Improved Information Rates for Bit-Interleaved Coded Modulation

Alfonso Martinez[1], Li Peng[2], Alex Alvarado[3], and Albert Guillén i Fàbregas[1,2,4]
[1]Universitat Pompeu Fabra, [2]University of Cambridge, [3]University College London, [4]ICREA
alfonso.martinez@ieee.org, alex.alvarado@ieee.org, guillen@ieee.org

***Abstract*—This paper shows that bit-interleaved coded modulation (BICM) over the Gaussian channel can achieve information rates larger than the so-called BICM capacity. For some labelings the improvement with respect to the BICM capacity is significant, especially at low and medium signal-to-noise ratios (SNR). Specifically, natural binary labeling is found to be both first- and second-order optimal at low SNR.**

***Index Terms*—Bit-interleaved coded modulation, mismatched decoding, Gaussian channel, LM rate.**

I. Introduction and Summary

In recent years, the combination of very good performance and simple implementation offered by bit-interleaved coded modulation (BICM) has led to its widespread adoption as a *de facto* standard in modern wireless and optical systems. Despite this popularity, BICM remains somewhat poorly understood from a fundamental theoretical perspective. In [1], Caire *et al.* built an equivalent channel model for BICM around the assumption of infinite interleaving and determined an achievable rate for this model, a quantity often referred to as *BICM capacity*. Later, Martinez *et al.* [2], [3] cast the operation of BICM as an instance of mismatched decoding (i.e. a decoder that does not operate according to the Maximum-Likelihood, i.e. ML, criterion) and found that Caire's BICM capacity coincides with the so-called generalized mutual information (GMI) of the BICM decoder, even for finite-length interleaving.

In general, mismatched decoders can reliably transmit information at rates above the GMI [4], [5], e.g. the LM rate [4], [5]. In this paper, we study the LM rate for BICM over the Gaussian channel and find that it generally exceeds the GMI. For some labelings, the improvement with respect to the BICM capacity is significant. Moreover, in the wideband regime, i.e. at vanishing signal-to-noise ratio (SNR) [6], where binary reflected Gray labeling is not first-order-optimal (FOO) but natural binary labeling is, we find that the LM rate of natural binary labeling is also second-order optimal (SOO).

II. Channel Model and Notation

At the encoder, a message is mapped onto a codeword which we denote by $\boldsymbol{x}$. A codeword is a sequence of N symbols $x_k \in \mathcal{X}$, $k = 1, \dots, N$ drawn from a constellation $\mathcal{X}$ of cardinality 2^m, where m denotes the number of bits per symbol. Codewords are constructed as the serial concatenation of a binary codeword of length $n = mN$, a bit-level interleaver and a binary labeling function that takes consecutive blocks of m bits and maps them to signal constellation symbols x, such that x_k is a function of $(b_{(k-1)m+1}, \dots, b_{km})$. We label each symbol by m bits, $b_1, \dots, b_m$, and let $b_i(x)$ denote the i-th bit in the binary label of x, for $i = 1, \dots, m$.

We consider transmission over the complex additive white Gaussian noise (AWGN) channel. The input X_k belongs to the alphabet $\mathcal{X}$, which is assumed to have zero mean, i.e. $\sum_x 2^{-m} x = 0$, and unit average input energy, i.e. $\sum_x 2^{-m} |x|^2 = 1$, under a uniform probability distribution $P(x) = \frac{1}{2^m}$, $x \in \mathcal{X}$. The k-th channel output Y_k is given by

$$Y_k = \sqrt{\mathsf{snr}} X_k + Z_k \qquad k = 1, \dots, N, \tag{1}$$

where snr is the average SNR and Z_k is a sample of i.i.d. circularly-symmetric complex Gaussian noise with zero mean and unit variance. With this setup, the channel transition probability density function $W(y|x)$ is given by

$$W(y|x) = \frac{1}{\pi} e^{-|y - \sqrt{\mathsf{snr}} x|^2}. \tag{2}$$

We define $2m$ sets $\mathcal{X}_b^i$ as the collection of symbols x satisfying $b_i(x) = b$. The labeling induces $2m$ different conditional symbol distributions $P_i(x|b)$ uniform in the set $\mathcal{X}_b^i$. We let $\overline{\mathcal{X}}_b^i$ denote the respective mean of $\mathcal{X}_b^i$ under $P_i(x|b)$.

