(Article begins on next page)

Constellation and Interleaver Design for BICM

Md. Jahangir Hossain, Alex Alvarado[§], Leszek Szczecinski[‡]
Institute for Telecommunications Research, University of South Australia, Australia
[§]Department of Signals and Systems, Communication Systems Group
Chalmers University of Technology, Gothenburg, Sweden
[‡] INRS-EMT, University of Quebec, Montreal,Canada
jahangir.hossain@unisa.edu.au, *alex.alvarado@chalmers.se*, *leszek@emt.inrs.ca*

***Abstract*—In this paper, we propose a new BICM design which considers hierarchical (nonequally spaced) constellations, a bit-level multiplexer, and multiple interleavers. It is shown that this new scheme outperform previous BICM designs because it exploits the temporal structure of the coded sequence. An analytical bound on the bit error rate (BER) of the proposed BICM scheme in terms of the constellation parameters and the multiplexing is developed for Nakagami-m fading channels. This bound is used to design a BICM transceiver with improved BER performance. Numerical results show that the gains compared to conventional BICM designs depend on the fading parameter and reach 2 dB for a target BER of 10^{-7} and $m = 5$.**

I. Introduction

Bit-interleaved coded modulation (BICM) [1]–[3] is used in most of the existing wireless communication standards, e.g., HSPA, IEEE 802.11a/g/n, DVB, etc. In BICM, the channel encoder and the modulator are separated by a bit-level interleaver which allows the designer to choose the code rate and the constellation independently. BICM appears as a simple out-of-the box coded modulation scheme, however, its full potential is achieved only if its design is optimized. For example, the unequal error protection (UEP) caused by the binary labeling of equally spaced (ES) constellations can be exploited to improve its performance [4]–[6]. More generally, if the equivalent channel seen by the encoder/decoder pair offers UEP, gains in terms of bit error rate (BER) can be obtained, cf. [4, Sec. I] and references therein.

In this paper, aiming at further gains compared to those in [4], [5], we propose to control the relative protection levels of different coded bits via non-equally spaced (NES) constellations. In particular, we use the so-called hierarchical constellations [7], which have received a great deal of attention in many applications, e.g., in simultaneous voice and multi-class data transmission [8], or in digital video broadcasting [9]. In these applications hierarchical constellations are used to transmit two or more independent data streams with different qualities at the same time. In this paper, however, we use hierarchical constellations to protect only one data stream against the channel impairments. By doing this, the BER performance of the single data stream is improved. The use of NES constellations for BICM has been analyzed in [10]–[13], but mainly using information-theoretic arguments that are not relevant for the system under consideration. The BER-minimizing approach presented in [5] should be used instead.

In order to exploit the UEP offered by the constellation, the interleaver must be properly designed. The most common interleaver considered in the literature is the single interleaver (S-interleaver) of [2], which eliminates the UEP caused by the binary labeling. UEP in BICM was in fact considered an “undesired feature” in [2, Sec. II]. Recently, the so-called multiple interleavers (M-interleavers) were shown to improve the performance of the system by exploiting the UEP in both BICM [4] and BICM with iterative decoding [6].

The previously proposed BICM systems with M-interleavers (BICM-M) [4]–[6] use a random bit-level multiplexing (R-MUX) that connects the encoder and the M-interleavers, i.e., the M-interleavers assign the coded bits to a particular bit position in the modulator in a pseudo-random fashion. By doing this, the dependency of adjacent coded bits is ignored. In this paper, we propose an multiplexing/interleaving inspired by the well-known puncturing strategy based on the periodic elimination of the bits according to a prescribed pattern that matches the temporal structure of the code, cf. [14]. We show that such a deterministic multiplexing of the coded bits outperforms the R-MUX used in [4], [5].

The main contribution of this paper is to propose and study an improved BICM system that outperform previous designs. The proposed system considers the use of hierarchical quadrature amplitude modulation (HQAM) constellations (HQAM-BICM), a deterministic (and periodic) bit-level multiplexer, and M-interleavers. To analyze the proposed system, we develop closed-form approximations for the probability density function (PDF) of the L-values as a function of the constellation parameters. These expressions are used to develop a union bound (UB) on the BER of the proposed BICM scheme for Nakagami-m fading channels. This UB is then used to optimize the BER performance of the system. Presented numerical examples show that the proposed system offers gains over previous BICM configurations (ES-QAM using an S-interleaver [2] or M-interleavers [4]). For the particular cases analyzed in this paper, the gains can reach 2 dB for a target of 10^{-7} and $m = 5$.

Research supported by the Australian Research Council (ARC) Linkage Project (under research grant #LP0991370), Australia, by NSERC, Canada (under research grant #249704-07), by the Swedish Research Council, Sweden (under research grant #2006-5599), and by The British Academy and The Royal Society (via the Newton International Fellowship scheme), UK.

Throughout this paper, we use boldface letters $\boldsymbol{c}_t = [c_{1,t}, \ldots, c_{N,t}]$ to denote row vectors and capital boldface letters $\boldsymbol{C} = [\boldsymbol{c}_1^{\mathrm{T}}, \ldots, \boldsymbol{c}_M^{\mathrm{T}}]^{\mathrm{T}}$ to denote a matrix of dimensions $M \times N$, where $(\cdot)^{\mathrm{T}}$ denotes transposition. We denote probability by $\Pr(\cdot)$ and the probability density function (PDF) of a random variable X by $f_X(x)$. A Gaussian distribution with mean value μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$, the Gaussian function with the same parameters by $\psi(\lambda; \mu, \sigma) \triangleq \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(\lambda-\mu)^2}{2\sigma^2})$, and the Q-function by $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{u^2}{2}\right) du$. We use $\mathcal{W}_i(l)$ to denote all the combinations of i nonegative integers such that their sum is l, i.e., $\mathcal{W}_i(l) \triangleq \{[w_1, \ldots, w_i] \in (\mathbb{N}_0)^i : w_1 + \ldots + w_i = l\}$.

