International Symposium on Information Theory and its Applications, ISITA2004
Parma, Italy, October 10–13, 2004

New Simple Evaluation of Error Probability of Bit-Interleaved Coded Modulation using the Saddlepoint Approximation

Alfonso MARTINEZ[†], Albert GUILLÉN i FÀBREGAS[‡], and Giuseppe CAIRE[‡]

[†] Technische Universiteit Eindhoven
Den Dolech 2, 5600 MB Eindhoven, The Netherlands
E-mail: alfonso.martinez@ieee.org

[‡] Institut Eurécom
2229 Rte. des Cretes, 06904 Sophia Antipolis, France
E-mail: {guillen,caire}@eurecom.fr

Abstract

This paper presents a simple and very accurate method to evaluate the error probability of bit-interleaved coded modulation. Modelling the channel as a binary-input continuous-output channel with a non-Gaussian transition density allows for the application of standard bounding techniques for Gaussian channels. The pairwise error probability is equal to the tail probability of a sum of random variables; this probability is then calculated with a saddlepoint approximation. Its precision is numerically validated for coded transmission over standard Gaussian noise and fully-interleaved fading channels for both convolutional and turbo-like codes. The proposed approximation is much tighter and simpler to compute than the existing techniques and reveals as a powerful tool for practical use.

1. INTRODUCTION

Binary linear codes over binary-input output symmetric channels (BIOS) have been widely studied and are relatively easy to analyze, thanks to the uniform error property [1] and to the fact that the pairwise error probability corresponds to the tail of the distribution of a sum of random variables $\sum_{j=1}^{d} \Lambda_j$, where

$$\Lambda = \log \frac{\Pr(\hat{c} = 1|\mathbf{z}, \mathbf{h})}{\Pr(\hat{c} = 0|\mathbf{z}, \mathbf{h})}. \tag{1}$$

Here $\mathbf{z}$ and $\mathbf{h}$ are the channel output and the channel state respectively. We shall refer to the variables Λ_j as a posteriori log-likelihood ratio. Figure 1 shows the location of Λ_j in the communication channel, after the demodulator. For the most usual BIOS channels, the sum has a known and easily manageable distribution. For example, for the binary symmetric channel (BSC) Λ_J are binomial random variables, for the binary-input additive white Gaussian noise (BI-AWGN) channel with signal-to-noise ratio ℓ they are normally distributed $\mathcal{N}(-4\ell, 8\ell)$. The problem of bounding error probability becomes much more complicated for codes over non-binary signal constellations, for non-symmetric channels, and for codes that do not possess the uniform error property.

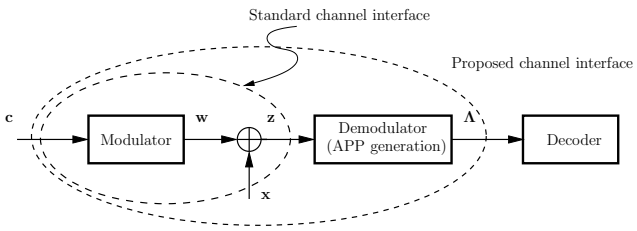


Figure 1: Channel interfaces: standard non-binary symbols at channel level, or at demodulator level, with binary symbols.

Bit-Interleaved Coded Modulation (BICM) is a pragmatic approach that maps a binary code over a sequence of high-order modulation symbols. Provided that the channel noise density function is symmetric, the performance of BICM under the assumption that the constituent binary code is linear can be studied by looking at the output of the BICM soft-demodulator. In a sense that will be made precise later, Λ_j collect the "a posteriori" statistics of the noise and fading realizations, as in binary transmission over AWGN, and, critically, of the bit index in the symbol mapping.

