



Copyright Notice

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This document was downloaded from Chalmers Publication Library (<http://publications.lib.chalmers.se/>), where it is available in accordance with the IEEE PSPB Operations Manual, amended 19 Nov. 2010, Sec. 8.1.9 (<http://www.ieee.org/documents/opsmanual.pdf>)

(Article begins on next page)

Exploiting UEP in QAM-based BICM: Interleaver and Code Design

Alex Alvarado, *Student Member, IEEE*, Erik Agrell, Leszek Szczecinski, *Senior Member, IEEE*, and Arne Svensson, *Fellow, IEEE*

Abstract—In this paper we formally analyze the interleaver and code design for QAM-based BICM transmissions using the binary reflected Gray code. We develop analytical bounds on the bit error rate and we use them to predict the performance of BICM when unequal error protection (UEP) is introduced by the constellation labeling. Based on these bounds the optimum design of interleaver and code is found, and numerical results for representative configurations are presented. When the new design is used, the improvements may reach 2 dB, and they are obtained without any increase on the transceiver’s complexity. We also introduce the concept of generalized optimum distance spectrum convolutional codes, which are the optimum codes for QAM-based BICM transmissions.

Index Terms—BICM, interleaver design, multiple interleavers, optimum distance spectrum codes, QAM, UEP.

I. INTRODUCTION

In bit-interleaved coded modulation (BICM) [1] with high-order constellations, the bit mapping causes the so-called unequal error protection (UEP) [2], i.e., depending on the bits’ position within the symbol, the bits experience different “protection”, which may be interpreted in terms of uncoded error probability or average mutual information. In this paper we formally analyze the problem of the interleaver and code design for unequally protected BICM transmissions.

BICM, first introduced by Zehavi [1] and later analyzed in detail by Caire *et al.* in [2], owes its popularity to the fact that the channel encoder and the modulator are separated by a bit-level interleaver. Because of this separation, the code rate and the constellation can be chosen independently allowing for a simple and flexible design [2, Sec. V]. At the receiver’s side, the reliability metrics are calculated for the coded bits in the form of logarithmic likelihood ratios, also known as *L-values*. These metrics are then deinterleaved and further used by the soft-input channel decoder. From a capacity point of

view, BICM with appropriately designed mapping introduces only a small penalty when compared to a coded modulation scheme (CM) where the channel encoder and mapper are jointly designed [2]. Ungerboeck’s trellis coded modulation (TCM) [3] is one of the most popular CM schemes, and it maximizes the minimum Euclidean distance between trellis paths corresponding to different code sequences. On the other hand, BICM maximizes the code diversity, and therefore, it outperforms TCM in fading channels. When compared to TCM, BICM decreases the minimum Euclidean distance, and consequently, it is suboptimal for the AWGN channel. Nevertheless, since this decrease is only marginal, BICM is very robust to variations of the channel characteristics [4, Sec. 14.6]. BICM is nowadays a de facto standard, and it is used in most of the existing wireless systems, e.g., HSPA, IEEE 802.11a/g, IEEE 802.16, etc.

When BICM is used with Gray-mapped 4-QAM, all the bits are equally treated by the modulator. On the other hand, if UEP is produced by the modulator, it can be exploited to improve the receiver’s performance. In this paper we are interested in UEP caused by the binary labeling of high-order constellations, however, we note that UEP can also be intentionally imposed. This can be done by using unequal power allocation for systematic/parity bits, an idea first used for turbo-encoded BICM (TC-BICM) in [5] and later analyzed in [6]–[10], or by simply deleting some bits (puncturing). The conclusions available in the literature about the best strategy to exploit the UEP are somehow contradictory. According to [5], [11] the performance of turbo-encoded transmissions can be improved if the parity bits are more protected, while in [6], [8], [10] it is shown that systematic bits must receive stronger protection. The influence of the block length and code rate for optimal power allocation was analyzed in [6], [12]. It has been shown in [13] and in [14, Sec. 9.3.2] that to improve the performance of TC-BICM, the systematic bits must be assigned to the most protected positions. According to [15], in the waterfall region, puncturing systematic bits (strong protection for parity bits) improves the performance, while in [16] the opposite is claimed. Interleaver design aiming to assign the coded bits to different bit positions for high-order modulation schemes was analyzed in [17]. UEP has been studied for LDPC codes in [18]–[22], and for turbo coded modulation schemes in [23], where the bits were grouped into different classes of importance.

To take advantage of the UEP caused by the modulator, and for a given channel code, the design of the interleaver connecting both entities becomes crucial. Following the frame-

Paper approved by XYYY, the Editor for ZZZZ of the IEEE Communications Society. Manuscript received Month Day, Year; revised Month Day, Year. Research supported by the Swedish Research Council, Sweden (under research grant #2006-5599), and by NSERC, Canada (under research grant #249704-07). Different parts of this work were presented at the European Wireless Conference EW2008, Prague, Czech Republic, June 2008, and at the International Conference on Communications ICC2009, Dresden, Germany, June 2009.

A. Alvarado, E. Agrell, and A. Svensson are with the Department of Signals and Systems, Communication Systems Group, Chalmers University of Technology, Gothenburg, Sweden (e-mail: {alex.alvarado,agrell,arnes}@chalmers.se)

L. Szczecinski is with the Institut National de la Recherche Scientifique, INRS-EMT, 800, Gauchetiere W. Suite 6900 Montreal, H5A 1K6, Canada (e-mail: leszek@emt.inrs.ca).

Digital Object Identifier 10.1109/TCOMM.2009.XXXXX.

work set in [2], for the analysis of BICM, a single interleaver (S-interleaver) is most often considered. This simplifies the analysis of the resulting system, but leads to sub-optimality already noted in the literature [24]. In fact, the original BICM paper of Zehavi [1] postulated the application of multiple interleavers (M-interleavers)¹ between each of the encoder's output and the corresponding modulator's input (e.g., using three interleavers for a 2/3-rate encoder, each of them feeding bits to one of the bits' positions in the 8-PSK symbol). Similar M-interleavers have been used for BICM [24], [28], for BICM with iterative demapping and decoding (BICM-ID) [29], for serially concatenated systems [25], and for BICM-OFDM [26]. M-interleavers have also been proposed in the 3GPP/HSPA standard [14], [30] with 16-QAM or 64-QAM. Their use in that context is relevant from an implementation point of view since two parallel interleavers in HSPA with 16-QAM (or three for 64-QAM) are constructed "re-using" the interleaver already implemented for 4-QAM. When such M-interleavers are used, the performance gains will strongly depend on the bit assignment between the encoder's output and the bit positions in the complex symbol.

Although previous works we cite noted the influence of the interleaver design and the UEP, to the best of our knowledge, this paper is the first to analyze formally this problem for BICM transmissions. More particularly, we present a methodology for the interleaver and code design for QAM-based BICM transmissions (BICM-QAM). To obtain simple design rules, we use the Gaussian model for the distribution of the L-values in QAM transmissions presented in [31] and the generalized transfer function of a code [27], [32], [33], which allows us to develop union bounds for the coded bit error rate (BER) of the system. Using these bounds, the optimum design of interleaver and code is presented, proving for example that the answer about the protection of systematic/parity bits cannot be given in abstraction of the code and the modulation. As another application of the developed bounds, we introduce the generalized optimum distance spectrum (GODS) codes as the answer to the problem of selecting good convolutional codes in BICM-QAM.

II. SYSTEM MODEL

Hereafter we use lowercase letters x to denote a scalar, and boldface letters \mathbf{x} to denote a vector of scalars. Capital letters X denote random variables, $\mathbb{P}(\cdot)$ denotes probability, and $f_X(x)$ denotes the probability density function (pdf) of the random variable X . Blackboard bold letters \mathbb{X} represent matrices or vectors.

We consider the BICM system shown in Fig. 1. The k_c vectors of N information bits $\mathbf{b}_l = [b_l(1), \dots, b_l(N)]$ are encoded by a rate $R = k_c/n$ channel encoder, where $l = 1, \dots, k_c$. The vectors of coded bits $\mathbf{c}_1, \dots, \mathbf{c}_n$ are then fed to the interleaver units where the p th output vector of the encoder is given by $\mathbf{c}_p = [c_p(1), \dots, c_p(N)]$. We emphasize here that the proposed scheme is different from the so-called

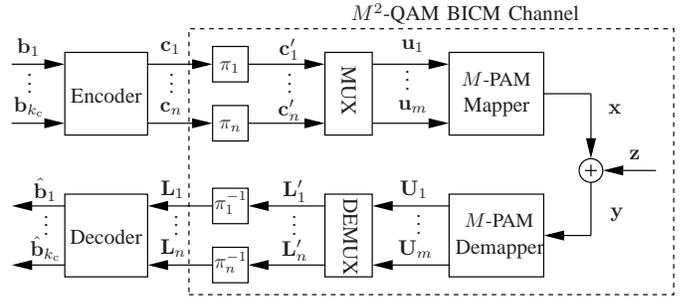


Fig. 1. Model of BICM-QAM transmission: a channel encoder followed by the interleavers (π_1, \dots, π_n), a multiplexing unit (MUX), the M -PAM mapper, the channel, and the inverse processes at the receiver's side.

multi-level coding [34] through the fact that only one encoder is present in the system.