II. System Model and Preliminaries

The proposed HQAM-BICM system model is shown in Fig. 1. Its functional blocks are described in what follows.

A. Encoder, Multiplexing, and Interleaving

The k_{c} vectors of information bits $\boldsymbol{i}_l = [i_{l,1}, \ldots, i_{l,N_{\mathrm{c}}}]$ with $l = 1, \ldots, k_{\mathrm{c}}$ are encoded by a rate $R = k_{\mathrm{c}}/n$ convolutional encoder (ENC). The vectors of coded bits $\boldsymbol{c}_p = [c_{p,1}, \ldots, c_{p,N_{\mathrm{c}}}]$ with $p = 1, \ldots, n$ are then fed to a deterministic multiplexing (S-MUX) unit which maps $\boldsymbol{C} = [\boldsymbol{c}_1^{\mathrm{T}}, \ldots, \boldsymbol{c}_n^{\mathrm{T}}]^{\mathrm{T}}$ onto $\boldsymbol{O} = [\boldsymbol{o}_1^{\mathrm{T}}, \ldots, \boldsymbol{o}_q^{\mathrm{T}}]^{\mathrm{T}}$, where the vectors $\boldsymbol{o}_k$ with $k = 1, \ldots, q$ have length N_{s} and fulfill $N_{\mathrm{s}} q = N_{\mathrm{c}} n$. The vector of bits after the S-MUX are fed to the independent interleavers $\pi_k, k = 1, \ldots, q$ that randomly permute the sequences of bits, i.e., $\boldsymbol{u}_k = \pi_k\{\boldsymbol{o}_k\}$. Each of the interleavers is connected to one of the bit positions in the hierarchical M-ary pulse amplitude modulation (M-PAM) constellation, where $q = \log_2 M$.

The S-MUX can be defined as a one-to-one mapping between the blocks of nN_{c} and qN_{s} bits, i.e., $\{0,1\}^{nN_{\mathrm{c}}} \leftrightarrow \{0,1\}^{qN_{\mathrm{s}}}$. We define it via an $n \times N_{\mathrm{c}}$ matrix $\tilde{\boldsymbol{K}}$, whose (p, t')th entry is a pair (k, t) where $k \in \{1, \ldots, q\}$ and $t \in \{1, \ldots, N_{\mathrm{s}}\}$. The entry (k, t) indicates that the bit $c_{p,t'}$ is assigned to the kth S-MUX's output at time instant t, i.e., $o_{k,t} = c_{p,t'}$. This definition of the S-MUX is entirely general but difficult to deal with, and thus, in this paper we only consider S-MUX configurations that operate periodically over blocks of nJ bits. We then represent $\tilde{\boldsymbol{K}}$ as a concatenation of N_{c}/J matrices $\boldsymbol{K}_\tau$, each of dimensions $n \times J$, i.e., $\tilde{\boldsymbol{K}} = [\boldsymbol{K}_0, \ldots, \boldsymbol{K}_{N_{\mathrm{c}}/J-1}]$, where the variable J is called the period of the S-MUX. The entries of $\boldsymbol{K}_\tau$ are pairs $(k, t + \tau nJ/q)$ with $k \in \{1, \ldots, q\}$ and $t \in \{1, \ldots, nJ/q\}$. Without loss of generality, we assume that $(N_{\mathrm{c}} \bmod J) = 0$ and that $(nJ \bmod q) = 0$. To clarify these definitions, consider the following example.

Example 1: Assume $k_{\mathrm{c}} = 1$ and $n = 2$ ($R = 1/2$), $J = 3$, and $q = 3$ (8-ary constellation). One S-MUX is

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

which results in

$$\tilde{\boldsymbol{K}} = \left[\begin{array}{ccc|ccc|c} (1,1) & (2,2) & (2,1) & (1,3) & (2,4) & (2,3) & \ldots \\ (1,2) & (3,2) & (3,1) & (1,4) & (3,4) & (3,3) & \ldots \end{array} \right].$$

The mapping between $\boldsymbol{C}$ and $\boldsymbol{O}$ is then

$$\boldsymbol{C} = \left[\begin{array}{ccc|ccc|c} c_{1,1} & c_{1,2} & c_{1,3} & c_{1,4} & c_{1,5} & c_{1,6} & \ldots \\ c_{2,1} & c_{2,2} & c_{2,3} & c_{2,4} & c_{2,5} & c_{2,6} & \ldots \end{array} \right] \Longleftrightarrow$$

$$\boldsymbol{O} = \left[\begin{array}{cc|cc|c} c_{1,1} & c_{2,1} & c_{1,4} & c_{2,4} & \ldots \\ c_{1,3} & c_{1,2} & c_{1,6} & c_{1,5} & \ldots \\ c_{2,3} & c_{2,2} & c_{2,6} & c_{2,5} & \ldots \end{array} \right]. \quad (2)$$

Since a matrix $\boldsymbol{K}_\tau$ is simply a permutation of the set $\{1, \ldots, q\} \times \{1, \ldots, nJ/q\}$, $(nJ)!$ different matrices $\boldsymbol{K}_\tau$ can be generated. However, this number can be reduced since *trivial operations* that do not affect the performance of the system can be applied to $\boldsymbol{K}_\tau$. For example, for the matrix in (1) with $\tau = 0$, consider the following two matrices:

$$\boldsymbol{K}_0' = \begin{bmatrix} (1,2) & (2,2) & (2,1) \\ (1,1) & (3,1) & (3,2) \end{bmatrix},$$

$$\boldsymbol{K}_0'' = \begin{bmatrix} (2,1) & (1,1) & (2,2) \\ (3,1) & (1,2) & (3,2) \end{bmatrix}.$$