The analysis presented in [2] provided simple expressions for the average mutual information and cut-off rate for BICM. As explained in [2], by possibly introducing an appropriate pseudo-random binary mapping between the coded bits and the modulation symbols, the channel from the output of the binary encoder to the output of the BICM demodulator is again a BIOS channel (see Figure 1). Then, the performance of BICM schemes can be obtained straightforwardly from the tail of the distribution of $\sum_{j=1}^{d} \Lambda_j$.

In [2], this tail was bounded using the simple Chernoff Battacharyya bound (reviewed in the following). Other tighter bounds involved the sum over a restricted set of error events and the exact computation of the

pairwise error probability using numerical integration. In this paper we use more refined saddlepoint approximations for approximating the tail of $\sum_{j=1}^{d} \Lambda_j$. Moreover, our saddlepoint approximations provide a theoretical foundation for the Gaussian approximation that the authors presented in [3], and in particular apply the technique to fully-interleaved fading channels.

2. SYSTEM MODEL

We consider the transmission of bit-interleaved coded modulation (BICM) over additive white Gaussian noise (AWGN) fading channels, for which the discrete-time received signal can be expressed as

$$y_k = \sqrt{\rho}\, h_k\, x_k + z_k, \quad k = 1, \ldots, L \tag{2}$$

where $x_k \in \mathcal{X}$ are complex-valued modulation symbols with $\mathrm{E}[|x_k|^2] = 1$, $\mathcal{X}$ is the complex signal constellation, e.g., phase-shift keying (PSK), quadrature-amplitude modulation (QAM), z_k denotes the k-th noise sample modeled as i.i.d. Gaussian $\mathcal{N}_{\mathbb{C}}(0,1)$. The standard AWGN and fully-interleaved Rayleigh fading channels are obtained from (2) by simply letting $h_k = 1$ and $h_k \sim \mathcal{N}_{\mathbb{C}}(0,1)$ respectively. In this way, the average signal-to-noise ratio is $\rho = E_s/N_0$. In the case of the fully-interleaved fading channel we assume perfect channel state information (CSI) at the receiver. However, the technique described here can also be applied to the non-perfect CSI case.

The codewords $\mathbf{x} = \big(x_1, \ldots, x_L\big) \in \mathcal{X}^L$ are obtained by bit-interleaving the codewords $\mathbf{c} = \big(c_1, \ldots, c_N\big)$ of a binary code $\mathcal{C} \subseteq \mathbb{F}_2^N$ of length N and rate $r_{\mathcal{C}} = K/N$, and mapping over the signal constellation $\mathcal{X}$ with the labeling rule $\mu : \{0,1\}^M \to \mathcal{X}$, with $M = \log_2 |\mathcal{X}|$, such that $\mu\big(c_{(k-1)M+1}, \ldots, c_{kM}\big) = x_k$. The resulting length of the BICM codeword is $L = N/M$ and the total spectral efficiency is $R = r_{\mathcal{C}} M$ bits/s/Hz.

At the receiver side, we consider the classical BICM decoder that does not perform iterations at the demapper side, for which the channel demodulator computes the bitwise a posteriori probabilities of bit $c_j \in \{0,1\}$, $j = (k-1)M + m$, where $m = 1, \ldots, M$:

$$\begin{aligned} \Pr\big(c_j | y_k, h_k\big) &\propto \sum_{x \in \mathcal{X}_m^{c_j}} \Pr\big(x | y_k, h_k\big) \\ &\propto \sum_{x \in \mathcal{X}_m^{c_j}} \exp\Big(-\big|y_k - \sqrt{\rho}\, h_k\, x\big|^2\Big), \end{aligned} \tag{3}$$

where $\mathcal{X}_m^b$ denotes the set of constellation symbols with a bit of value b at the label position m.