A. The interleavers and the multiplexing unit

The interleavers (π_1, \dots, π_n) in Fig. 1 are assumed to be infinite and independent (ideal), yielding randomly permuted sequences of the coded bits $\mathbf{c}'_p = \pi_p\{\mathbf{c}_p\}$. This idealizing assumption lets us focus on the essential features of the design and is also justified by the fact that the resulting design's optimality does not seem to be affected by finite-length interleavers used during the simulations. We note that a more realistic analysis would consider finite-length (i.e., non-ideal) interleavers, however, this requires a different and more complex approach.

The multiplexing unit (MUX) assigns the coded and interleaved bits to the different bit positions in the M^2 -QAM symbol. The mapping considered here is based on the so-called binary reflected Gray code (BRGC)² [37], [38], so each symbol is a superposition of independently modulated real/imaginary parts [39]. Consequently, we focus on the equivalent M -PAM constellation (cf. Fig. 1) where $M = 2^m$.

For a fully general approach, we define the multiplexing unit using a matrix $\mathbb{K}_{n \times m} \equiv \mathbb{K}$ of dimensions $n \times m$, whose elements, $0 \leq \kappa_{p,q} \leq 1$, denote the fraction of bits \mathbf{c}'_p that will be assigned to the q th output \mathbf{u}_q . As all the vectors \mathbf{u}_q for $q = 1, \dots, m$ have the same length, so the constraint $\sum_{p=1}^n \kappa_{p,q} = \frac{n}{m}$ must be satisfied, and since all the bits in the vector \mathbf{c}'_p must be assigned to one of the m outputs, the condition $\sum_{q=1}^m \kappa_{p,q} = 1$ must also be fulfilled. The matrix \mathbb{K} can be then written as shown in (1), where the last row and the last column of \mathbb{K} take into account the constraints imposed on $\kappa_{p,q}$, and consequently, when designing \mathbb{K} , only $\kappa_{p,q}$ for $p = 1, \dots, n-1$ and $q = 1, \dots, m-1$ may be freely set (considering also $0 \leq \kappa_{p,q} \leq 1 \forall p, q$).

We emphasize that \mathbb{K} in (1) represents the multiplexing unit, i.e., it defines how the coded bits are assigned to the inputs of the modulator. This matrix and the multiple (parallel) interleavers in Fig. 1 model the whole interleaving, and allow us to consider its different configurations. For this reason we

¹Different names have been given to this interleaver: for example, "in-line" [25], "intralevel" [26], "M" [24], "dual" [14], or "modular" [27] interleavers. Its formal definition will be presented in Sec. II-A.

²The BRGC is selected for our analysis due to its relevance in practical systems, its optimality in terms of BER in uncoded transmissions [35], and also because it maximizes the "BICM capacity" for a wide range of SNRs and constellation sizes [36].

$$\mathbb{K} = \begin{bmatrix} \kappa_{1,1} & \cdots & \kappa_{1,m-1} & 1 - \sum_{q=1}^{m-1} \kappa_{1,q} \\ \kappa_{2,1} & \cdots & \kappa_{2,m-1} & 1 - \sum_{q=1}^{m-1} \kappa_{2,q} \\ \vdots & \ddots & \vdots & \vdots \\ \kappa_{n-1,1} & \cdots & \kappa_{n-1,m-1} & 1 - \sum_{q=1}^{m-1} \kappa_{n-1,q} \\ \frac{n}{m} - \sum_{p=1}^{n-1} \kappa_{p,1} & \cdots & \frac{n}{m} - \sum_{p=1}^{n-1} \kappa_{p,m-1} & 1 - n + \frac{n}{m} + \sum_{p=1}^{n-1} \sum_{q=1}^{m-1} \kappa_{p,q} \end{bmatrix}. \quad (1)$$

will refer to “interleaver design” as the process of selecting the elements $\kappa_{p,q}$ defining \mathbb{K} . For example, for $n = m$, if $\mathbb{K} = \mathbb{I}_n$ (\mathbb{I}_n being the identity matrix), the system is transformed into the Zehavi’s configuration where all the bits from the same encoder’s output are assigned to the same modulator’s input. Exchanging the rows of this matrix allows us to consider different ways of connecting the encoder to the modulator. If we consider $\kappa_{p,q} = \frac{1}{m}$ for all p and q , a uniform distribution of the coded bits at the inputs of the modulator is achieved. When comparing our model to the S-interleaver (single interleaver) in [2] we note that due to the infinite interleaver assumption, the S-interleaver also results in a uniform distribution, and therefore our model and the interleaver in [2] become equivalent.

At any time instant t , the coded and interleaved bits $[u_1(t), \dots, u_m(t)]$ are mapped to an M -PAM symbol $x(t) \in \mathcal{X}$ using a binary memoryless mapping $\mathcal{M} : \{0, 1\}^m \rightarrow \mathcal{X}$, where $\mathcal{X} = \{(1 - M)\Delta, (3 - M)\Delta, \dots, (M - 1)\Delta\}$ is the set of M -PAM symbols³, and where 2Δ is the minimum distance between them. The constellation is normalized to unit average energy so $\Delta = \sqrt{\frac{3}{2(M^2-1)}}$. The result of the transmission of N_s symbols is given by $\mathbf{y} = \mathbf{x} + \mathbf{z}$, where $\mathbf{x} = [x(1), \dots, x(N_s)]$, and $\mathbf{z} \in \mathbb{R}^{N_s}$ is a vector with samples of zero-mean and independent Gaussian random variables with variance $N_0/2$. The signal-to-noise ratio (SNR) per complex symbol is given by $\gamma = \frac{1}{N_0}$.

At the receiver’s side, the reliability metrics of the transmitted bits are calculated in the form of logarithmic likelihood ratios (L-values)⁴ for each bit position as [1], [2], [40], [41]

$$U_q(t) = \gamma \left(\min_{a \in \mathcal{X}_{q,0}} \{(x(t) - a)^2\} - \min_{a \in \mathcal{X}_{q,1}} \{(x(t) - a)^2\} \right), \quad (2)$$

where $\mathcal{X}_{q,b}$ is the set of symbols labelled with the q th bit equal to b . Since the mapping is memoryless, from now on we drop the time index t , e.g., $U_q(t) \equiv U_q$.

It is worth to mention that (2) is a suboptimal metric since it is based on the max-log approximation. This simplification, proposed in the early works of Zehavi and Caire *et al.*, is recommended by the 3GPP working groups [41] as it has small impact on the receiver’s performance when Gray-mapped constellations are used [42]–[44].

The vector of soft information \mathbf{U}_q is demultiplexed (\mathbf{L}'_p), deinterleaved (\mathbf{L}_p) and then passed to a channel decoder which

³The M^2 -QAM constellation is formed by the direct product of two M -PAM constellations, i.e., $\mathcal{X} \times \mathcal{X}$.

⁴L-values convey information about the bits’ probabilities and are often used in practice. Alternative implementations can use different metrics or the actual probabilities.

produces an estimate of the transmitted bits $\hat{\mathbf{b}}$.

B. Equivalent Channel Model

Using the results presented in [31] it is possible to build an equivalent model for the M^2 -QAM BICM channel shown in Fig. 1. In this model each bit u_q after the MUX can be seen as being sent over a *virtual channel* whose output L-value U_q has a distribution that depends on the bit’s position q and the symbol sent, i.e., the value of the other bits $u_v, v \neq q$. We explain it briefly below while more details may be found in [31]. Let $d_q(x)$ denote the Euclidean distance between the symbol x and the closest symbol in the constellation with the opposite value of the bit labeling x at position q , i.e., if $x \in \mathcal{X}_{q,b}$, $b \in \{0, 1\}$, $d_q(x) = \min_{a \in \mathcal{X}_{q,1-b}} |x - a|$. Due to the properties of the BRGC, symbols with the q th bit set to 0 or 1 are clustered so that $d_q(x)$ may be at a distance that varies from 2Δ to $2\Delta \frac{M}{2^q}$. That is, when $q = m$, there is always an adjacent symbol (at distance 2Δ) with the opposite value of the bit. On the other hand, for $q = 1$, the number of possible distances is $M/2$. Since $d_q(x)$ determines the “protection” experienced by the bit, different values of $d_q(x)$ cause UEP. For $q = m$ the bits have always the same “weak” protection but for $q = 1$, depending on the value of other bits in the modulating codeword, the protection may be relatively “strong”. In Fig. 2 we show the 8-PAM constellation with BRGC and also the distances $d_q(x)$ for some symbols. All the distances are listed in Table I.