The matrix $\boldsymbol{K}_0'$ is obtained by permuting the elements of $\boldsymbol{K}_0$ such that the first elements in the entries of $\boldsymbol{K}_0$ are not altered. Because M-interleavers are used after the S-MUX (cf. Fig. 1), the temporal structure of the sequences $\boldsymbol{o}_k$ is randomized, and thus, the second elements of the entries (which determines to which time the bit $c_{p,t'}$ is assigned) are not relevant. Consequently, the performance of the system using $\boldsymbol{K}_0$ or $\boldsymbol{K}_0'$ will be the same. The matrix $\boldsymbol{K}_0''$ is obtained by cyclically rotating the columns of the matrix $\boldsymbol{K}_0$, which will produce $\tilde{\boldsymbol{K}}$ with columns shifted to the left or right. For long coded sequences, i.e., $J \ll N_{\mathrm{c}}$, the original matrix and its shifted version will yield the same performance.

In Fig. 1, the blocks ENC and S-MUX are grouped in what is called an "equivalent code". This will be explained in Sec. IV-A and Sec. IV-B.

B. HPAM Constellations

In this paper we consider HQAM constellations labeled by the binary reflected Gray coded (BRGC) [15] under an average symbol energy constraint. If square QAM constellations labeled by BRGC are used, each symbol is a superposition of independently modulated real/imaginary parts, which allows us to focus on the real part of the constellation, i.e., the constituent one dimensional HPAM constellation, cf. Fig. 1. At any time instant t, the coded and interleaved bits $[u_{1,t}, \ldots, u_{q,t}]$ are mapped to an HPAM symbol $x^{\mathrm{I}}(t) \in \mathcal{X} = \{x_0^{\mathrm{I}}, \ldots, x_{M-1}^{\mathrm{I}}\}$ using a binary memoryless mapping $\mathcal{M} : \{0,1\}^q \to \mathcal{X}$. Since the mapper is memoryless, from now on we drop the time index t. We analyze HPAM constellations labeled with the BRGC as the one shown in Fig. 2 ($M = 8$). We use $k = 1, \ldots, q$ to denote the bit position of the binary labeling, where $k = 1$ represents the left most bit position. The bit value of $k = 1$ selects one of the two squares in Fig. 2. Similarly, for a given value of the first bit, the bit value for the next position ($k = 2$) selects one of the two triangles that surround the previously selected square. Finally, given the bit

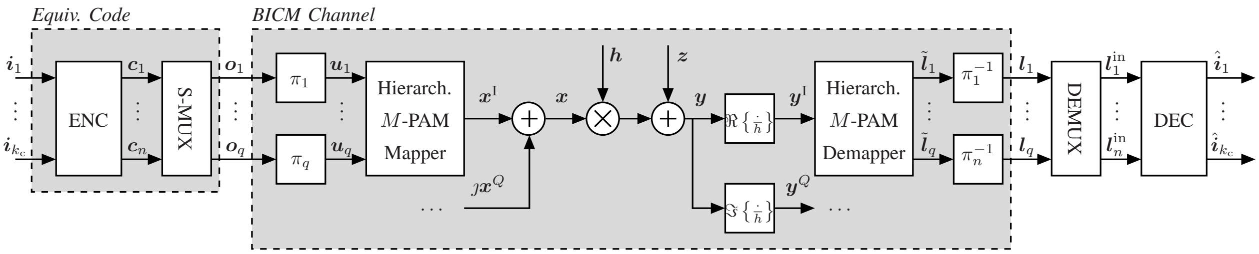

Fig. 1. Model of HQAM-BICM transmission: a channel encoder followed by the multiplexer (S-MUX), the interleavers ($\pi_1, \ldots, \pi_q$), the hierarchical M-PAM mapper, the channel, and the inverse processes at the receiver's side.

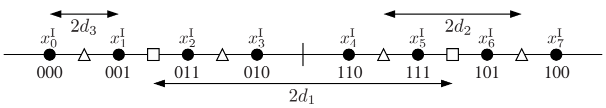

Fig. 2. HPAM ($M = 8$) constellation labeled with the BRGC.

values for $k = 1$ and $k = 2$, the bit value of bit position $k = 3$ selects one of the two black symbols that surround the previously selected triangle. This selected symbol (black circle) is finally transmitted by the modulator.

We denote the base-2 representation of the integer $0 \leq j \leq M-1$ by the vector $\boldsymbol{b}(j) = [b_1(j), \ldots, b_q(j)]$, where $b_1(j)$ is the most significant bit of j and $b_q(j)$ the least significant one. This allows us to express the elements $x_j^{\mathrm{I}} \in \mathcal{X}$ of the HPAM constellation as $x_j^{\mathrm{I}} = \sum_{k=1}^{q} (-1)^{b_k(j)-1} d_k$, where d_k are the distances between symbols in the constellation. Using this relationship, and for equiprobable symbol transmission, the average symbol energy can be expressed as $E_{\mathrm{s}} = (1 + \sum_{k=1}^{q-1} \alpha_k^2) d_1^2$, where $\alpha_k \triangleq \frac{d_{k+1}}{d_1}$ are the constellation parameters. In order to have $E_{\mathrm{s}} = 1$, $d_1 = \frac{1}{\sqrt{\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_{q-1}^2 + 1}}$. To assure a BRGC-labeled constellation, the following inequalities should hold

$$\alpha_k \geq \sum_{j=k+1}^{q-1} \alpha_j, \sum_{k=1}^{q-1} \alpha_k \leq 1, \text{and } \alpha_{q-1} \geq 0. \tag{3}$$

These constraints are found solving $(x_{j+1} - x_j) \geq 0$ with $j = 0, \ldots, M-2$.