By applying the belief propagation algorithm to the BICM code dependency graph [4], in [5], it is shown that the general iterative decoder in (3) shows substantial performance gain when $\mathcal{C}$ is a trellis-terminated convolutional code and μ is not Gray. However, this approach does not seem to show any gain when applied to turbo-like codes. Moreover, we consider that $\mathcal{C}$ is decoded with a maximum-likelihood (ML) decoder in the BICM channel. In general this does not perform ML decoding for the whole BICM, but is nevertheless a good approximation for the usual cases of turbo-like coding. Notice, however, that this particular decoder is shown to be near-optimal when coupled with Gray mapping [2] and that it is commonly employed in practical systems with capacity-approaching code ensembles.

3. GENERALIZED UNION BOUND

3.1. Binary-Input Continuous-Output Channel Equivalent for BICM

As introduced in [2] the BICM equivalent channel can be made BIOS using a time-varying mapping that uses μ and its complement $\bar{\mu}$ with probability 1/2. We further consider that binary code $\mathcal{C}$ is linear, and therefore, we consider that the all-zero codeword $\mathbf{c} = \mathbf{0}$ has been transmitted. We now define the a posteriori log-likelihood ratio as

$$\Lambda = \log \frac{\Pr(\hat{c} = 1 | y, h)}{\Pr(\hat{c} = 0 | y, h)}. \tag{4}$$

where we have dropped the time dependence for simplicity.

The a posteriori probabilities in (3) and therefore the density of the log-likelihood ratio Λ clearly depend on the transmitted symbol x, channel fading h, noise realization z and the bit position in the label m, in contrast to the binary case, where the dependence on the symbol x and on m is absent. Under the assumption of sufficient interleaving, it was shown in [2] that we can consider both m and x as nuisance parameters to be in characterized statistically, rather than with an exact analysis. The log-likelihood ratio Λ is a continuous random variable with density

$$p_\Lambda(\Lambda) = \sum_{\substack{m,x,h,z \,|\\ \log \frac{\Pr(c_m = \bar{c} | x,h,z)}{\Pr(c_m = c | x,h,z)} = \Lambda}} \Pr(m) \Pr(x) \Pr(h) \Pr(z), \tag{5}$$

where $\Pr(z)$ and $\Pr(h)$ correspond to the (continuous) probability densities of the noise and the channel state respectively. The sum (possibly an integral) is performed over all bit positions m, and over all symbols x with a bit c at position m which are all assumed

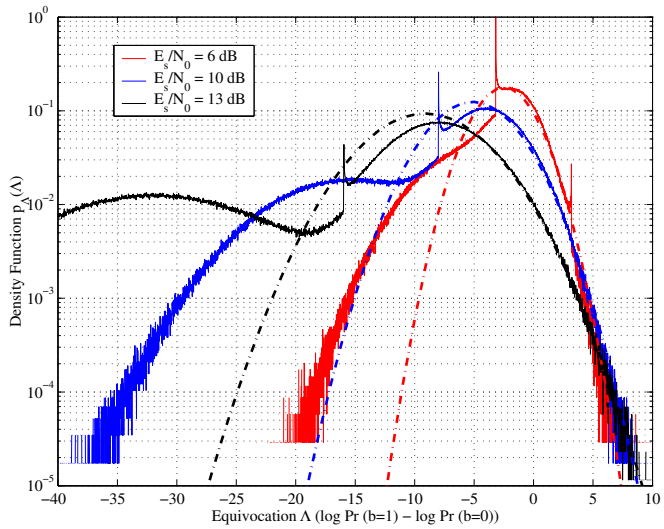


Figure 2: Density of the a posteriori log-likelihood ratio Λ and 16-QAM with Gray mapping.

equiprobable. This yields

$$p_\Lambda(\Lambda) = \sum_{\substack{m,x,h,z| \\ \log \frac{\Pr(c_m=\overline{c}|x,h,z)}{\Pr(c_m=c|x,h,z)}=\Lambda}} \frac{1}{M2^M}\, p_h(h)\, p_z(z). \tag{6}$$

The transition probabilities (5) and (6) now take into account not only the additive noise and the fading, but also the interleavers, or equivalently the label positions. This describes the binary-input continuous output equivalent BICM channel, used in [2,3] to evaluate the error probability of BICM.