According to [31], there are $M/2$ different Gaussian distributions that can be used to model the L-values. A bit transmitted at position q passes through the virtual channel Θ_j when it is sent using a symbol x such that $d_j(x) = 2\Delta j$. Then, the L-value U_q has a distribution that may be approximated as Gaussian with mean μ_j and variance σ^2 , where

$$(\mu_j, \sigma^2) = (4\gamma\Delta^2(2j - 1), 8\gamma\Delta^2), \quad (3)$$

with $j = 1, \dots, M/2$. It is worth to mention that the equivalent model presented in this section is slightly different from the one presented in [2]. While both of them consider m parallel binary-input soft-output channels, in our model we use the knowledge of the densities of the L-values. These densities were previously calculated in [31] and are based on the use of the max-log approximation. Moreover, in order to make the analysis tractable, we use the simplified Gaussian model for these densities as proposed in [31].

The probability that an L-value at bit position q is distributed

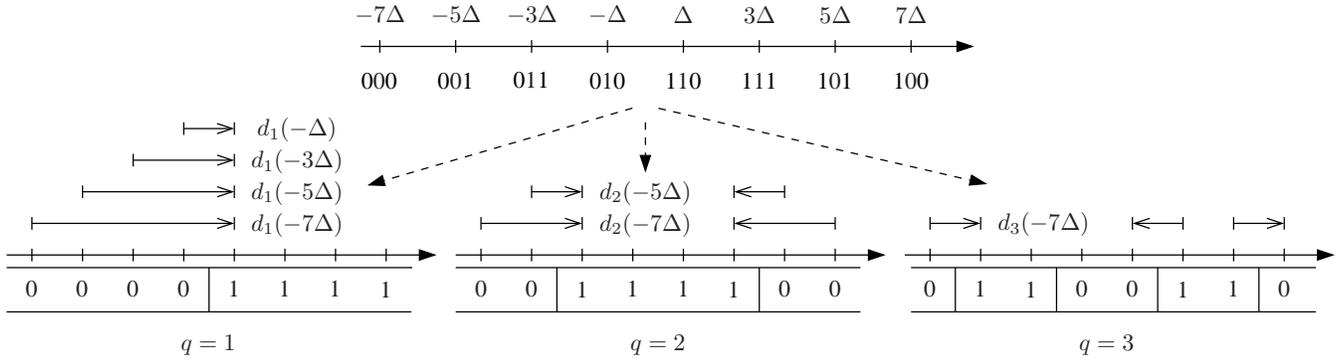


Fig. 2. 8-PAM constellation with BRGC. The binary labelings per position are shown together with the distances $d_q(x)$ for some symbols. The weaker protection of the bit position $q = 3$ is evident due to the smaller (on average) values of $d_3(x)$.

TABLE I
UEP CAUSED BY THE BRGC: MODULATING CODEWORDS, 8-PAM
SYMBOLS, DISTANCES $d_q(x)$, AND VIRTUAL CHANNELS Θ_j .

$[u_1 \dots u_m]$	000	001	011	010	110	111	101	100
x	-7Δ	-5Δ	-3Δ	$-\Delta$	Δ	3Δ	5Δ	7Δ
$d_1(x)$	8Δ	6Δ	4Δ	2Δ	2Δ	4Δ	6Δ	8Δ
Θ_j	Θ_4	Θ_3	Θ_2	Θ_1	Θ_1	Θ_2	Θ_3	Θ_4
$d_2(x)$	4Δ	2Δ	2Δ	4Δ	4Δ	2Δ	2Δ	4Δ
Θ_j	Θ_2	Θ_1	Θ_1	Θ_2	Θ_2	Θ_1	Θ_1	Θ_2
$d_3(x)$	2Δ							
Θ_j	Θ_1							

with parameters (μ_j, σ^2) is given by

$$\omega_{q,j} = \begin{cases} \frac{1}{2^{m-q}} & \text{if } j = 1, \dots, 2^{m-q} \\ 0 & \text{if } j = 2^{m-q} + 1, \dots, M \end{cases}, \quad (4)$$

that is, the virtual channel Θ_1 can be used by the bit for all positions q , Θ_2 only for $q \leq m-1$, Θ_3 and Θ_4 only for $q \leq m-2$, $\Theta_5, \dots, \Theta_8$ for $q \leq m-3$, and so on. It is worth to mention that for the BRGC, all the points in the constellation have only one closest neighbor with the opposite bit label at the same distance (cf. Fig. 2). This is a property of the mapping analyzed in this paper, and it does not hold in general.

To fully characterize the equivalent M^2 -QAM BICM channel we define the matrix $\mathbb{O}_{m \times \frac{M}{2}} \equiv \mathbb{O}$ of dimensions $m \times M/2$ where each element $\omega_{q,j}$ in \mathbb{O} is the probability that a transmitted bit at position q is transmitted using the channel Θ_j . The resulting equivalent channel model is schematically shown in Fig. 3.

Based on the previous discussion, the M^2 -QAM BICM channel of Fig. 1 can be replaced by a ‘‘compound’’ channel completely defined by the matrices \mathbb{K} (interleaver) and \mathbb{O} (mapping). If we define the matrix \mathbb{X} as

$$\mathbb{X} \triangleq \mathbb{K}\mathbb{O} = \begin{bmatrix} \sum_{q=1}^m \kappa_{1,q} \omega_{q,1} & \dots & \sum_{q=1}^m \kappa_{1,q} \omega_{q,M/2} \\ \vdots & \ddots & \vdots \\ \sum_{q=1}^m \kappa_{n,q} \omega_{q,1} & \dots & \sum_{q=1}^m \kappa_{n,q} \omega_{q,M/2} \end{bmatrix}, \quad (5)$$

then the p th output $L_p \in \mathbb{R}$ of this channel is associated with the p th binary input c_p , where L_p is a Gaussian mixture with density given by

$$f_{L_p}(\lambda) = \sum_{j=1}^{M/2} \xi_{p,j} \Phi(\mu_j, \sigma^2; \lambda), \quad (6)$$

where $\Phi(\mu_j, \sigma^2; \lambda) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\lambda-\mu_j)^2}{2\sigma^2}\right)$ is a Gaussian function, and $\xi_{p,j}$ is the (p, j) th element of \mathbb{X} which denotes the probability that the p th bit passes through the channel Θ_j .

Example 1: Consider a rate $R = 1/3$ ($n = 3$) code and an 8-PAM constellation ($m = 3$, $M = 8$) presented in Table I. In this table the virtual channels associated with the different symbols and bit positions are shown. For this case, we consider two matrices \mathbb{K}

$$\mathbb{K}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbb{K}'' = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}, \quad (7)$$

and the matrix \mathbb{O} is given by

$$\mathbb{O} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \quad (8)$$

While the matrix \mathbb{K}' represents Zehavi’s configuration, we note that the entries of the matrix \mathbb{K}'' are equal to $1/m$, which means that—thanks to the infinite interleaving—the encoder output bits are uniformly distributed over all m inputs of the modulator, and therefore, the M -interleaver represented by \mathbb{K}'' is equivalent to the S -interleaver postulated by Caire *et al.* in [2].

III. INTERLEAVER AND CODE DESIGN

In this section, based on the model introduced in Sec. II and using a generalized transfer function of a code, we develop union bounds on the BER of BICM-QAM. Based on these bounds the optimum design of interleaver and code is found and later used in Sec. IV to answer simple questions such as: What are the attainable gains obtained by using M -interleavers? Which bits (systematic/parity) should receive stronger protection? What are the optimum convolutional codes in this scenario?

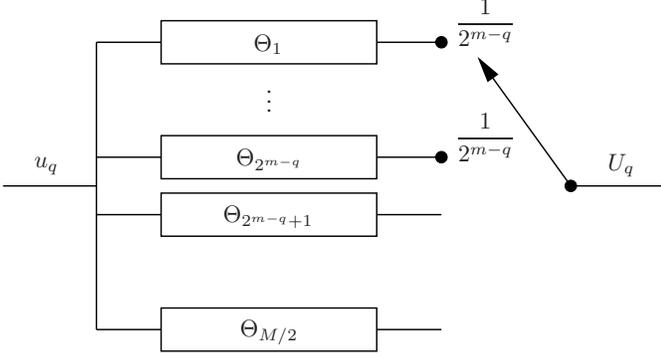


Fig. 3. Equivalent channel model: the virtual channels $\Theta_j, j = 1 \dots, 2^{m-q}$ are selected with equal probability, while the channels $\Theta_j, j = 2^{m-q} + 1, \dots, M/2$ are not available for the bit at position q .