Finally, we also define m bit distributions $P_i(b) = \frac{1}{2}$, $b = 0, 1$, and $2m$ possibly different channel transition probabilities $W_i(y|b) \triangleq \sum_x P_i(x|b) W(y|x) = \sum_{x \in \mathcal{X}_b^i} \frac{2}{2^m} W(y|x)$.

The BICM decoder chooses its estimate of the transmitted codeword according to the following rule

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}} \sum_{k=1}^{N} \sum_{i=1}^{m} d_i(b_i(x_k), y_k), \tag{3}$$

where the i-th bit decoding metric $d_i(b, y)$ is given by

$$d_i(b, y) = \log W_i(y|b). \tag{4}$$

The defining feature of the BICM mismatched decoder is that it treats the m bits in a symbol as if they were independent. The corresponding symbol decoding metric $d(x, y)$ is given by

$$d(x, y) = \sum_{i=1}^{m} d_i\big(b_i(x), y\big). \tag{5}$$

This work has been funded in part by the European Research Council under ERC grant agreement 259663, by the European Union's 7th Framework Programme under grant agreement 303633, by the Spanish Ministry of Economy and Competitiveness under grants RYC-2011-08150 and TEC2012-38800-C03-03, and by the Engineering and Physical Sciences Research Council (EPSRC) project UNLOC (EP/J017582/1), United Kingdom.

As the BICM decoder makes its decision according to a metric other than ML, we have an instance of mismatched decoding. While the formulation in terms of decoding metric seems the most natural for the purposes of this paper, it is possible to present the decoder operation in terms of log-likelihood ratios, as in [1]. To any extent, this decoder characterizes the performance of BICM without requiring Caire's assumption of infinite interleaving [1] as observed in [2], [3].

III. Achievable Information Rates

Achievable information rates are often determined by considering the ensemble performance of collections of random independent identically distributed (i.i.d.) codewords [4], [5]. For instance, the GMI is the largest achievable rate when the codeword symbols are randomly selected in an i.i.d. manner according to a general distribution P, not necessarily uniform. The general expression for the GMI is [4]

$$I_0(\mathsf{snr}) \triangleq \sup_{s\geq 0} \mathbb{E}\left[\log \frac{e^{sd(X,Y)}}{\mathbb{E}[e^{sd(X',Y)}|Y]}\right], \tag{6}$$

where we denote the GMI by $I_0(\mathsf{snr})$ to emphasize the dependence on the SNR. The expectations $\mathbb{E}[\cdot]$ are carried out according to the distributions $(X,Y)\sim P\times W$ and $X'\sim P$.

The performance of the i.i.d. ensemble is weakened by the codewords whose empirical distribution is very different from the original distribution P. In contrast, considering codewords whose empirical distribution coincides with P [4] leads to an improved ensemble. In this case, the associated rate is the LM rate, which we denote by $I_1(\mathsf{snr})$. This rate is given by

$$I_1(\mathsf{snr}) \triangleq \sup_{s\geq 0, a(\cdot)} \mathbb{E}\left[\log \frac{e^{a(X)}e^{sd(X,Y)}}{\sum_{x'} P(x')e^{a(x')}e^{sd(x',Y)}}\right], \tag{7}$$

where $a(x)$ is a real-valued function with finite mean $\bar{a} = \mathbb{E}[a(X)]$. Recently, it has been observed that $a(x)$ admits an interpretation in terms of a pseudo-cost function [5], [7]. Concretely, the LM rate can be achieved by choosing codewords that satisfy a pseudo-cost constraint, $N\bar{a}-\delta < a^n(\boldsymbol{x}) \leq N\bar{a}$, where $a^n(\boldsymbol{x}) \triangleq \sum_{k=1}^{N} a(x_k)$ and the constant $\delta > 0$ limits the allowed codeword cost. Since the GMI can be recovered from the LM rate by setting $a(x)=0$ for all $x\in\mathcal{X}$, we have that $I_0(\mathsf{snr}) \leq I_1(\mathsf{snr})$. In general, however, the LM rate is not the largest achievable rate with mismatched decoding.