C. Demultiplexing, Deinterleaving and Decoding

The result of the transmission of a complex symbol $x = x^{\mathrm{I}} + \jmath x^Q$ is given by $y = hx + z$, where $h = h^{\mathrm{I}} + \jmath h^Q$ is the complex channel gain, and the complex noise z is a sum of two independent Gaussian random variables with zero mean and variance $N_0/2$. We assume the amplitude of the channel gain $|h|$ follows a Nakagami-m distribution [16], and thus, the instantaneous signal to noise ratio (SNR), defined as $\gamma \triangleq \frac{|h|^2}{N_0}$, follows a Gamma distribution, i.e., $f_\Gamma(\gamma; \overline{\gamma}) = \frac{\gamma^{m-1}}{G(m)} \left(\frac{m}{\overline{\gamma}}\right)^m \exp\left(-\frac{m\gamma}{\overline{\gamma}}\right)$ [17, eq. (2.21)], [18, eq. (3)], where $G(m)$ is the Gamma function, $\overline{\gamma} = \mathbb{E}_\Gamma[\gamma]$ is the average SNR, and Γ is the random variable that represents the instantaneous SNR. We consider the Nakagami-m fading channel because it encompasses the nonfading and the Rayleigh fading channels.

At the receiver's side, the real part of the received signal normalized by the channel gain h ($y^{\mathrm{I}} = \Re\{y/h\}$) is passed to the demapper which computes logarithmic likelihood ratios (L-values) for each bit in the transmitted symbol. The approximated kth L-value (after using the max-log approximation) given the transmitted symbol x_j and the instantaneous SNR γ can be expressed as [1]–[4]

$$\tilde{l}_k(y^{\mathrm{I}}|x_j, \gamma) \approx \gamma \left[\min_{a \in \mathcal{X}_{k,1}} \left\{ (y^{\mathrm{I}} - a)^2 \right\} - \min_{a \in \mathcal{X}_{k,0}} \left\{ (y^{\mathrm{I}} - a)^2 \right\} \right], \tag{4}$$

where $\mathcal{X}_{k,b}$ is the set of symbols labeled with the kth bit equal to b.

The vectors of L-values calculated by the demapper are then deinterleaved, generating the sequence $\boldsymbol{l}_k = \pi_k^{-1}(\tilde{\boldsymbol{l}}_k)$ with $k = 1, \ldots, q$, cf. Fig. 1. These bits are finally demultiplexed by the demultiplexer (DEMUX), where the DEMUX is defined as $l_{p,t'}^{\mathrm{in}} = l_{k,t}$, and then passed to the channel decoder which produces an estimate of the transmitted bits.

III. L-values and Equivalent Channel Model

In order to evaluate the performance of the system in terms of BER, we develop closed-form expressions for the PDF of the L-values for HQAM-BICM transmission as a function of the constellation parameters. From now on, all the analysis is made for the constituent HPAM constellation (cf. Fig. 1). With a slight abuse of notation, we use x and y to denote the real part of the transmitted symbol and the real part of the received signal, respectively, i.e., we skip the superscript $(\cdot)^{\mathrm{I}}$.

In order to compute the PDF of $\tilde{l}_k$ in (4), we propose a generalization of the so-called consistent model (CoMod) introduced in [19]. For a given transmitted symbol x_j, the CoMod approximation states that

$$\tilde{l}_k(y|x_j, \gamma) \approx \gamma(\delta_{k,j} y + \rho_{k,j}), \tag{5}$$

where the parameters $\delta_{k,j}$ and $\rho_{k,j}$ in (5) depend on the transmitted symbol x_j and the symbol $\hat{a}_{k,j}$ closest to x_j with the opposite bit value. Using this, the symmetries of the BRGC constellations, and by expanding (4), we obtain the following expressions: $\delta_{k,j} = 2(-1)^{1-\tilde{b}_{k,j}}(\hat{a}_{k,j} - x_j)$ and $\rho_{k,j} = (-1)^{1-\tilde{b}_{k,j}}(x_j^2 - \hat{a}_{k,j}^2)$, where $\tilde{\boldsymbol{b}}_j = [\tilde{b}_{1,j}, \ldots, \tilde{b}_{q,j}]$ is the binary label of the symbol x_j. Using the values of $\delta_{k,j}$ and $\rho_{k,j}$, the L-values in (5) can be expressed as

$$\tilde{l}_k(y|x_j, \gamma) = (-1)^{1-\tilde{b}_{k,j}} (\hat{a}_{k,j} - x_j) \left[2y - \hat{a}_{k,j} - x_j\right] \gamma. \tag{6}$$

TABLE I
VALUES OF $\mu_{k,j}$ FOR $M = 8$ IN (8).

k	x_j	$\mu_{k,j}$
1	x_0	$+4d_1^2$
	x_1	$+4(d_1 - d_3)^2$
	x_2	$+4(d_1 - d_2)^2$
	x_3	$+4(d_1 - d_2 - d_3)^2$
2	x_0	$+4d_2^2$
	x_1	$+4(d_2 - d_3)^2$
3	x_0	$+4d_3^2$