A. Generalized weight distribution spectrum

For any convolutional code (CC) it is possible to define a *generalized transfer function* (GTF) which enumerates not only the number of non-zero output bits over a path, but the *location* of those bits, i.e., it indicates which branch the non-zero outputs are associated with [1], [27]. For a rate- k_c/n CC we define the GTF of the code as

$$T(\mathbf{W}, I, L) = \sum_{\mathbf{w}} \sum_i \sum_l t_{\mathbf{w},i,l} I^i L^l \prod_{p=1}^n W_p^{w_p}, \quad (9)$$

where the *generalized weight* $\mathbf{w} = (w_1, \dots, w_n)$ gathers the weight w_p of the p th output of the encoder, and $\mathbf{W} = (W_1, \dots, W_n)$, I , and L are dummy variables. The coefficient $t_{\mathbf{w},i,l}$ enumerates the number of paths diverging from the zero state and merging with the zero state after l steps, associated with an input sequence of weight i , and an output sequence of generalized weight \mathbf{w} . The coefficients $t_{\mathbf{w},i,l}$ can be calculated using standard techniques [45, Ch. 4]. Efficient methods for this calculation include the recursive algorithm of Divsalar *et al.* [46], or a breadth first search algorithm [47].

Using the GTF, it is possible to obtain a generalized weight distribution spectrum (GWDS) of the code [1], [45, Ch. 4]

$$\beta(\mathbf{w}) = \frac{1}{k_c} \frac{1}{\prod_{p=1}^n w_p!} \left[\frac{\partial^w}{\partial \mathbf{W}^{\mathbf{w}}} \frac{\partial}{\partial I} T(\mathbf{W}, I, L) \right] \Big|_{\mathbf{w}=0, I=L=1},$$

where $\frac{\partial^w}{\partial \mathbf{W}^{\mathbf{w}}} = \frac{\partial^{w_1}}{\partial W_1^{w_1}} \dots \frac{\partial^{w_n}}{\partial W_n^{w_n}}$ and $w = w_1 + \dots + w_n$.

If a turbo code (TC) is considered, the concept of *uniform interleaver* introduced by Benedetto *et al.* [48] can be used to calculate the spectrum of the code. The extension to a GWDS is straightforward; more details can be found in [46], [48], [49].

B. Union bounds for BICM-QAM

In order to use the GWDS of the code to calculate union bounds for the BER, 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., $\mathcal{W}_i(l) \triangleq \{(w_1, \dots, w_i) \in (\mathbb{Z}^+)^i : w_1 + \dots + w_i = l\}$. Using the GWDS of the code, the union bound (UB) on the BER for both convolutionally and turbo

coded BICM is given by

$$\text{BER} \leq \text{UB} = \sum_{l=w_{\text{free}}}^{\infty} \sum_{\mathbf{w} \in \mathcal{W}_n(l)} \beta(\mathbf{w}) \text{PEP}(\mathbf{w}), \quad (10)$$

where w_{free} is the free distance of the code, and $\text{PEP}(\mathbf{w})$ is the pairwise error probability which represents the probability of detecting a codeword with generalized weight \mathbf{w} instead of the transmitted all-one codeword.⁵ Obviously, and for practical reasons, the bound in (10) is calculated using only a limited number of terms in the first sum. This means that (10) is not a UB anymore, but rather its approximation. Nevertheless, throughout this paper we will use the name UB to refer to approximations of the true bound.

To calculate the PEP we need to calculate the probability that the decoder selects a codeword with generalized weight \mathbf{w} instead of the transmitted all-one codeword. To this end, we note that the decision is made based on the sum of $w_1 + \dots + w_n$ L-values in the divergent path. Let Z be the decision variable where

$$Z = \sum_{i=1}^{w_1} L_1^{(i)} + \dots + \sum_{i=1}^{w_n} L_n^{(i)} = \sum_{p=1}^n \sum_{i=1}^{w_p} L_p^{(i)}, \quad (11)$$

i.e., a sum of l independent random variables, where the random variable associated with the i th output is a sum of i.i.d. Gaussian mixtures given by (6). Consequently, for a given value of \mathbf{w} , the PEP can be calculated as the tail integral of the pdf of Z , i.e.,

$$\text{PEP}(\mathbf{w}) = \mathbb{P}(Z < 0) = \int_{-\infty}^0 f_Z(\lambda) d\lambda. \quad (12)$$

To calculate $f_Z(\lambda)$ we first define the j -fold self convolution operator as follows. Let L be a random variable with density $f_L(\lambda)$, its j -fold self convolution is denoted by

$$[f_L(\lambda)]^{*(j)} \triangleq \underbrace{f_L(\lambda) * \dots * f_L(\lambda)}_{j \text{ times}}, \quad (13)$$

which corresponds to the PDF of the sum of j i.i.d. random variables L .

Using the above notation and (6), we can calculate the PDF of the decision variable Z in (11) as

$$f_Z(\lambda) = [f_{L_1}(\lambda)]^{*(w_1)} * \dots * [f_{L_n}(\lambda)]^{*(w_n)}, \quad (14)$$

where the p th term in (14) can be approximated⁶ by

⁵We note that the constellation labeling produces a non-symmetric channel, i.e., the conditional channel transition probability for a bit $b = 0$ is not the same that for $b = 1$. Consequently, the exact value of the PEP in (10) depends on both \mathbf{w} and the transmitted codeword. However, the symmetry condition can be easily fulfilled if the bits at the encoder output are randomly negated and the sign of the L-values at the decoder input changed afterwards. Moreover, numerical results showed that this symmetrization causes negligible impact on the performance of QAM-based BICM transmissions.

⁶The approximation refers to the fact that the Gaussian model for the L-values is used instead of the exact densities.

$$[f_{L_p}(\lambda)]^{*(w_p)} \approx \left[\sum_{j=1}^{M/2} \xi_{p,j} \Phi(\mu_j, \sigma^2; \lambda) \right]^{*(w_p)} \quad (15)$$

$$= \sum_{j_1=1}^{M/2} \dots \sum_{j_{w_p}=1}^{M/2} \Phi \left(\sum_{i=1}^{w_p} \mu_{j_i}, w_p \sigma^2; \lambda \right) \prod_{i=1}^{w_p} \xi_{p,j_i} \quad (16)$$

$$= \sum_{\mathbf{r} \in \mathcal{W}_{M/2}(w_p)} \binom{w_p}{\mathbf{r}} \Phi \left(\sum_{j=1}^{M/2} r_j \mu_j, w_p \sigma^2; \lambda \right) \prod_{j=1}^{M/2} \xi_{p,j}^{r_j}. \quad (17)$$

To pass from (15) to (16) we have expanded the convolution of sums as sums of convolutions and then applied $\Phi(\mu_i, \sigma_i^2; \lambda) * \Phi(\mu_j, \sigma_j^2; \lambda) = \Phi(\mu_i + \mu_j, \sigma_i^2 + \sigma_j^2; \lambda)$. To pass from (16) to (17) we note that a Gaussian function with parameters $(r_1 \mu_1 + \dots + r_{M/2} \mu_{M/2}, w_p \sigma^2)$ can be generated by different combinations of (j_1, \dots, j_{w_p}) . Furthermore, the number of combinations (multiplicities) for a given value of $\mathbf{r} = (r_1, \dots, r_{M/2})$ are the multinomial coefficients given by

$$\binom{w_p}{\mathbf{r}} \triangleq \frac{w_p!}{r_1! \cdot \dots \cdot r_{M/2}!}. \quad (20)$$

Using (17) in (14) we get the final and exact expression for the density of Z shown in (18) and (19), where

$$g(\mathbf{r}_1, \dots, \mathbf{r}_n) = \prod_{p=1}^n \left[\binom{w_p}{\mathbf{r}_p} \prod_{j=1}^{M/2} \xi_{p,j}^{r_{p,j}} \right]. \quad (21)$$

Based on the previous discussion, we present three propositions which are the main results of this section. They will help us to simplify the design of the system (cf. Sec. IV).