For BICM the supremum is attained in (6) at $s=1$ [3]. With independent and uniformly distributed bits, the GMI coincides with the BICM capacity,

$$I_0(\mathsf{snr}) = \sum_{i=1}^{m} I(B_i;Y) \tag{8}$$

$$= \sum_{i=1}^{m} \mathbb{E}\left[\log \frac{W_i(Y|B)}{\sum_{b'} \frac{1}{2} W_i(Y|b')}\right], \tag{9}$$

where $I(B_i;Y)$ represents the mutual information between the bit B_i in position i, $i=1,\dots,m$, and the channel output Y and the expectation in (9) is done with respect to the distribution $(B_i,Y)\sim P_i\times W_i$.

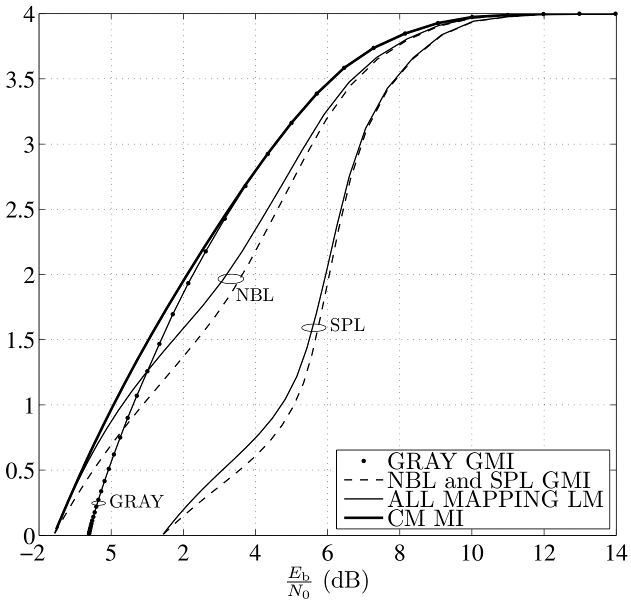

Fig. 1. Comparison of $I_0(\mathsf{snr})$ and $I_1(\mathsf{snr})$ for BICM with 16QAM modulation, an AWGN channel and different labelings. The mutual information of CM with 16QAM modulation is also shown for reference.

For a fixed input distribution, the GMI is concave in s and the LM rate is jointly concave in s and $a(x)$. One can therefore evaluate these rates easily using convex optimization routines. As no such simple decomposition exists for the LM rate, we keep the optimization problem in (7) with uniform $P(x)$:

$$I_1(\mathsf{snr}) = \sup_{s\geq 0, a(\cdot)} \mathbb{E}\left[\log \frac{e^{a(X)}\prod_{i=1}^{m} W_i\big(Y|b_i(X)\big)^s}{\sum_{x'} \frac{1}{2^m} e^{a(x')}\prod_{i=1}^{m} W_i\big(Y|b_i(x')\big)^s}\right], \tag{10}$$

where the expectation is done according to $(X,Y)\sim P\times W$.

We consider three different labelings, namely binary reflected Gray labeling (BRGL), natural binary labeling (NBL) and set-partitioning labeling (SPL), (e.g. the one in [1, Fig. 2(a)]). Fig. 1 shows the GMI and LM rate of BICM with 16QAM modulation over an AWGN channel for different labelings. We also show the mutual information of CM transmission with 16QAM for reference. We observe considerable improvement of the LM rate over the GMI for NBL and SPL, and little improvement for BRGL. The GMI of BICM with Gray labeling is expected to be larger than that with NBL or SPL and close to the mutual information of CM transmission except in the low-SNR regime [1], [2], [8]. We observe that this also happens for the LM rate.

