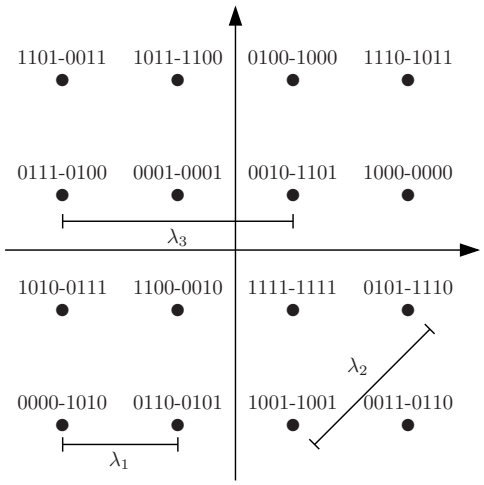


Figure 9: Optimized mapping M16a of [16] (left side of '-') and M16c (right side of '-'). The first three elements in the EDS $\mathcal{L}$ are shown.

$\mathcal{L}$	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9
λ_p	1	$\sqrt{2}$	2	$\sqrt{5}$	$2\sqrt{2}$	3	$\sqrt{10}$	$\sqrt{13}$	$3\sqrt{2}$
M16a	0	0	0	16	4	0	4	8	0
M16c	4	6	5	7	1	3	3	2	1

Table 1: EDS with EFF of 16-QAM for the optimized mapping M16a of [16] and for M16c.

the BRGC is also shown. In the axis of these figures, and with a slight abuse of notation, we denote the AMI between the coded bits and the extrinsic L-values generated by the decoder $I^{\mathrm{e}}_{\mathcal{D}}$, and the AMI generated by the demapper by $I^{\mathrm{e}}_{\mathcal{P}}$. Analogously, the a priori AMI (superscript a) used by the decoder and the demapper are denoted by $I^{\mathrm{a}}_{\mathcal{D}}$ and $I^{\mathrm{a}}_{\mathcal{P}}$ respectively.

In Fig. 10 we can appreciate that the tunnel between both curves just opened up, which means that the so-called turbo cliff is reached at this SNR value ($\overline{\gamma} = 7.76$ dB). The iterative process reaches (after many iterations) the upper right corner of the curves, which means that the BER will be quite low. In this case, a small increase in the SNR will produce a substantial decrease in the BER, however, if the SNR decreases, the tunnel will be closed, and the BER might be quite high (the iterative process does not

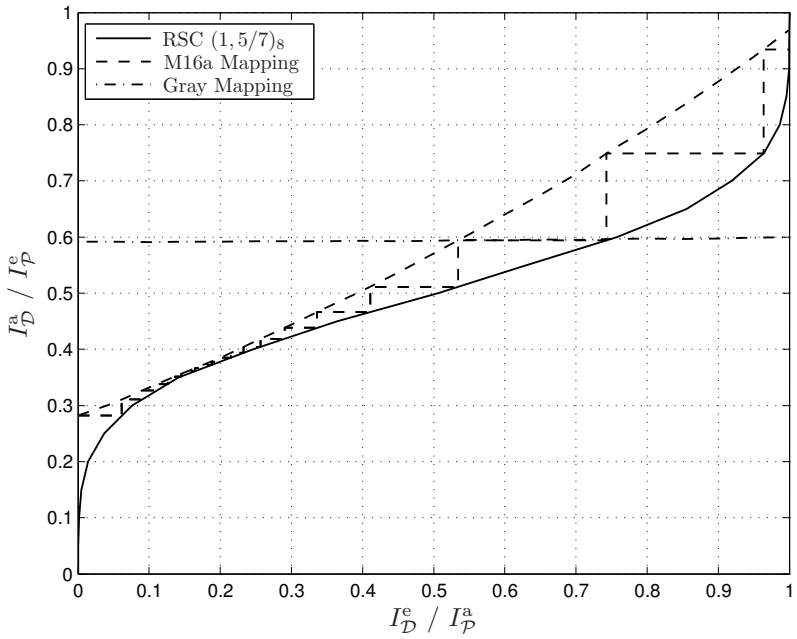


Figure 10: EXIT functions and decoding trajectories for BICM-ID formed by the concatenation of the $(1, 5/7)_8$ code and the M16a mapping in an AWGN channel for 16QAM and SNR $\overline{\gamma} = 7.76$ dB.

converge). If we analyze Fig. 11, we can see that the tunnel is widely open, but the iterative process will stop after approximately 7-8 iterations (when both curves cross each other). This will produce a higher BER, however, the second scheme will converge for a lower SNR, i.e., a small decrease on the SNR will not close the tunnel. This is the so-called early convergence property which can also be appreciated in the design of turbo codes [40].