The received signal follows a Gaussian distribution, i.e., $Y \sim \mathcal{N}(x_j, 1/(2\gamma))$. Thus, the L-values in (6) are Gaussian distributed too, i.e.,

$$\tilde{L}_k(x_j, \gamma) \sim \mathcal{N}(\gamma\mu_{k,j}, \gamma\sigma_{k,j}^2), \tag{7}$$

where

$$\mu_{k,j} = (-1)^{2-\tilde{b}_{k,j}}(\hat{a}_{k,j} - x_j)^2, \tag{8}$$
$$\sigma_{k,j}^2 = 2(\hat{a}_{k,j} - x_j)^2 = 2|\mu_{k,j}|. \tag{9}$$

From (8) and (9), it is clear that the Gaussian distribution in (7) is entirely determined by $\mu_{k,j}$ in (8) shown also in Table I in terms of the distances d_k for $M = 8$. These values can be obtained by direct inspection of Fig. 2. Due to the symmetry of the constellation, in this table, and for a given bit position, we only consider symbols with $\tilde{b}_{k,j} = 0$ that give different mean values. Table I shows that for $k = 1$ there are four different mean values, for $k = 2$ two different mean values, and for $k = 3$ only one. This idea can be generalized, i.e., from Table I, it is possible to infer that for a given k, there are $M_k \triangleq \frac{M}{2^k}$ different mean values, which are determined by the first M_k values of j. This idea was previously used in [19]. From the evolution of $\mu_{k,j}$ in Table I, we can write a generic closed-form expression for $\mu_{k,j}$ (for any M) as follows:

$$\mu_{k,j} = 4\left(d_k - \sum_{k'=k+1}^{q} \breve{b}_{k'-k}(j) d_{k'}\right)^2, \tag{10}$$

where $j = 0, 1, \ldots, M_k - 1$, $k = 1, 2, \ldots, q$, and $\breve{\boldsymbol{b}}(j) = [\breve{b}_q(j), \ldots, \breve{b}_1(j)]$ is the binary representation of the integer j where $\breve{b}_1(j)$ is the least significant bit.

Using the approximation for the PDF of the L-values presented previously, it is possible to build an equivalent model for the *BICM channel* shown in Fig. 1. A given bit $o_{k,t} = 0$ can be transmitted through one of the M_k *virtual channels*, where the jth virtual channel is determined by $\mu_{k,j}$. Because of the interleavers, the probability of $o_{k,t}$ being transmitted through the jth virtual channel is $\xi_{k,j} = \frac{1}{M_k}$, and therefore, the approximated PDF of the L-value L_k passed to the decoder can be expressed as a Gaussian mixture with density given by

$$f_{L_k}(\lambda; \gamma) = \frac{2^k}{M} \sum_{j=1}^{M/2^k} \psi(\lambda; \gamma\mu_{k,j}, 2\gamma\mu_{k,j}). \tag{11}$$

IV. PERFORMANCE ANALYSIS

In this section, and based on (11), we develop a UB on the BER of the HQAM-BICM system proposed in Sec. II.

A. Union Bound

The ENC and the S-MUX can be grouped into an "equivalent code" shown, as shown in Fig. 1. The (truncated) UB on the BER can be then expressed as

$$\text{BER} \leq \text{UB} = \frac{1}{k_c} \sum_{w=w^{\text{free}}}^{\hat{w}} \sum_{\boldsymbol{w} \in \mathcal{W}_q(w)} \beta_{\mathbb{K}}(\boldsymbol{w}) \text{PEP}(\boldsymbol{w}; \overline{\gamma}), \tag{12}$$

where $\text{PEP}(\boldsymbol{w}; \overline{\gamma})$ is the (pairwise error) probability, which represents the probability that the decoder selects a codeword with generalized weight $\boldsymbol{w} = [w_1, \ldots, w_q] \in (\mathbb{N}_0)^q$ (cf. [6]) instead of the transmitted all-zero codeword, and $\beta_{\mathbb{K}}(\boldsymbol{w})$ is the weight distribution spectrum (WDS) of the equivalent code.

B. Computation of $\beta_{\mathbb{K}}(\boldsymbol{w})$

In order to compute the WDS $\beta_{\mathbb{K}}(\boldsymbol{w})$, we need to consider an average over all path diverging at different time instants. This is similar to the computation of the WDS of punctured convolutional codes, cf. [14, Sec. II-B]. Due to the modularity of the S-MUX, we obtain

$$\beta_{\mathbb{K}}(\boldsymbol{w}) = \frac{1}{J} \sum_{j=1}^{J} \beta_{\mathbb{K}}^{(t+j)}(\boldsymbol{w}), \tag{13}$$

where $\beta_{\mathbb{K}}^{(t)}(\boldsymbol{w})$ represents the WDS for divergent paths at time t. The spectrum $\beta_{\mathbb{K}}(\boldsymbol{w})$ can be numerically calculated using a breadth first search algorithm and must be truncated so that only diverging paths with total Hamming weight $w_1 + \ldots + w_q \leq \hat{w}$ are considered.