Proposition 1: The UB on the BER for BICM-QAM can be approximated as

$$\text{UB} \approx \sum_{l=w_{\text{free}}}^{\infty} \sum_{\mathbf{w} \in \mathcal{W}_n(l)} \beta(\mathbf{w}) \sum_{\mathbf{r}_1, \dots, \mathbf{r}_n} g(\mathbf{r}_1, \dots, \mathbf{r}_n) \cdot \mathbb{Q}(h(\mathbf{r}_1, \dots, \mathbf{r}_n)), \quad (22)$$

where

$$h(\mathbf{r}_1, \dots, \mathbf{r}_n) = \frac{\sum_{p=1}^n \sum_{j=1}^{M/2} r_{p,j} \mu_j}{\sqrt{l \sigma^2}}, \quad (23)$$

$g(\mathbf{r}_1, \dots, \mathbf{r}_n)$ is given by (21), $\mathbb{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt$, and $\mathbf{r}_p \in \mathcal{W}_{M/2}(w_p)$ for $p = 1, \dots, n$.

Proof: From (10), (12), and (19). ■

Analyzing the expression in (22), it is possible to see that it is composed of three terms: $\beta(\mathbf{w})$ which depends only on the code, $\mathbb{Q}(h(\mathbf{r}_1, \dots, \mathbf{r}_n))$ which depends only on the channel [cf. (23)], and $g(\mathbf{r}_1, \dots, \mathbf{r}_n)$ which depends on the interleaver [cf. (5)]. Expressing the UB in this way shows how to optimize the BICM-QAM transmissions. In particular, we note that the channel properties defined by \mathbb{Q} are fixed for a given value of M , and that the optimum performance of the system will be achieved by a joint design of the interleaver *and* the code. We also note that all combinations in (18) are in general tedious

to evaluate (especially for large values of n and/or m), thus we seek further approximations.

The simplification presented in the following proposition is based on considering, for each l , only the Gaussian density with the smallest mean-to-standard deviation ratio. The intuition behind this approximation is that the error coefficients generated by other Gaussian densities are less important.

Proposition 2: The UB in (22) can be further approximated by

$$\text{UB}' = \sum_{l=w_{\text{free}}}^{\infty} \mathbb{Q}(\sqrt{2l\gamma\Delta^2}) \sum_{\mathbf{w} \in \mathcal{W}_n(l)} \beta(\mathbf{w}) \prod_{p=1}^n \xi_{p,1}^{w_p}. \quad (24)$$

Proof: Approximate $\mathcal{W}_{M/2}(w_p)$ in the third sum of (22) by its leading element $\mathbf{r}_p = (w_p, 0, \dots, 0)$. Then $g(\mathbf{r}_1, \dots, \mathbf{r}_n) = \prod_{p=1}^n \xi_{p,1}^{w_p}$ from (21) and $h(\mathbf{r}_1, \dots, \mathbf{r}_n) = \sqrt{l} \mu_1 / \sigma = \sqrt{2l\gamma}\Delta$ from (23) and (3). Now (24) follows from (22). ■

We emphasize here that (24) is quite simple to evaluate compared with the original expression in (22), and it still takes into account the parameters to optimize the transmission (interleaver and code).

The following proposition presents an even simpler asymptotic approximation of the original expression in (22), i.e., when the SNR goes to infinity. This result will provide us with the new criteria to select the optimum code and interleaver design (cf. Sec. IV-B).

Proposition 3: The asymptotic performance of BICM-QAM is given by

$$\text{UB}'' = \mathbb{Q}(\sqrt{2\gamma\Delta^2 w_{\text{free}}}) \sum_{\mathbf{w} \in \mathcal{W}_n(w_{\text{free}})} \beta(\mathbf{w}) \prod_{p=1}^n \xi_{p,1}^{w_p}. \quad (25)$$

Proof: The bound (22) is a sum of weighted Q-functions, whose argument $h(\mathbf{r}_1, \dots, \mathbf{r}_n)$ depends on the number of bits that were transmitted using the different virtual channels. If $\gamma \rightarrow \infty$, only one of those Q-functions will dominate the bound, i.e., the Q-function with the smallest argument. For a given value of \mathbf{w} we need to choose the combination of $(\mathbf{r}_1, \dots, \mathbf{r}_n)$ that minimizes $h(\mathbf{r}_1, \dots, \mathbf{r}_n)$, i.e.,

$$\begin{aligned} \min_{\mathbf{r}_1, \dots, \mathbf{r}_n} \left\{ h(\mathbf{r}_1, \dots, \mathbf{r}_n) \right\} &= \min_{\mathbf{r}_1, \dots, \mathbf{r}_n} \left\{ \frac{\sum_{p=1}^n \sum_{j=1}^{M/2} r_{p,j} \mu_j}{\sqrt{l \sigma^2}} \right\} \\ &\triangleq \min_{\mathbf{r}_1, \dots, \mathbf{r}_n} \left\{ \sum_{p=1}^n \sum_{j=1}^{M/2} r_{p,j} \mu_j \right\} \\ &= \min_{\mathbf{r}_1, \dots, \mathbf{r}_n} \left\{ \sum_{j=1}^{M/2} r_{1,j} \mu_j + \dots + \sum_{j=1}^{M/2} r_{n,j} \mu_j \right\}. \quad (26) \end{aligned}$$

Since $\mu_j > 0, j = 1, \dots, M/2$ and $\mu_j > \mu_1, j = 2, \dots, M/2$, it is clear that $\mathbf{r}_p = (w_p, 0, \dots, 0) \forall p$ minimizes (26).

Using the previous result and the definitions of μ_j and σ^2 in (3), it can be seen that the function $h(\mathbf{r}_1, \dots, \mathbf{r}_n)$ has a minimum value of $\sqrt{2\gamma\Delta^2 l}$. Moreover, if l is increased, the argument of the dominant Q-function will increase and consequently, the minimum is obtained when $l = w_{\text{free}}$, i.e., when all the w_{free} bits were transmitted using the least

$$f_Z(\lambda) = \sum_{\mathbf{r}_1 \in \mathcal{W}_{M/2}(w_1)} \binom{w_1}{\mathbf{r}_1} \Phi \left(\sum_{j=1}^{M/2} r_{1,j} \mu_j, w_1 \sigma^2; \lambda \right) \prod_{j=1}^{M/2} \xi_{1,j}^{r_{1,j}} * \dots * \sum_{\mathbf{r}_n \in \mathcal{W}_{M/2}(w_n)} \binom{w_n}{\mathbf{r}_n} \Phi \left(\sum_{j=1}^{M/2} r_{n,j} \mu_j, w_n \sigma^2; \lambda \right) \prod_{j=1}^{M/2} \xi_{n,j}^{r_{n,j}} \quad (18)$$

$$= \sum_{\mathbf{r}_1 \in \mathcal{W}_{M/2}(w_1)} \dots \sum_{\mathbf{r}_n \in \mathcal{W}_{M/2}(w_n)} g(\mathbf{r}_1, \dots, \mathbf{r}_n) \Phi \left(\sum_{p=1}^n \sum_{j=1}^{M/2} r_{p,j} \mu_j, \sigma^2 \sum_{p=1}^n w_p; \lambda \right). \quad (19)$$

protected channel Θ_1 . The weighting coefficient in (25) can be obtained using the definition of \mathbb{X} in (5).

By combining the results presented above, (25) can be obtained. \blacksquare

For the numerical evaluation of (22) and (24), l will be limited between w_{free} and l_{max} .

IV. NUMERICAL RESULTS

A. UB for BICM-QAM

In this section we contrast the bound in (22) with the results obtained based on numerical simulations. With these results we aim to quantify the potential gains when M-interleavers are used instead of S-interleavers, and also to confirm the analytical developments presented in Sec. III.

For a spectral efficiency of 1 bit/s/Hz, two cases are analyzed. A rate-1/2 TC or CC is used in conjunction with 16-QAM ($n = 2$ and $m = 2$), and a rate-1/3 TC or CC is used with 64-QAM ($n = 3$ and $m = 3$). For the CC we use ODS codes from [50] with polynomials given in octal notation and where the p th polynomial generator is associated with the p th encoder's output. For the TC, two identical rate-1/2 recursive systematic convolutional (RSC) encoders are concatenated in parallel separated by a single interleaver of length N . Even if formally the rate-1/2 TC has three outputs (systematic bits, parity bits from the RSC1 and from the RSC2), here we make no distinction between the parity bits, and we consider them to be one output.

For $n = m = 2$ we see from (1) that there is only one degree of freedom when selecting \mathbb{K} ($\kappa_{1,1}$). In Fig. 4 the bound (22) is compared with the simulation results⁷ for the values of $\kappa_{1,1}$ that yield the two M-interleavers ($\kappa_{1,1} \in \{0, 1\}$) and the S-interleaver ($\kappa_{1,1} = 1/2$).