The importance of the EXIT analysis presented previously for BICM-ID is that it allows the designer to select the mappings *and* the code according to some optimization criteria, for example low error floor or early convergence, without many full BER simulations. A major drawback of this analysis is that the curves are obtained by numerical simulations since the PDFs of the L-values (which allows us to calculate the AMI) are not known.

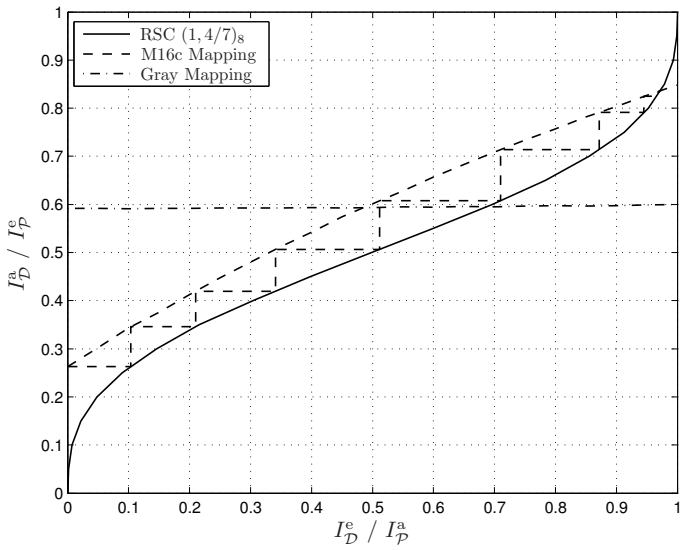


Figure 11: EXIT functions and decoding trajectories for BICM-ID formed by the concatenation of the $(1, 4/7)_8$ code and the M16c mapping in an AWGN channel for 16QAM and SNR $\overline{\gamma} = 7.76$ dB.

5 PDF of L-values

As detailed in Section 3, the BICM system can be regarded as a BISO channel which is completely defined by the PDF of the outputs (L-values). Consequently, knowing the PDF of the L-values allows us to calculate the BICM capacity and also to predict the coded performance in terms of BER. In this section, previous works addressing the issue of modeling the PDF of the L-values are presented, and a simple performance evaluation method is reviewed.

5.1 Previous Work

Analytical expressions for the PDF of the L-values are known for BPSK and quadrature phase shift keying (QPSK) modulations in fading and non-fading channels [40] for BICM. The difficulty in finding the PDF for high-order modulations, recognized in [19], can be alleviated when the so-called max-log approximation is used in the receiver to calculate the L-values. Thanks to this approach, analytical expressions for the PDF were developed in [50] for 16 and 64-QAM, and in [51] for 8 and 32-PSK. A general methodology to

find closed-form expressions for the PDF of the L-values, which is applicable to any constellation and mapping, was also presented in [52]. However, this general methodology is algorithmic in nature, i.e., requires nontrivial programming.

5.2 Performance Evaluation

Evaluation of performance of BICM for high-order modulations was usually limited by the lack of formal description of the metrics used in the decision process. This problem was partially palliated by bounding techniques, e.g., [1] or the so-called expurgated bounds in [2]. However, these bounding techniques are either very loose, or when tight, they must be algorithmic.

To analyze the performance of any linear code under maximum likelihood (ML) decoding, union bound techniques are often used. The union bound (UB) on the BER for a convolutional code—which represents an upper bound on the BER of the code P_{b}—is given by [33, Sec. 4.4]

$$P_{\mathrm{b}} \leq \mathrm{UB} = \sum_{d=d_{\mathrm{free}}}^{\infty} \beta_d \cdot \mathrm{PEP}(d), \tag{23}$$

where d_{free} is the free distance of the code, β_d is the weight distribution spectrum of the code, and $\mathrm{PEP}(d)$ is the pairwise error probability, which represents the probability of detecting a codeword with Hamming weight d instead of the transmitted all-zero codeword.