More surprisingly, the LM rate with NBL is very close to the CM mutual information in the low-SNR regime. Although the GMI with NBL is known to achieve the minimum energy-per-bit [8], there is a gap in terms of wideband slope [6]: the GMI for NBL is first-order optimal, but it is not second-order optimal. In the next section, we analyze the wideband regime of the LM rate and prove that the LM rate for NBL is both FOO and SOO, while the LM rate of BRGL shows no improvement with respect to that of the GMI.

IV. Wideband Regime Expansion

We wish to find the wideband-regime (Taylor) series expansions of the GMI and LM rates in powers of snr, namely

$$I_0(\mathsf{snr}) = c_{1,\text{GMI}}\mathsf{snr} + c_{2,\text{GMI}}\mathsf{snr}^2 + O(\mathsf{snr}^3) \quad (11)$$

$$I_1(\mathsf{snr}) = c_{1,\text{LM}}\mathsf{snr} + c_{2,\text{LM}}\mathsf{snr}^2 + O(\mathsf{snr}^3), \quad (12)$$

for some coefficients $c_{1,\text{GMI}}$, $c_{2,\text{GMI}}$, $c_{1,\text{LM}}$, and $c_{2,\text{LM}}$. The summands $O(\mathsf{snr}^3)$ represent unspecified remainder terms that grow as snr^3 and can be neglected in the wideband regime. We recall that a communication scheme is FOO if $c_1 = 1$; the scheme is SOO if $c_1 = 1$ and $c_2 = -\frac{1}{2}$.

The following theorem summarizes our main findings:

Theorem 1: The coefficients c_1 and c_2 of the GMI and LM rates for BICM are given by

$$c_{1,\text{GMI}} = c_{1,\text{LM}} = \sum_{i=1}^{m}\sum_{b}\frac{1}{2}|\overline{\mathcal{X}}_b^i|^2 \quad (13)$$

$$c_{2,\text{GMI}} = -\frac{1}{2}\left(m\kappa(\mathcal{X}) - \sum_{i=1}^{m}\sum_{b}\frac{1}{2}\kappa(\mathcal{X}_b^i)\right), \quad (14)$$

$$c_{2,\text{LM}} = c_{2,\text{GMI}} + \frac{1}{2}\sum_{\bar{x}}\frac{1}{2^m}\left(\sum_{i,j=1;i\neq j}^{m} r\big(\overline{\mathcal{X}}^i_{b_i(x)}, \overline{\mathcal{X}}^j_{b_j(x)}\big)\right)^2. \quad (15)$$

Moreover, the wideband regime of the LM rate is achieved for $s = 1$, i.e. letting s vary with snr cannot improve the value of the coefficient, and

$$a(x) = -2\sum_{i,j=1;i\neq j}^{m} r\big(\overline{\mathcal{X}}^i_{b_i(x)}, \overline{\mathcal{X}}^j_{b_j(x)}\big)\mathsf{snr}. \quad (16)$$

In the theorem, we have used the functions $r(a,b) \triangleq ab^* + a^*b$ and $\kappa(\mathcal{X}_\mathcal{S})$, defined for a constellation $\mathcal{X}_\mathcal{S}$ with cardinality $|\mathcal{X}_\mathcal{S}|$ and mean $\overline{\mathcal{X}}_\mathcal{S}$ under the probability distribution $P_\mathcal{S}$ as

$$\kappa(\mathcal{X}_\mathcal{S}) \triangleq \left(\sum_{x\in\mathcal{X}_\mathcal{S}}\frac{1}{|\mathcal{X}_\mathcal{S}|}\big|x-\overline{\mathcal{X}}_\mathcal{S}\big|^2\right)^2 + \left|\sum_{x\in\mathcal{X}_\mathcal{S}}\frac{1}{|\mathcal{X}_\mathcal{S}|}\big(x-\overline{\mathcal{X}}_\mathcal{S}\big)^2\right|^2. \quad (17)$$