Let us first analyze the CC case. From Fig. 4 we note that the simulation results perfectly match the analytical bounds. For this particular code, the best interleaver design—denoted by \mathbb{K}_B —is obtained when $\kappa_{1,1} = 1$, i.e., when the bits coming from the first encoder's output (generator polynomial 23) are more protected by the channel than the second encoder's output. The worst interleaver design—denoted by \mathbb{K}_W —is obtained when $\kappa_{1,1} = 0$, while the S-interleaver—denoted by \mathbb{K}_S —gives a performance between \mathbb{K}_B and \mathbb{K}_W . From the two-dimensional GWDS of this particular code, we observed that the non-zero elements $\mathbf{w} = (w_1, w_2) \in \mathcal{W}_n(w_{\text{free}})$ are not “balanced”, i.e., the weights w_1 are on average larger than the weights w_2 . Using this code property in Proposition 3, one

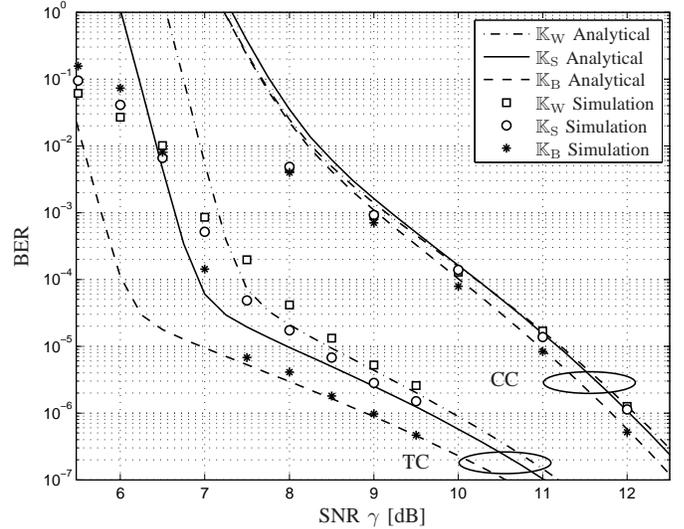


Fig. 4. UB (22) (Proposition 1) and simulated BER for BICM-QAM for TC and CC: $n = 2$, $m = 2$ ($R = 1/2$ and 16-QAM) and different interleaver configurations (23, 35). The CC is the ODS code with $K = 5$ and polynomial generators (23, 35). The TC is a parallel concatenation of two identical RSCs defined by their polynomial generators (1, 5/7). Alternate puncturing of the parity bits is performed to reach $R = 1/2$. The interleaver size is $N = 1000$ and 10 iterations are performed by the turbo decoder.

can easily demonstrate that protecting more the bits from the first output will decrease the UB. The difference between the two configurations is relatively small (0.3 dB at $\text{BER} = 10^{-6}$, cf. Fig. 4), however, we will see in the following that for other codes, or code rates, the gains can be much more important.

If the rate-1/2 TC is used instead, the optimum interleaver \mathbb{K}_B is achieved setting $\kappa_{1,1} = 0$, i.e., when the parity bits are more protected than the systematic bits (and \mathbb{K}_W if $\kappa_{1,1} = 1$). This contradicts [14, Sec. 9.3.2] and [13], where it is claimed that systematic bits should always be sent to the more reliable positions. However, using the developed bounds, we see that the optimum assignment depends on the code defined by its GWDS. In Fig. 4 these results are presented, where the bound (22) perfectly predicts the error floor of the TC. We emphasize that for this code, and for a target BER of 10^{-6} , the difference between \mathbb{K}_B and \mathbb{K}_W is 1 dB, which is obtained without complexity increase but only by properly assigning the coded bits to the bit positions in the QAM symbol.

If we analyze the asymptotic behaviour of this code using Proposition 3, we discover that the bound (25) is tight only for very high SNR values ($\text{BER} \approx 10^{-12}$). The reason behind this is the so-called spectral thinning property of the TCs, i.e.,

⁷To calculate the bound in (22) numerically, we used $l_{\text{max}} = 100$ for the TC and $l_{\text{max}} = 50$ for the CC. The interleaver size for the TC is $N = 1000$.

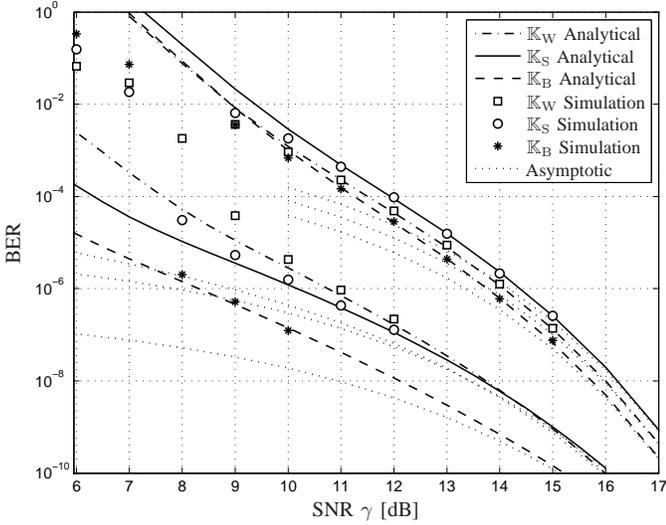


Fig. 5. UB (22) (Proposition 1) and simulated BER for BICM-QAM for TC and CC: $n = 3$, $m = 3$ ($R = 1/3$ and 64-QAM) and different interleave configurations. The CC is the ODS code with $K = 5$ and polynomial generators (25, 33, 37). The TC is a parallel concatenation of two identical RSCs defined by their polynomial generators (1, 5/7). The interleaver size is $N = 1500$ and 10 iterations are performed by the turbo decoder. The asymptotic bounds based on (24) (Proposition 2) for the TC and on (25) (Proposition 3) for the CC are also shown.

the values of the GWDS for $\mathbf{w} \in \mathcal{W}_n(w_{\text{free}})$ are quite small. To analyze the TC in the error floor region, we will thus use Proposition 2 since it considers more terms in the spectrum, cf. (24).

In Fig. 5 we present the bounds and the result of numerical simulations for a rate-1/3 TC or CC used in conjunction with 64-QAM ($n = 3$ and $m = 3$). In this case, the optimization space is formed by the variables $\kappa_{1,1}$, $\kappa_{1,2}$, $\kappa_{2,1}$, and $\kappa_{2,2}$, under the constraints presented in Sec. II-A. The variables of the optimization space are in general continuous, however, we only analyze the six possible M-interleavers ($\kappa_{p,q} \in \{0, 1\}$) and the S-interleaver ($\kappa_{p,q} = 1/3$). The results presented in Fig. 5 are for the best and worst M-interleaver found, and also the S-interleaver. The best (or worst) M-interleaver was found by selecting the matrix \mathbb{K} that minimizes (resp. maximizes) the UB at a given target BER. The selected target BER was 10^{-6} , however, we noted that changing the target BER to any other value of practical interest (between 10^{-4} and 10^{-7}) does not change the conclusion about the best (or worst) M-interleaver. For this particular code, the matrices found are

$$\mathbb{K}_B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \mathbb{K}_W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (27)$$

For this configuration we used $N = 1500$ in order to double check the correct computation of the GWDS of the TC and the bounds. In this figure we can see again that the bound (22) match the simulation results, and that for a target BER of 10^{-6} there is difference of approximately 2 dB between \mathbb{K}_W and \mathbb{K}_B .

In order to calculate the bound (22) for $n = m = 3$ (cf. Fig. 5), we used $l_{\text{max}} = 50$ for the TC and $l_{\text{max}} = 25$ for

the CC. As mentioned before, when m and/or n increase, counting all the combinations in (22) becomes tedious, and consequently, the maximum value of l considered must be relatively small. In Fig. 5 we also present results for the (asymptotic) simplifications presented in Sec. III-B. For the CC we calculate UB'' using (25) and $l_{\text{max}} = 50$, and for the TC we calculate UB' using (24) and $l_{\text{max}} = 100$. The computations for these simplifications are very simple compared with (22), and yet they predict the asymptotic performance of the system as shown in Fig. 5.

From the results presented in Fig. 4 and Fig. 5, we can draw the following interesting conclusions:

- For a given target BER of 10^{-6} , the SNR gains between the best and the worst interleave configuration are between some tenths of dB and up to 2 dB (cf. TC in Fig. 5).
- The bound (22) is tight for BER values less than 10^{-3} for the CC and for the error floor region of the TC, while (24) and (25) can be used to predict the asymptotic performance of a TC and a CC respectively.
- Optimized M-interleavers were always better than S-interleavers for the analyzed cases.
- Improperly designed M-interleavers (\mathbb{K}_W) can degrade the system performance compared to \mathbb{K}_S . Thus, when using M-interleavers, the optimization of \mathbb{K} becomes a mandatory step.
- \mathbb{K}_S can be worse than \mathbb{K}_W (cf. for example the CC in Fig. 5), so S-interleavers cannot, in general, be considered as a “conservative” solution between \mathbb{K}_B and \mathbb{K}_W .