The weight distribution spectrum of the code β_d is the number of bit errors (the information error weight) for error events of distance d. Although this concept was first used for convolutional codes, it is also possible to calculate the weight distribution spectrum of a turbo code using the *uniform interleaver* concept introduced in [53, 54]. Different algorithms for calculating spectrums can be used, for example the recursive algorithm presented in [55].

For a symmetric and memoryless channel, the pairwise error probability in (23) can be written as the tail probability of a sum of d L-values in the nonzero path [19, 56], i.e.,

$$\begin{aligned} \mathrm{PEP}(d) &= \Pr\left\{ \sum_{j=1}^{d} L(j) > 0 \right\} \\ &= \int_0^{\infty} \mathrm{p}_d^{\Sigma}(\lambda)\, d\lambda, \end{aligned} \tag{24}$$

where $\mathrm{p}_d^{\Sigma}(\lambda)$ represents the PDF of the sum of d L-values. Since the L-values are assumed to be independent (due to the ideal interleaver), the PDF of their sum is a convolution of the individual PDFs. The pairwise

error probability is then calculated as an integral of this convolution over the positive values of the argument λ.

Pairwise error probability calculations are straightforward for BPSK or QPSK, however, for high-order QAM modulations, they become nontrivial. Important contributions addressing this issue have been published recently by A. Martinez, A. Guillén i Fàbregas and G. Caire who presented in [57] a method to approximate the BISO BICM channel by a BPSK channel with scaled SNR. Another approach presented by the same authors to tackle this problem is the so-called saddlepoint approximation [19], which has also been used in [56] for computing bounds on the BER for BPSK over fading channels. The main drawback of these approaches is that both rely on numerical integration.

6 Purpose and Contributions

The purpose of this thesis is to develop methods to characterize the performance of BICM when high-order modulation schemes are used. The main contribution is the development of analytical expressions for the PDF of L-values for BICM and BICM-ID. Based on these results the BICM capacity is efficiently calculated and the coded performance is predicted using union bound techniques. Moreover, BICM with a packet retransmission scheme is analyzed, and the EXIT functions of the demapper for BICM-ID are also calculated.

The use of histograms (obtained through Monte-Carlo simulation) or other estimation methods (for example the cumulant method or the Gaussian mixture models) as estimates of the PDFs do not allow for an analytical treatment and are computationally costly alternatives when compared to the use of closed-form expressions. This provides the motivation to derive analytical expressions for the PDF.

The three contributions included in this thesis are summarized in the following list.

Paper A – On the Distribution of Extrinsic L-values in Gray-mapped 16-QAM

In this paper the issue of the probabilistic modeling of the extrinsic L-values for BICM-ID is addressed. Starting with a simple piece-wise linear model of the L-values obtained via the max-log approximation, expressions for the cumulative distribution functions (CDFs) of the L-values are found. The desired forms of the PDFs for Gray-mapped 16-QAM are found differentiating the CDF. The developed analytical expressions are then used to efficiently compute the so-called EXIT functions of the demapper for different values

of SNR. The proposed analytical expressions are also compared with the histograms of the L-values obtained through numerical simulations.

Paper B – Distribution of L-values in Gray-mapped M^2-QAM: Closed-form Approximations and Applications

In this paper closed-form expressions for the PDFs of the L-values in BICM for QAM constellations with Gray mapping are developed. Based on these expressions, two simple Gaussian mixture approximations that are analytically tractable are also proposed. The developments are used to efficiently calculate the BICM capacity, and to develop bounds on the coded bit-error rate when a convolutional code is used. The coded performance of an hybrid automatic repeat request (HARQ) based on constellation rearrangement is also evaluated.

Paper C – Distribution of Max-Log Metrics for QAM-based BICM in Fading Channels

In this paper closed-form expressions for the PDFs of the L-values in BICM transmissions over fully-interleaved fading channels are derived. The expressions are valid for the relevant case of QAM schemes with Gray mapping when the metrics are calculated via the so-called max-log approximation. The results presented are particularized for a Rayleigh fading channel, however, developments for the general case of a Nakagami-m case are also included. Using the developed expressions, the performance of BICM transmissions using convolutional and turbo codes is efficiently evaluated, i.e., without resorting to otherwise required two-dimensional numerical integration. The BICM capacity for different fading channels and constellation sizes is also evaluated.