C. Computation of $\text{PEP}(\boldsymbol{w}; \overline{\gamma})$

For fading channels, and because of the interleavers, each L-value passed to the decoder is affected by a different fading realization. We denote these channel realizations by the vector $\boldsymbol{\gamma} \triangleq [\gamma_{1,1}, \ldots, \gamma_{1,w_1}, \ldots, \gamma_{q,1}, \ldots, \gamma_{q,w_q}]$. The conditional PEP($\boldsymbol{w}; \boldsymbol{\gamma}$) can be expressed in terms of the decision variable $D(\boldsymbol{w}; \boldsymbol{\gamma})$ as

$$\text{PEP}(\boldsymbol{w}; \boldsymbol{\gamma}) = \Pr\{D(\boldsymbol{w}; \boldsymbol{\gamma}) > 0\}, \tag{14}$$

where

$$D(\boldsymbol{w}; \boldsymbol{\gamma}) = \sum_{l=1}^{w_1} L_1^{(l)}(\gamma_{1,l}) + \ldots + \sum_{l=1}^{w_q} L_q^{(l)}(\gamma_{q,l}), \tag{15}$$

and where $L_k^{(l)}(\gamma_{k,l})$ are samples of random variables representing the L-values in the kth bit position (outputs of the BICM channel in Fig. 1) whose PDF is given by (11).

The conditional PEP in (14) can be written as

$$\text{PEP}(\boldsymbol{w}; \boldsymbol{\gamma}) = \int_0^{\infty} f_{D(\boldsymbol{w}; \boldsymbol{\gamma})}(\lambda; \boldsymbol{\gamma}) d\lambda, \tag{16}$$

and since the L-values in (15) are independent, the PDF of the decision variable can be expressed as

$$f_{D(\boldsymbol{w};\boldsymbol{\gamma})}(\lambda;\boldsymbol{\gamma}) = f_{L_1}(\lambda;\gamma_{1,1}) * \ldots * f_{L_1}(\lambda;\gamma_{1,w_1}) * \ldots \\ * f_{L_q}(\lambda;\gamma_{q,1}) * \ldots * f_{L_q}(\lambda;\gamma_{q,w_q}), \quad (17)$$

where $*$ denotes convolution.

Using (16), the PEP for Nakagami-m fading channels can be expressed as

$$\mathrm{PEP}(\boldsymbol{w};\overline{\gamma}) = \int_{(\mathbb{R}^+)^w} f_{\boldsymbol{\Gamma}}(\boldsymbol{\gamma};\overline{\gamma}) \int_0^\infty f_{D(\boldsymbol{w};\boldsymbol{\gamma})}(\lambda;\boldsymbol{\gamma})\mathrm{d}\lambda\mathrm{d}\boldsymbol{\gamma}, \quad (18)$$

where $\mathbb{R}^+$ denotes the set of nonegative real numbers, and due to the interleavers, we have

$$f_{\boldsymbol{\Gamma}}(\boldsymbol{\gamma};\overline{\gamma}) = \prod_{k=1}^{q}\prod_{l=1}^{w_k} f_{\Gamma_{k,l}}(\gamma_{k,l}). \quad (19)$$

Theorem 1: The UB for HQAM-BICM for Nakagami-m fading channels is given by

$$\mathrm{UB} \approx \frac{1}{k_{\mathrm{c}}} \sum_{w=w^{\mathrm{free}}}^{\hat{w}} \sum_{\mathbf{w}\in\mathcal{W}_q(w)} \beta_{\mathbb{K}}(\mathbf{w}) \sum_{\underline{\boldsymbol{r}}\in\mathcal{R}(\boldsymbol{w})} g(\underline{\boldsymbol{r}})\cdot i(\underline{\boldsymbol{r}}), \quad (20)$$

where $\underline{\boldsymbol{r}} = [\boldsymbol{r}_1,\ldots,\boldsymbol{r}_q]$, $\boldsymbol{r}_k \in \mathcal{W}_{M_k}(w_k)$ with $k=1,\ldots,q$, $\mathcal{R}(\boldsymbol{w}) \triangleq \{\mathcal{W}_{M_1}(w_1)\times\ldots\times\mathcal{W}_{M_q}(w_q)\}$, and where

$$g(\underline{\boldsymbol{r}}) = \prod_{k=1}^{q}\left(\frac{w_k!\prod_{j=1}^{M_k}\xi_{k,j}^{r_{k,j}}}{\prod_{j=1}^{M_k} r_{k,j}!}\right) \quad (21)$$

$$i(\underline{\boldsymbol{r}}) = \frac{1}{\pi}\int_0^{\pi/2}\prod_{k=1}^{q}\prod_{j=1}^{M_k}\left(1+\frac{\mu_{k,j}\overline{\gamma}}{4m\sin^2(\phi)}\right)^{-mr_{k,j}}\mathrm{d}\phi. \quad (22)$$

Proof: The proof is given in the Appendix. ■

Theorem 1 show the the UB for the HQAM-BICM system for Nakagami-m fading channels. We note that the developments in this subsection can be considered a generalization of the PEP calculation presented in [18] in the sense that we consider multiple interleavers and HQAM constellations.

V. Numerical Results

We use a rate $R = 1/2$ optimum distance spectrum (ODS) convolutional codes (CC) with constraint length $K = 3$ and generator polynomial $(5,7)_8$. The decoding is based on the soft-input Viterbi algorithm without memory truncation, and the block length used for simulation is $N_{\mathrm{c}} = 24.000$. We consider 8-ary constellations ($q = 3$), which together with the rate $R = 1/2$ code ($n = 2$). In this case, the constellation is defined by the pair (α_1,α_2). We consider S-MUX configurations with the shortest possible period, i.e., $J = 3$, for which there will be a total of thirty different S-MUX configurations. The UB are evaluated for $\hat{w} = 30$ and optimized numerically via an exhaustive search for given values of $\bar{\gamma}$ and m. We note that the objective function (UB) is potentially non-convex and the optimization space is one- or two-dimensional. Because of this, an exhaustive seach is a feasible and robust alternative compared to other (more complex) optimization approaches.