The assignment of the coded bits to the positions we presented can be seen as a code-dependent interleaver design that does not modify the flexibility of BICM which allows the designer to choose the encoder independently of the mapping. The proposed scheme should not be confused for example with TCM where code and mapping are jointly designed. The only difference with previous BICM designs is that here we propose an optimum way of connecting the encoder and mapper. Also note that for given values of n and m , the problem of selecting the optimum interleaver configuration (selection of \mathbb{K}) is a multidimensional optimization problem, however, the optimization was performed over only a limited number of points.

B. Optimum Interleaver and Code Design for BICM with Convolutional Codes

It is well known that ODS codes—tabulated for example in [50]—are the optimum convolutional codes for binary transmissions. However, according to (25), when UEP is introduced by the channel, the optimization criterion is different to [50, Sec. II], namely, the interleaver and the GWDS of the code must be taken into account. In this section we define the generalized optimum distance spectrum (GODS) codes, which are the optimum codes for this scenario.

For a given constraint length K , code rate R , constellation size m , and assuming that the optimum free distance w_{free} for that family of codes is known (cf. for example [50, Table I, II

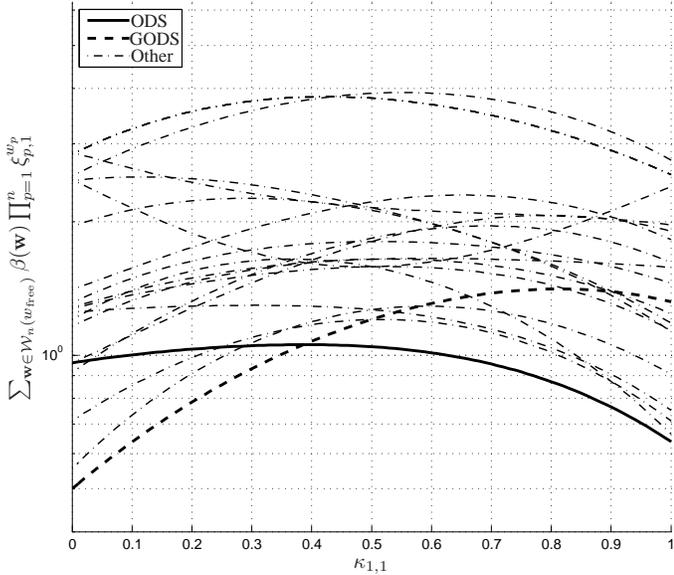


Fig. 6. Cost function in (28) for all possible codes with optimum w_{free} for $R = 1/2$, 16-QAM, and $K = 9$ as a function of the interleaver parameter. The thick solid line represents the ODS code (561, 753), and the thick dashed line the new code (515, 677).

or III]), any combination of code and interleaver will produce an asymptotic BER given by (25).

Definition 1: A GODS convolutional code ($\mathcal{C}_{\text{GODS}}$) is a code that—using an optimized interleaver configuration (\mathbb{K}_{GODS})—produces an asymptotic BER which is a minimum compared to the values that any other encoder and interleaver combination can generate, i.e.,

$$[\mathcal{C}_{\text{GODS}}, \mathbb{K}_{\text{GODS}}] = \arg \min_{\mathcal{C}, \mathbb{K}} \left\{ \sum_{\mathbf{w} \in \mathcal{W}_n(w_{\text{free}})} \beta(\mathbf{w}) \prod_{p=1}^n \xi_{p,1}^{w_p} \right\}, \quad (28)$$

where \mathcal{C} belongs to the set of all codes with optimum w_{free} .

Using the previous definition, an exhaustive search for pairs $[\mathcal{C}_{\text{GODS}}, \mathbb{K}_{\text{GODS}}]$ with constraint length up to $K = 10$ was performed. Three different configurations were tested: code rate $R = 1/2$ ($n = 2$) and 16-QAM ($m = 2$), 64-QAM ($m = 3$) or 256-QAM ($m = 4$). The optimization space for \mathbb{K} in these cases was $\kappa_{1,1} \in \{0, 1/2, 1\}$ for $m = 2$, $\kappa_{1,1}, \kappa_{1,2} \in \{0, 1/3, 2/3\}$ for $m = 3$, and $\kappa_{1,1}, \kappa_{1,2}, \kappa_{1,3} \in \{0, 1/2, 1\}$ for $m = 4$. The results are presented in Table II, where the asterisks denote codes found that are different from the ODS codes listed in [50]. Among the 24 combinations studied, 7 resulted in new optimal codes. Extension to any other combination of code rate and modulation order is straightforward.

In Fig. 6 the cost function in (28), which is the interleaver-dependent factor of UB'' , for $R = 1/2$, 16-QAM, and $K = 9$ is presented as a function of the interleaver parameter $\kappa_{1,1}$. The ODS code (561, 753) is marked with a black thick line. Analyzing this curve, it is clear that the performance of this code can be optimized by setting $\kappa_{1,1} = 1$, and that the curve has a maximum for $\kappa_{1,1} = 0.4$ which will result in the worst interleaver design for this particular code. The cost function obtained for the code (515, 677) (thick dashed line)

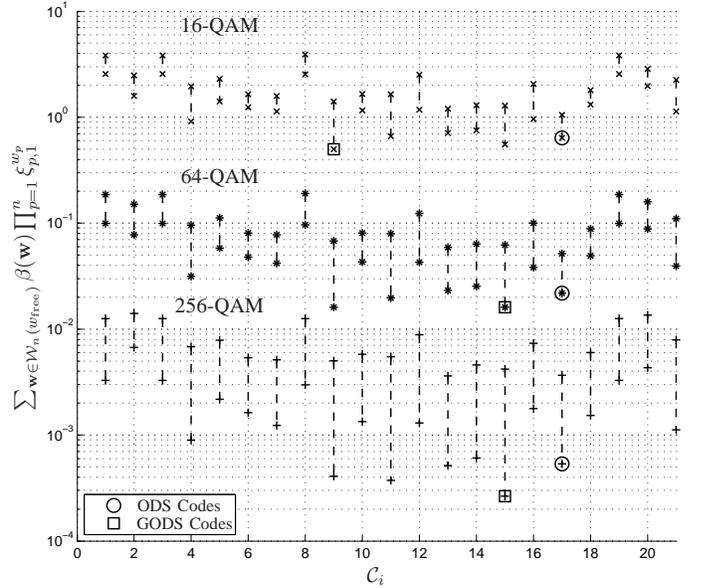


Fig. 7. Weighting coefficient of the UB in (25) for the best (\mathbb{K}_{B}) and worst (\mathbb{K}_{W}) interleave design, $K = 9$, and the 21 possible codes with $w_{\text{free}} = 12$ for $k_c = 1$, $n = 2$ ($R = 1/2$) and $m = 2$ (*), $m = 3$ (*), and $m = 4$ (+). The dashed lines represent the range of variation between the best and the worst interleave design.

is the smallest among all other codes (including the ODS one). Consequently, if the multiplexing unit is adequately designed setting $\kappa_{1,1} = 0$ (best M-interleaver), this code is the optimal code for this particular transmission with no increase of complexity. However, if the interleaver is not optimized, for example setting $\kappa_{1,1} = 1/2$ (S-interleaver), the new code is not optimal anymore.

Finally, in Fig. 7 the performance of the optimum design $[\mathcal{C}_{\text{GODS}}, \mathbb{K}_{\text{GODS}}]$ can be compared with all codes with $K = 9$ (and $w_{\text{free}} = 12$) using the best and the worst interleave design (\mathbb{K}_{B} and \mathbb{K}_{W}). The dashed lines represent the range of variation between the best and the worst interleave design, i.e., any other interleaver configuration will have a coefficient between the corresponding pair of markers. We note that the optimum design may significantly outperform other codes, e.g., 256-QAM and \mathcal{C}_{15} in Fig. 7. The improvement with respect to ODS codes is less evident but clear. Thus, the results presented in this section indicate that finding the interleaver and code should be a mandatory step in the design of BICM-QAM.

V. CONCLUSIONS

In this paper we developed analytical bounds to predict the performance of BICM with QAM schemes when UEP is introduced by the constellation labeling. Together with the original union bound, two asymptotic expressions which are simple to evaluate were developed. The analytical developments were supported by simulation results yielding accurate results.