7 Future Work

Some open issues related to the work presented in this thesis are described in the following list:

- It is still not clear which mappings are the ones which maximize the BICM capacity. As discussed previously, a Gray mapping was conjectured to be the optimal, however, this has been been recently disproved. Consequently, an open problem is to find the optimum mapping for any constellation size. Based on the results available in the literature, which are available only for small constellation sizes where it is still possible to do a full search, it seems that there is no unique optimum mapping. However, the BRGC is a good candidate since

numerical results have shown that this mapping is the optimum for most of the analyzed cases and SNR values of interest.

- Although good mappings for BICM-ID are well-known in the literature, there are no systematic methods to construct them. All the good mappings are found based on some algorithmic search which has obvious limitations when the constellation size increases. It is known that mappings for BICM-ID can be characterized by its EDS which is analogous to the distance spectrum of a convolutional code. In this sense, it would be possible to calculate the minimum free distance of the mapping for a given constellation size, and eventually come up with a method to systematically construct good mappings based on optimum EDS.

- It is well-known that the optimum BICM-ID design can be achieved if the code and the mapping are jointly designed. In the literature however, this approach is considered too difficult, and in general the design of mappings is tackled for a given code and vice versa. Moreover, the interleaver design in this case is also crucial, but in most of the cases is completely ignored.

- Since the binary mapping of high-order modulation schemes produces unequal error protection, the code and the interleaver can be designed to take into account this property. With a proper selection of code and interleaver design, a performance improvement should be expected when compared to the standard codes and the (pseudo)random interleaver used up to now.

8 Related Contributions

Other related publications by the author, which are not in included in this thesis, are:

- "On Adaptive BICM with Finite Block-Length and Simplified Metrics Calculation", A. Alvarado, H. Carrasco, and R. Feick, *IEEE Vehicular Technology Conference 2006, VTC-2006 Fall*, Montreal, Canada, Sep. 2006.

- "Distribution of L-values in Gray-mapped M^2-QAM Signals: Exact Expressions and Simple Approximations", A. Alvarado, L. Szczecinski, R. Feick, and L. Ahumada, *IEEE Global Telecommunications Conference GLOBECOM 2007*, Washington, USA, Nov. 2007.

- "Probability Density Functions of Reliability Metrics for 16-QAM-Based BICM Transmission in Rayleigh Channel", L. Szczecinski, A. Alvarado, and R. Feick, *IEEE International Conference on Communications, ICC 2007*, Glasgow, UK, June 2007.

- "Closed-form approximation of Coded BER in QAM-based BICM Faded Transmission", L. Szczecinski, A. Alvarado, and R. Feick, *IEEE Sarnoff Symposium 2008*, Princeton, NJ, USA, Apr. 2008.

References

[1] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 3, pp. 873–884, May 1992.

[2] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[3] A. Goldsmith, *Wireless Communications*. New York, NY: Cambridge University Press, 2005.

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

[5] IEEE 802.11, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High-speed physical layer in the 5GHz band," IEEE Std 802.11a-1999(R2003), Tech. Rep., Jul. 1999.

[6] I. Koffman and V. Roman, "Broadband wireless access solutions based on OFDM access in IEEE 802.16," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 96–103, Apr 2002.

[7] A. Batra *et al.*, "Multi-band OFDM physical layer proposal for IEEE 802.15 task group 3a," Mar. 2004, Document IEEE P802.15-03/268r3, available at http://grouper.ieee.org/groups/802/15/.

[8] T. Lestable *et al.*, "D2.2.3 modulation and coding schemes for the WINNER II system," WINNER II, Tech. Rep. IST-4-027756, November 2007, available at https://www.ist-winner.org.

[9] E. Biglieri, "Coding and modulation for a horrible channel," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 92–98, May 2003.

[10] F. Schreckenbach, "Iterative decoding of bit-interleaved coded modulation," Ph.D. dissertation, Munich University of Technology, Munich, Germany, 2007.

[11] P. Örmeci, X. Liu, D. L. Goeckel, and R. D. Wesel, "Adaptive bit-interleaved coded modulation," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1572–1581, Sep. 2001.

[12] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 1, no. 6, pp. 169–171, Nov. 1997.

[13] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping for QPSK modulation," *IEE Electronics Letters*, vol. 34, no. 15, pp. 1459–1460, Jul. 1998.

[14] A. Chindapol and J. A. Ritcey, "Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 944–957, May 2001.