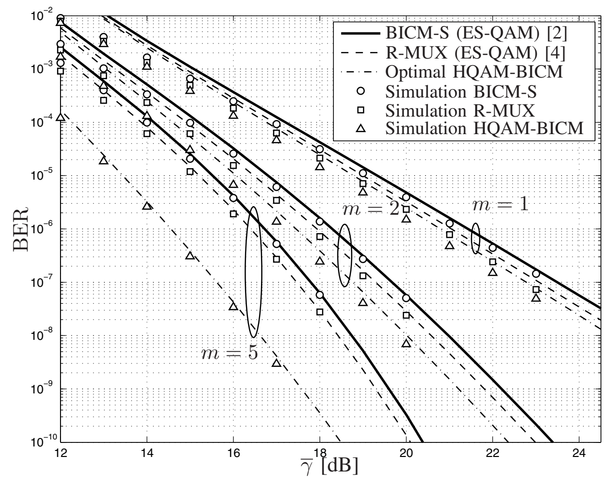

Fig. 3. UB (lines) for $n = 2$ and $q = 3$ (1.5 bit/symbol) for Nakagami-m fading channels given by Theorem 1 for $m = 1, 2, 5$. Numerical simulations are also included with markers.

Since the UB in Theorem 1 depends on $\mathbb{K}$, (α_1,α_2), m, and $\overline{\gamma}$, in general, the optimization must be done jointly over all these parameters. However, we have observed that for a given value of m, the optimal constellation and S-MUX do not change significantly for the SNR of interest. Motivated by this observation, we have found the optimal constellation for an average SNR that gives a BER of approximately 10^{-7}, and we have used that particular constellation for the range of SNRs in Fig. 3. The obtained values are $\mathbb{K}^*_{m=1} = \mathbb{K}^*_{m=2} = \mathbb{K}^{(1)}$ (cf. Example 1),

$$\mathbb{K}^*_{m=5} = \begin{bmatrix}(1,1) & (2,1) & (3,1)\\ (3,2) & (2,2) & (1,2)\end{bmatrix}, \quad (23)$$

and

$$(\alpha_1^*,\alpha_2^*)\big|_{m=1}^{\mathrm{BER}=10^{-7}} = (0.48, 0.20), \quad (24)$$

$$(\alpha_1^*,\alpha_2^*)\big|_{m=2}^{\mathrm{BER}=10^{-7}} = (0.47, 0.17), \quad (25)$$

$$(\alpha_1^*,\alpha_2^*)\big|_{m=5}^{\mathrm{BER}=10^{-7}} = (0.42, 0.01). \quad (26)$$

In Fig. 3 we present the UB obtained for HQAM-BICM using (23)–(26) and we compare it against the UB of two previous BICM designs: the BICM system based on an S-interleaver and an ES 8-PAM constellation of [2], and the BICM system with the optimal random MUX (R-MUX) in [4]. The simulations results are also shown to match well the developed bounds, in particular for high values of m. The proposed HQAM-BICM outperforms previous BICM designs. When compared to the configuration in [2], the proposed system offer gains up to 2 dB for $m = 5$ for at BER target of 10^{-7}. The gains compared to the configuration in [4] are smaller yet quite large (particularly for $m = 5$). This figure also shows that the achievable gains increase when the fading is less severe (m increases). Due to space limitations, we omit the results obtained for the nonfading channel ($m \to \infty$), however, it is shown in [20] that the gains in this case can reach up to 3 dB.

VI. Conclusions

In this paper we proposed and studied a new BICM design that uses HQAM constellations, a deterministic bit-level multiplexer, and M-interleavers. This new design exploits both the temporal structure of the code and the shaping gains offered by an unequally spaced constellation, which results in improvements of a few decibels compared to previous BICM designs. The gains were shown to depend on the fading parameter, and they increase when the fading is less severe.

Appendix
Proof of Theorem 1

We denote by $\Phi_{L_k}(s;\gamma_{k,l})$ the two-sided Laplace transform of $f_{L_k}(\lambda;\gamma_{k,l})$ in (11). The Laplace transform of the decision variable $D(\boldsymbol{w};\boldsymbol{\gamma})$ can be expressed as

$$\Phi_{D(\boldsymbol{w};\boldsymbol{\gamma})}(s) = \prod_{k=1}^{q}\prod_{l=1}^{w_k}\Phi_{L_k}(s;\gamma_{k,l}) \tag{27}$$

$$= \prod_{k=1}^{q}\prod_{l=1}^{w_k}\sum_{j=1}^{M_k}\xi_{k,j}\exp\left(\mu_{k,j}\gamma_{k,l}(s^2-s)\right), \tag{28}$$

where to pass from (27) to (28) we used the pair of two-sided Laplace transform $\mathcal{N}(\lambda;\mu,2\mu)\Leftrightarrow\exp(\mu(s^2-s))$.

By introducing the vectors $\boldsymbol{j}_k = [j_{k,1},\ldots,j_{k,w_k}]$, and the sets $\mathcal{M}_k \triangleq \{1,\ldots,M_k\}$, and $\mathcal{J} \triangleq \mathcal{M}_1^{w_1}\times\ldots\times\mathcal{M}_q^{w_q}$, we can express $\Phi_{D(\boldsymbol{w};\boldsymbol{\gamma})}(s)$ in (28) as

$$\Phi_{D(\boldsymbol{w};\boldsymbol{\gamma})}(s) = \sum_{\boldsymbol{j}_1,\ldots,\boldsymbol{j}_q\in\mathcal{J}}\prod_{k=1}^{q}\prod_{l=1}^{w_k}\xi_{k,j_{k,l}} \cdot\exp\left(\sum_{k=1}^{q}\sum_{l=1}^{w_k}\mu_{k,j_{k,l}}\gamma_{k,l}(s^2-s)\right). \tag{29}$$