Sketch of Proof: It is convenient to express $I_0(\mathsf{snr})$ and $I_1(\mathsf{snr})$ in terms of a function $R_1(\gamma)$ for fixed s and $a(x)$:

$$R_1(\gamma) \triangleq \sum_{\chi}\frac{1}{2^m}\int W(y|\chi)\left(\log\frac{e^{a(\chi)+sd(\chi,\chi,z)}}{\sum_x P(x)e^{a(x)+sd(x,\chi,z)}}\right)\mathrm{d}y \quad (18)$$

$$= \sum_{\chi}\frac{1}{2^m}\int W(z)\left(g_\chi(\chi,z) - \log\left(\sum_x\frac{1}{2^m}e^{g_\chi(x,z)}\right)\right)\mathrm{d}z, \quad (19)$$

where $\gamma \triangleq \sqrt{\mathsf{snr}}$, we write $d(x,\chi,z)$ instead of $d(x,y)$, with χ the transmitted symbol (while decoders have no access to the value of χ, we can safely assume it known at the evaluation of the information rates) and, again with some abuse of notation, we defined a noise distribution $W(z) \triangleq W(\gamma\chi + z|\chi) = \frac{1}{\pi}e^{-|z|^2}$, and let $g_\chi(x,z) \triangleq a(x) + sd(x,\chi,z)$ be a function of χ, x, z. Conveniently, only the functions g_χ depend on γ.

As the expansions (11) and (12) are related to the first four derivatives of $R_1(\gamma)$ wrt γ evaluated at $\gamma = 0$, we first evaluate each of the four derivatives and then use them to find the coefficients $c_{1,\text{GMI}}$, $c_{2,\text{GMI}}$, $c_{1,\text{LM}}$, and $c_{2,\text{LM}}$. We do this by finding the Taylor series expansion of (19) via the series expansion of various functions of $g_\chi(x,z)$: for the expansion of $g_\chi(x,z)$, we need the power series of $d(x,\chi,z)$ and to obtain this series we require in turn that of $d_i(b,\chi,z)$ (a variant of $d_i(b,y)$ that has access to the transmitted symbol χ).

In the optimization over s and $a(x)$, and with some loss of generality, we restrict our attention to smooth functions of the SNR. More precisely, we consider functions that can be approximated by a Taylor expansion in powers of $\gamma = \sqrt{\mathsf{snr}}$. We let both s and $a(x)$ depend smoothly on SNR and expand them in powers of $\gamma = \sqrt{\mathsf{snr}}$, namely

$$s = \sum_{\ell=0,\frac{1}{2},1,\frac{3}{2},2}\frac{1}{(2\ell)!}\,s_\ell\gamma^{2\ell} \quad (20)$$

$$a(x) = \sum_{\ell=0,\frac{1}{2},1,\frac{3}{2},2}\frac{1}{(2\ell)!}\,a_\ell(x)\gamma^{2\ell}, \quad (21)$$

where the various coefficients are to be determined. The subindices ℓ denote powers of snr. As the mean of $a(x)$ has to be finite, we can safely subtract a non-zero value with no effect on the rate and we assume that $\sum_x\frac{1}{2^m}a(x) = 0$. We also let $\sum_x\frac{1}{2^m}a_\ell(x) = 0$ for all ℓ. The derivation continues by isolating the various powers of γ in each derivative to find the corresponding series coefficient. If the coefficient depends on s and/or $a(x)$, we optimize its value to maximize the rate. The process is tedious, although straightforward, and we do provide only a summary of the main steps involved. Details are omitted for the sake of conciseness.

Zeroth-order term: The zero-th order term, i.e. proportional to γ^0 or a constant, vanishes by choosing $a_0(x) = 0$. Moreover, it is possible to show that this is the best choice.

First-order term: The first-order term, i.e. proportional to γ^1, vanishes, that is $R_1'(0)$, which is consistent with the absence of a term in γ in the series expansion of the rates. Moreover, it is possible to show that this term imposes no additional constraints on the values of s or $a(x)$.