[15] M. Tüchler, "Design of serially concatenated systems depending on the block length," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 209–218, Feb. 2004.

[16] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 7, no. 12, pp. 593–595, Dec. 2003.

[17] L. Szczecinski, H. Chafnaji, and C. Hermosilla, "Modulation doping for iterative demapping of bit-interleaved coded modulation," *IEEE Commun. Lett.*, vol. 9, no. 12, pp. 1031–1033, Dec. 2005.

[18] X. Li, A. Chindapol, and J. A. Ritcey, "Bit-interlaved coded modulation with iterative decoding and 8PSK signaling," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 1250–1257, Aug. 2002.

[19] A. Martinez, A. Guillén i Fàbregas, and G. Caire, "Error probability analysis of bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 262–271, Jan. 2006.

[20] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *International Conference on Communications, ICC 1993*, Geneva, Switzerland, May 1993, pp. 1064–1070.

[21] A. J. Viterbi, "An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 260–264, Feb. 1998.

[22] M. K. Simon and R. Annavajjala, "On the optimality of bit detection of certain digital modulations," *IEEE Trans. Commun.*, vol. 53, no. 2, pp. 299–307, Feb. 2005.

[23] B. Classon, K. Blankenship, and V. Desai, "Channel coding for 4G systems with adaptive modulation and coding," *IEEE Wireless Commun. Mag.*, vol. 9, no. 2, pp. 8–13, Apr. 2002.

[24] A. Alvarado, H. Carrasco, and R. Feick, "On adaptive BICM with finite block-length and simplified metrics calculation," in *IEEE Vehicular Technology Conference 2006, VTC-2006 Fall*, Montreal, Canada, Sep. 2006.

[25] Ericsson, Motorola, and Nokia, "Link evaluation methods for high speed downlink packet access (HSDPA)," TSG-RAN Working Group 1 Meeting #15, TSGR1#15(00)1093, Tech. Rep., Aug. 2000.

[26] K. Hyun and D. Yoon, "Bit metric generation for Gray coded QAM signals," *IEE Proc.-Commun.*, vol. 152, no. 6, pp. 1134–1138, Dec. 2005.

[27] M. S. Raju, R. Annavajjala, and A. Chockalingam, "BER analysis of QAM on fading channels with transmit diversity," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 481–486, Mar. 2006.

[28] S. Lin and J. D. J. Costello, *Error Control Coding*, 2nd ed. Upper Saddle River, New Jersey 07458, USA: Prentice-Hall, Inc., 2003.

[29] R. H. Morelos-Zaragoza, *The Art of Error Correcting Codes*, 1st ed. New York, NY: John Wiley & Sons, 2002.

[30] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *IEEE International Conf. on Commun., ICC 95*, Jun. 1995, pp. 1009–1013.

[31] M. Valenti and J. Sun, "The UMTS turbo code and an efficient decoder implementation suitable for software defined radios," *International Journal on Wireless Information Networks*, vol. 8, no. 4, pp. 203–216, Oct. 2001.

[32] W. Gross and P. Gulak, "Simplified MAP algorithm suitable for implementation of turbo decoders," *Electronics Letters*, vol. 34, no. 16, pp. 1577–1578, Aug. 1998.

[33] A. J. Viterbi and J. K. Omura, *Principles of Digital Communications and Coding*. McGraw-Hill, 1979.

[34] C. E. Shannon, "A mathematical theory of communications," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and Oct. 1948.

[35] T. Cover and J. Thomas, *Elements of Information Theory*. New York, USA: Wiley series in telecommunications, John Wiley & Sons, 1991.

[36] C. Stierstorfer and R. Fischer, "Adaptive interleaving for bit-interleaved coded modulation," in *7th International ITG Conference on Source and Channel Coding (SCC)*, Ulm, Germany, Jan. 2008.

[37] A. Abedi, E. Thompson, and A. K. Khandani, "Application of cumulant method in performance evaluation of turbo-like codes," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 2037–2041, Nov. 2007.

[38] R. Redner and H. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Review*, vol. 26, pp. 195–239, Apr. 1994.

[39] J. Lindblom and J. Samuelsson, "Bounded support Gaussian mixture modeling of speech spectra," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 1, pp. 88–99, Jan. 2003.