We quantified the attainable gains when using optimized M-interleavers over S-interleavers for convolutionally-encoded and turbo-encoded schemes. These improvements can be up to 2 dB for the analyzed cases, and they can be obtained without

TABLE II
OPTIMUM INTERLEAVERS AND CODES FOR $R = 1/2$ AND 16, 64, AND 256-QAM. ASTERISK (*) DENOTES A NEW CODE, BETTER THAN THE ODS CODES.

K	w_{free}	16-QAM ($m = 2$)		64-QAM ($m = 3$)			256-QAM ($m = 4$)			
		$\mathcal{C}_{\text{GODS}}$	κ_{11}	$\mathcal{C}_{\text{GODS}}$	κ_{11}	κ_{12}	$\mathcal{C}_{\text{GODS}}$	κ_{11}	κ_{12}	κ_{13}
3	5	(5, 7)	0	(5, 7)	0	1/3	(5, 7)	1/2	1/2	0
4	6	(15, 17)	1	(15, 17)	2/3	1/3	(15, 17)	1/2	1/2	0
5	7	(23, 35)	1	(27, 31)*	0	1/3	(23, 35)	1/2	1/2	0
6	8	(53, 75)	0	(53, 75)	0	1/3	(53, 75)	1/2	1/2	0
7	10	(133, 171)	1	(135, 147)*	0	1/3	(135, 147)*	0	0	1/2
8	10	(247, 371)	1	(225, 373)*	0	1/3	(247, 371)	1/2	1/2	0
9	12	(515, 677)*	0	(557, 751)*	0	1/3	(457, 755)*	1/2	1/2	0
10	12	(1151, 1753)	0	(1151, 1753)	0	1/3	(1151, 1753)	1/2	1/2	0

complexity increase but only if the assignment of the coded bits to the bit positions in the complex symbol is optimized. We also introduced the concept of GODS codes, which are the optimum codes for the analyzed scenario.

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] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.
- [4] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw-Hill, 2008.
- [5] J. Hokfelt and T. Maseng, "Optimizing the energy of different bitstreams of turbo code," in *International Turbo Coding Seminar*, Lund, Sweden, Aug. 1996, pp. 59–63.
- [6] M. M. Salah, R. A. Raines, M. A. Temple, and T. G. Bailey, "Energy allocation strategies for turbo codes with short frames," in *The International Conference on Information Technology: Coding and Computing, ITCC 2000*, Las Vegas, USA, Mar. 2000, p. 408.
- [7] W. Zhang and X. Wang, "Optimal energy allocations for turbo codes based on distributions of low weight codewords," *Electronics Letters*, vol. 20, no. 19, pp. 1205–1206, Sep. 2004.
- [8] T. Duman and M. Salehi, "On optimal power allocation for turbo codes," in *International Symposium on Information Theory, ISIT 1997*, Ulm, Germany, June 1997, p. 104.
- [9] A. H. S. Mohammadi and A. K. Khandani, "Unequal error protection on the turbo-encoder output bits," in *IEEE International Conference on Communications, ICC 1997*, vol. 2, Montreal, Canada, June 1997, pp. 730–734.
- [10] —, "Unequal power allocation to the turbo-encoder output bits with applications to CDMA systems," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1609–1610, Nov. 1999.
- [11] M. A. Kousa and A. H. Mugaibel, "Puncturing effects on turbo codes," *Proc. IEE*, vol. 149, no. 3, pp. 132–138, June 2002.
- [12] Y. M. Choi and P. J. Lee, "Analysis of turbo codes with symmetric modulation," *Electronics Letters*, vol. 35, no. 1, pp. 35–36, Jan. 1999.
- [13] S. Le Goff, "Bandwidth-efficient turbo coding over Rayleigh fading channels," *Int. J. Commun. Syst.*, vol. 15, pp. 621–633, July 2002.
- [14] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, *3G Evolution: HSPA and LTE for Mobile Broadband*, 1st ed. Academic Press, 2007.
- [15] I. Land and P. Hoeher, "Partially systematic rate 1/2 turbo codes," in *International Symposium on Turbo Codes and Related Topics*, Brest, France, Sep. 2000, pp. 287–290.
- [16] C. Wang and Q. Vo, "Unequal protection with turbo decoding for high-order modulated signaling," in *IEEE Wireless Communications and Networking Conference, WCNC 2005*, vol. 2, New Orleans, USA, Mar. 2005, pp. 1041–1045.
- [17] E. Rosnes and Ø. Ytrehus, "On the design of bit-interleaved turbo-coded modulation with low error floors," *IEEE Trans. Commun.*, vol. 54, no. 9, pp. 1563–73, Sep. 2006.
- [18] H. Sankar, N. Sindhusayana, and K. R. Narayanan, "Design of low-density parity-check (LDPC) codes for high order constellations," in *Global Telecommunications Conference, GLOBECOM 2004*, vol. 5, Dallas, USA, Nov. 2004, pp. 3113–3117.
- [19] N. von Deetzen and S. Sandberg, "Design of unequal error protection LDPC codes for higher order constellations," in *IEEE International Conference on Communications, ICC 2007*, Washington, DC, USA, June 2007, pp. 926–931.
- [20] Y. Li, C. K. Ho, Y. Wu, and S. Sun, "Bit-to-symbol mapping in LDPC coded modulation," in *IEEE Vehicular Technology Conference, VTC-2005 Spring*, vol. 1, Stockholm, Sweden, June 2005, pp. 683–686.
- [21] R. D. Maddock and A. H. Banihashemi, "Reliability-based coded modulation with low-density parity-check codes," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 403–406, Mar. 2006.
- [22] Y. Li and W. E. Ryan, "Bit-reliability mapping in LDPC-coded modulation systems," *IEEE Commun. Lett.*, vol. 9, no. 1, pp. 1–3, Jan. 2005.
- [23] M. Aydinlik and M. Salehi, "Turbo coded modulation for unequal error protection," *IEEE Trans. Commun.*, vol. 56, no. 4, pp. 555–564, Apr. 2008.
- [24] I. Abramovici and S. Shamai, "On turbo encoded BICM," *Annales des Telecommunications*, vol. 54, no. 3-4, pp. 225–234, Mar. - Apr. 1999.
- [25] A. Nilsson and T. M. Aulin, "On in-line bit interleaving for serially concatenated systems," in *International Conference on Communications ICC 2005*, vol. 1, May 2005, pp. 547–552.
- [26] C. Stierstorfer and R. Fischer, "Intralevel interleaving for BICM in OFDM scenarios," in *12th International OFDM Workshop*, Hamburg, Germany, Aug. 2007.
- [27] X. Li, A. Chindapol, and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding and 8PSK signaling," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 1250–1257, Aug. 2002.
- [28] U. Hansson and T. Aulin, "Channel symbol expansion diversity—improved coded modulation for the Rayleigh fading channel," in *IEEE International Conference on Communications, ICC 1996*, vol. 2, Dallas, USA, June 1996, pp. 891–895.
- [29] 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.
- [30] 3GPP, "Universal mobile telecommunications system (UMTS); multiplexing and channel coding (FDD)," 3GPP, Tech. Rep. TS 25.212, V8.5.0 Release 8, Mar. 2009.
- [31] 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.
- [32] Y. S. Leung, S. G. Wilson, and J. W. Ketchum, "Multifrequency trellis coding with low delay for fading channels," *IEEE Trans. Commun.*, vol. 41, no. 10, pp. 1450–1459, Oct. 1993.
- [33] D. N. Rowitch and L. B. Milstein, "Convolutional coding for direct sequence multicarrier CDMA," in *Military Communications Conference 1995, MILCOM'95*, vol. 1, San Diego, CA, USA, 1995, pp. 55–59.
- [34] U. Wachsmann, R. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361–1391, July 1999.
- [35] E. Agrell, J. Lassing, E. G. Ström, and T. Otosson, "Gray coding for multilevel constellations in Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 224–235, Jan. 2007.
- [36] 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.
- [37] F. Gray, "Pulse code communications," U. S. Patent 2 632 058, Mar. 1953.

- [38] 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.
- [39] 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.
- [40] 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.
- [41] 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.
- [42] 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.
- [43] 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.
- [44] 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.
- [45] A. J. Viterbi and J. K. Omura, *Principles of Digital Communications and Coding*. McGraw-Hill, 1979.
- [46] 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.
- [47] 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.
- [48] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.
- [49] L. C. Perez, J. Seghers, and D. J. Costello, Jr., "A distance spectrum interpretation of turbo codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 16, pp. 1698–1709, Nov. 1996.
- [50] P. Frenger, P. Orten, and T. Ottosson, "Convolutional codes with optimum distance spectrum," *IEEE Trans. Commun.*, vol. 3, no. 11, pp. 317–319, Nov. 1999.