Second-order term: The second-order term, i.e. proportional to γ^2, gives the coefficients $c_{1,\text{GMI}}$ and $c_{1,\text{LM}}$ in the expansions (11) and (12). For the GMI, $c_{1,\text{GMI}}$ is given by [3]

$$c_{1,\text{GMI}} = \sum_{i=1}^{m}\sum_{b}\frac{1}{2}|\overline{\mathcal{X}}_b^i|^2. \quad (22)$$

The only possible contributions to the second-order term of (11) and (12) come from the first- and second-order derivatives of $g_\chi(x,z)$, as the zero-th order derivative vanishes. A tedious calculation gives

$$R_1''(0) = 2(2s_0 - s_0^2)\left(\sum_{i=1}^{m}\sum_{b}\frac{1}{2}|\overline{\mathcal{X}}_b^i|^2\right). \quad (23)$$

This quantity is maximized for $s_0 = 1$, the value corresponding to $I_0(\mathsf{snr})$. As $a(x)$ does not have an effect on $R_1''(0)$, we conclude that the coefficients for I_0 and I_1 coincide, that is

$$c_{1,\mathrm{GMI}} = \frac{1}{2} R_1''(0)\big|_{s=1,a(x)=0} = \frac{1}{2} \max_{s,a(x)} R_1''(0) = c_{1,\mathrm{LM}}. \quad (24)$$

After optimization over s and $a(x)$, $c_{1,\mathrm{LM}}$ and $c_{1,\mathrm{GMI}}$ coincide. Consequently, labelings which are FOO for the GMI rate (e.g. NBL [8]) remain so for the LM rate. Besides, no other labelings have an FOO LM rate.

Third-order term: The third-order term, i.e. proportional to γ^3, vanishes, i.e. $R_1'''(0) = 0$. Moreover, we find that the third-order term imposes no additional constraints on the values of s and $a(x)$.

Fourth-order term: The fourth-order term, i.e. proportional to γ^4, gives the coefficients $c_{2,\mathrm{GMI}}$ and $c_{2,\mathrm{LM}}$ in the expansions (11) and (12). For the GMI, $c_{2,\mathrm{GMI}}$ is given by [3]

$$c_{2,\mathrm{GMI}} = -\frac{1}{2}\left(m\kappa(\mathcal{X}) - \sum_{i=1}^{m} \sum_{b=0,1} \frac{1}{2}\kappa(\mathcal{X}_b^i) \right), \quad (25)$$

where $\kappa(\mathcal{X}_{\mathcal{S}})$ is defined in (17).

The only possible contributions to the fourth-order term come from the first-, second-, third-, and fourth-order derivatives of $g_{\mathcal{X}}(x,z)$. As intermediate steps in its evaluation, we find the optimum choices $s_{\frac{1}{2}} = 0$ in (20) and $a_{\frac{1}{2}}(x) = 0$ in (21). The fourth-order derivative $R_1^{(\mathrm{iv})}(0)$ is given by

$$R_1^{(\mathrm{iv})}(0) = 24 c_{2,\mathrm{GMI}} - 3 \sum_x \frac{1}{2^m}\Big(a_1(x)^2 + 2a_1(x)\zeta(x) \Big), \quad (26)$$

where

$$\zeta(x) = -2\sum_{i=1}^{m} \big|\overline{\mathcal{X}}^i_{b_i(x)}\big|^2 + 2\left|\sum_{i=1}^{m} \overline{\mathcal{X}}^i_{b_i(x)}\right|^2 \quad (27)$$

$$= 2 \sum_{i,j=1; i\neq j}^{m} r\big(\overline{\mathcal{X}}^i_{b_i(x)}, \overline{\mathcal{X}}^j_{b_j(x)}\big). \quad (28)$$

It is possible to show that

$$\sum_x \zeta(x) = 0. \quad (29)$$

In (27) we find the optimum $a_1(x)$ by using Lagrange multipliers. Let the Lagrangian $\mathcal{L}$ be

$$\mathcal{L} = \sum_x \frac{1}{2^m}\Big(a_1(x)^2 + 2a_1(x)\zeta(x) \Big) + 2\lambda \sum_x \frac{1}{2^m} a_1(x). \quad (30)$$