[40] S. ten Brink, "Convergence behaviour of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[41] S. ten Brink, M. Tüchler, and J. Hagenahuer, "Measures for tracing convergence of iterative decoding algorithms," in *4th International ITG Conference on Source and Channel Coding (SCC)*, Berlin, Germany, Jan. 2002.

[42] J. Tan and G. Stüber, "Analysis and design of interleaver mappings for iteratively decoded BICM," in *IEEE International Conference on Communications, ICC 2002*, New York, USA, May 2002.

[43] L. Zhao, L. Lampe, and J. Huber, "Study of bit-interleaved coded space-time modulation with different labeling," in *IEEE Information Theory Workshop*, Paris, France, Mar. 2003, pp. 199–202.

[44] T. Clevorn, S. Godtmann, and P. Vary, "Optimized mappings for iteratively decoded BICM on rayleigh channels with IQ interleaving," in *IEEE Vehicular Technology Conference 2006, VTC-2006 Spring*, vol. 5, Melbourne, Australia, May 2006, pp. 2083–2087.

[45] J. Tan and G. Stüber, "Analysis and design of symbol mappers for iteratively decoded BICM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 662–672, Mar. 2005.

[46] E. Agrell, J. Lassing, E. G. Ström, 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.

[47] E. Agrell, J. Lassing, E. G. Ström, and T. Ottosson, "Gray coding for multilevel constellations in Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 224–235, Jan. 2007.

[48] C. Stierstorfer and R. Fischer, "(Gray) Mappings for bit-interleaved coded modulation," in *IEEE Vehicular Technology Conference 2007, VTC-2007 Spring*, Dublin, Ireland, Apr. 2007.

[49] F. Schreckenbach and G. Bauch, "Bit-interleaved coded irregular modulation," *Eur. Trans. on Telecommun.*, vol. 17, no. 2, pp. 269–282, mar.-apr. 2006.

[50] M. Benjillali, L. Szczecinski, and S. Aissa, "Probability density functions of logarithmic likelihood ratios in rectangular QAM," in *Twenty-Third Biennial Symposium on Communications*, Kingston, Canada, May 2006, pp. 283–286.

[51] L. Szczecinski and M. Benjillali, "Probability density functions of logarithmic likelihood ratios in phase shift keying BICM," in *IEEE Global Telecommunications Conference GLOBECOM 2006*, San Francisco, USA, Nov. 2006.

[52] L. Szczecinski, R. Bettancourt, and R. Feick, "Probability density function of reliability metrics in BICM with arbitrary modulation: Closed-form through algorithmic approach," in *IEEE Global Telecommunications Conference GLOBECOM 2006, San Francisco, USA*, Nov. 2006.

[53] S. Benedetto and G. Montorsi, "Average performance of parallel concatenated block codes," *Electronics Letters*, vol. 31, no. 3, pp. 156–158, Feb. 1995.

[54] ——, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.

[55] D. Divsalar, S. Dolinar, R. J. McEliece, and F. Pollara, "Transfer function bounds on the performance of turbo codes," JPL, Cal. Tech., TDA Progr. Rep. 42-121, Aug. 1995.

[56] A. Martinez, A. Guillén i Fàbregas, and G. Caire, "A closed-form approximation for the error probability of BPSK fading channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2051–2054, Jun. 2007.

[57] A. Guillén i Fàbregas, A. Martinez, and G. Caire, "Error probability of bit-interleaved coded modulation using the Gaussian approximation," in *38th Conference on Information Sciences and Systems CISS 2004*, Princeton University, NJ, USA, Mar. 2004.

Part II

Included Papers

Paper A

On the Distribution of Extrinsic L-values in Gray-mapped 16-QAM

Alex Alvarado, Leszek Szczecinski, and Rodolfo Feick

Published in
Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing, IWCMC'07
ISBN:978-1-59593-695-0, pp. 329–336
Honolulu, Hawaii, USA

Paper B

Distribution of L-values in Gray-mapped M^2-QAM: Closed-form Approximations and Applications

Alex Alvarado, Leszek Szczecinski, Rodolfo Feick, and Luciano Ahumada

Accepted for publication in
IEEE Transactions on Communications.

Paper C

Distribution of Max-Log Metrics for QAM-based BICM in Faded Channels

Leszek Szczecinski, Alex Alvarado, and Rodolfo Feick

Accepted for publication in
IEEE Transactions on Communications.