By taking the inverse Laplace transform of (29) and using (18), the $\mathrm{PEP}(\boldsymbol{w};\overline{\gamma})$ can be evaluated as

$$\mathrm{PEP}(\boldsymbol{w};\overline{\gamma}) = \sum_{\boldsymbol{j}_1,\ldots,\boldsymbol{j}_q\in\mathcal{J}}\left[\prod_{k=1}^{q}\prod_{l=1}^{w_k}\xi_{k,j_{k,l}}\right]\int_{(\mathbb{R}^+)^w} f_{\boldsymbol{\Gamma}}(\boldsymbol{\gamma};\overline{\gamma}) \cdot Q\left(\sqrt{\frac{1}{2}\sum_{k=1}^{q}\sum_{l=1}^{w_k}\mu_{k,j_{k,l}}\gamma_{k,l}}\right)\,\mathrm{d}\boldsymbol{\gamma}. \tag{30}$$

Using the alternative representation of Q-function $Q(x) = \frac{1}{\pi}\int_0^{\pi/2}\exp\left(-\frac{x^2}{2\sin^2(\phi)}\right)\,\mathrm{d}\phi$ and (19), the PEP in (30) can be calculated as

$$\mathrm{PEP}(\boldsymbol{w};\overline{\gamma}) = \sum_{\boldsymbol{j}_1,\ldots,\boldsymbol{j}_q\in\mathcal{J}}\frac{1}{\pi}\prod_{k=1}^{q}\prod_{l=1}^{w_k}\xi_{k,j_{k,l}} \cdot\int_0^{\pi/2}\prod_{k=1}^{q}\prod_{l=1}^{w_k}\left(1+\frac{\mu_{k,j_{k,l}}\overline{\gamma}}{4m\sin^2(\phi)}\right)^{-m}\mathrm{d}\phi \tag{31}$$

$$= \sum_{\underline{\boldsymbol{r}}\in\mathcal{R}(\boldsymbol{w})}\frac{1}{\pi}\prod_{k=1}^{q}\left(\frac{w_k!\prod_{j=1}^{M_k}\xi_{k,j}^{r_{k,j}}}{\prod_{j=1}^{M_k}r_{k,j}!}\right) \cdot\int_0^{\pi/2}\prod_{k=1}^{q}\prod_{j=1}^{M_k}\left(1+\frac{\mu_{k,j}\overline{\gamma}}{4m\sin^2(\phi)}\right)^{-mr_{k,j}}\mathrm{d}\phi, \tag{32}$$

where to pass from (31) to (32), we re-grouped the terms in the outer summation. The UB in (20) is obtained by using (32) in (12), which completes the proof.

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. Guillén i Fàbregas, A. Martinez, and G. Caire, "Bit-interleaved coded modulation," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 1–2, pp. 1–153, 2008.

[4] A. Alvarado, E. Agrell, L. Szczecinski, and A. Svensson, "Exploiting UEP in QAM-based BICM: Interleaver and code design," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 500–510, Feb. 2010.

[5] Md. J. Hossain, A. Alvarado, and L. Szczecinski, "BICM transmission using non-uniform QAM constellations: Performance analysis and design," in *IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, May 2010.

[6] A. Alvarado, L. Szczecinski, E. Agrell, and A. Svensson, "On BICM-ID with multiple interleavers," *IEEE Commun. Lett.*, vol. 14, no. 9, pp. 785–787, Sep. 2010.

[7] P. K. Vitthaladevuni and M.-S. Alouini, "A recursive algorithm for the exact BER computation of generalized hierarchical QAM constellations," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 297–307, Jan. 2003.

[8] Md. J. Hossain, P. K. Vitthaladevuni, M.-S. Alouini, V. K. Bhargava, and A. J. Goldsmith, "Adaptive hierarchical constellations for simultaneous voice and multi-class data transmission over fading channels," *IEEE Trans. Veh. Technol.*, vol. 55, no. 4, pp. 1181–1194, July 2006.

[9] ETSI, "Digital video broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television," ETSI, Tech. Rep. ETSI EN 300 744 V1.6.1 (2009-01), Jan. 2009.

[10] D. Sommer and G. P. Fettweis, "Signal shaping by non-uniform QAM for AWGN channels and applications using turbo coding," in *International ITG Conference on Source and Channel Coding (SCC)*, Munich, Germany, Jan. 2000.

[11] S. Y. L. Goff, "Signal constellations for bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 307–313, Jan. 2003.

[12] M. Barsoum, C. Jones, and M. Fitz, "Constellation design via capacity maximization," in *IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[13] E. Agrell and A. Alvarado, "On optimal constellations for BICM at low SNR," in *IEEE Information Theory Workshop (ITW)*, Taormina, Italy, Oct. 2009.

[14] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, vol. 36, no. 4, pp. 389–400, Apr. 1988.

[15] 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.

[16] M. Nakagami, "The m-distribution—A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*, W. G. Hoffman, Ed. Oxford, U.K.: Pergamon, 1960.

[17] M. K. Simon and M. S. Alouini, *Digital Communications over fading channels: A unified approach to perfomance analysis*, 1st ed. New York, NY: John Wiley & Sons, 2000.

[18] A. Kenarsari-Anhari and L. Lampe, "An analytical approach for performance evaluation of BICM over Nakagami-m fading channels," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1090–1101, Apr. 2010.

[19] 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.

[20] Md. J. Hossain, A. Alvarado, and L. Szczecinski, "Towards fully optimized bicm transceivers," *IEEE Trans. Commun.*, 2011 (to appear), available at http://arxiv.org/abs/1012.1799.