Taking the derivative of the Lagrangian with respect to $a_1(x)$, setting it to zero, and solving for $a_1(x)$, we conclude that $a_1(x) = -\lambda - \zeta(x)$. Then, we adjust λ so that the constraint on the mean of $a_1(x)$ is satisfied. As $\sum_x \zeta(x) = 0$ (see (29)), we conclude that $\lambda = 0$ and

$$a_1(x) = -\zeta(x). \quad (31)$$

Substituting the optimum $a_1(x)$ in (26), we obtain

$$\max_{a_1(x)} \left\{ -\sum_x \Big(a_1(x)^2 + 2a_1(x)\zeta(x) \Big) \right\} = \sum_x \zeta(x)^2. \quad (32)$$

Putting (32) back into (26), we obtain $c_{2,\mathrm{LM}}$, namely

$$c_{2,\mathrm{LM}} = \max_{s,a(x)} \frac{1}{24} R_1^{(\mathrm{iv})}(0) \quad (33)$$

$$= c_{2,\mathrm{GMI}} + \frac{1}{2}\sum_x \frac{1}{2^m}\left(\left|\sum_{i=1}^{m} \overline{\mathcal{X}}^i_{b_i(x)}\right|^2 - \sum_{i=1}^{m} \big|\overline{\mathcal{X}}^i_{b_i(x)}\big|^2 \right)^2 \quad (34)$$

$$= c_{2,\mathrm{GMI}} + \frac{1}{2}\sum_{\bar{x}} \frac{1}{2^m}\left(\sum_{i,j=1; i\neq j}^{m} r\big(\overline{\mathcal{X}}^i_{b_i(x)}, \overline{\mathcal{X}}^j_{b_j(x)}\big) \right)^2. \quad (35)$$

The wideband regime of the LM rate is thus achieved by setting $a_0(x) = a_{\frac{1}{2}}(x) = 0$ and $a_1(x)$ in (21) and $s = 1$, as letting s vary with snr does not improve the coefficients. ■

Table I shows the coefficients c_1 and c_2 for GMI and LM and square 2^m-QAM and uniform symbol probabilities. As it was known, BRGL is not FOO [9]. Moreover, we find that the coefficients c_1 and c_2 for BRGL are the same for both GMI and LM rates. In contrast, NBL turns out to be both FOO and SOO if we consider the LM rate, while it is FOO only if we consider the GMI. An interesting question arising from our analysis is whether all FOO-labelings in terms of GMI turn our to be also SOO when we consider the LM rate.

TABLE I
COEFFICIENTS c_1 AND c_2 FOR GMI AND LM AND SQUARE 2^m-QAM

	c_1		c_2	
	GMI	LM	GMI	LM
BRGL	$\frac{3\cdot 2^{2m}}{4\cdot(2^{2m}-1)} < 1$	$\frac{3\cdot 2^{2m}}{4(\cdot 2^{2m}-1)} < 1$	(14)	(35)
NBL	1	1	(14) $(< -\frac{1}{2})$	$-\frac{1}{2}$

REFERENCES

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

[2] A. Guillén i Fàbregas, A. Martinez, and G. Caire, *Bit-Interleaved Coded Modulation*. Foundations and Trends on Communications and Information Theory, Now Publishers, 2008, vol. 5, no. 1-2.

[3] A. Martinez, A. Guillén i Fàbregas, G. Caire, and F. M. J. Willems, “Bit-interleaved coded modulation revisited: A mismatched decoding perspective,” *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2756–2765, Jun. 2009.

[4] I. Csiszár and P. Narayan, “Channel capacity for a given decoding metric,” *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 35–43, Jan. 1995.

[5] A. Ganti, A. Lapidoth, and İ. E. Telatar, “Mismatched decoding revisited: general alphabets, channels with memory, and the wideband limit,” *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Jul. 2000.

[6] S. Verdú, “Spectral efficiency in the wideband regime,” *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[7] S. Shamai (Shitz) and I. Sason, “Variations on the Gallager bounds, connections, and applications,” *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3029–3051, Dec. 2002.

[8] E. Agrell and A. Alvarado, “Optimal alphabets and binary labelings for BICM at low SNR,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6650–6672, Oct. 2011.

[9] A. Martinez, A. Guillén i Fàbregas, G. Caire, and F. M. J. Willems, “Bit-interleaved coded modulation in the wideband regime,” *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5447–5455, Dec. 2008.