Even though a closed expression for $p_\Lambda(\Lambda)$ may be difficult to obtain, it is nevertheless simple to approximate it by computer simulation. Figure 2 shows some simple results of 16-QAM signaling, with Gray mapping in the AWGN channel. We have estimated the density of Λ with computer simulations for several values of E_s/N_0, i.e., $\rho = 6\,\text{dB}$, $\rho = 10\,\text{dB}$, and $\rho = 13\,\text{dB}$. We plot the results in logarithmic scale, to better appreciate the tail behavior.

3.2. Derivation of a Generalized Union Bound

Under the assumption of BIOS channel and of $\mathcal{C}$ linear, following the standard derivation of [9] we obtain a union bound on the codeword error probability of the form

$$P_e \le \sum_d A_d \Pr\left(\sum_{j=1}^{d} \Lambda_j > 0\right) \tag{7}$$

where A_d is the distance spectrum of $\mathcal{C}$ and accounts for the number of codewords with Hamming weight d. Similarly, in order to estimate the bit-error probability P_b we need $A_{i,d}$, the number of codewords of $\mathcal{C}$ with Hamming weight d generated with input sequences with Hamming weight i. Thus we can write,

$$P_b \le \sum_d \sum_i \frac{i}{K} A_{i,d} \Pr\left(\sum_{j=1}^{d} \Lambda_j > 0\right). \tag{8}$$

Therefore, (7) and (8) have the same form, and the problem reduces then to calculating the tail probability of a sum of independent identically distributed (iid) random variables, Λ_j, with distribution $p_\Lambda(\Lambda)$ given in (6). In the next sections we shall use accurate approximations to this probability.

3.3. Bhattacharyya Union Bound

Most efficient bounds to the tail probability of a sum of random variables Λ make use of the cumulant transform $\kappa(s) = \log \mathrm{E}\left[e^{s\Lambda}\right]$, $s \in \mathbb{R}$. Using the definition of Λ, we can rewrite $\kappa(s)$ as

$$\kappa(s) = \log \mathrm{E}_\Lambda \left[\left(\frac{\Pr(\hat{c}=1|y,h)}{\Pr(\hat{c}=0|y,h)}\right)^s\right], \tag{9}$$

where the subscript Λ indicates that the expectation is taken with respect to Λ. Since the equivocation variable is a function of the channel and modulation parameters, we can in fact perform the expectation over the joint distribution (6)

$$\kappa(s) = \log \mathrm{E}_{m,w,h,n} \left[\left(\frac{\Pr(\hat{c}=1|y,h)}{\Pr(\hat{c}=0|y,h)}\right)^s\right]. \tag{10}$$

It is well known that $\kappa(s)$ is a convex function of its argument [6], and its minimum is reached at a point $\hat{s}$, the saddlepoint, such that $\kappa'(\hat{s}) = 0$. From symmetry considerations in (10) it can be easily seen that $\hat{s} = 1/2$.

The Chernoff bound [7] sets an upper bound to the tail/error probability as

$$\Pr\left(\sum_{j=1}^{d} \Lambda_j > 0\right) \le e^{d\kappa(\hat{s})}. \tag{11}$$

As $\hat{s} = 1/2$, we have

$$\Pr\left(\sum_{j=1}^{d} \Lambda_j > 0\right) \le \left(\mathrm{E}_{m,w,h,n}\left[\sqrt{\frac{\Pr(\hat{c}=1|y,h)}{\Pr(\hat{c}=0|y,h)}}\right]\right)^d,$$

which is nothing but the Bhattacharyya union bound, derived by other means in [2] where

$$B = \mathrm{E}_\Lambda\left[e^{\frac{1}{2}\Lambda}\right] = \mathrm{E}_{m,w,h,n}\left[\sqrt{\frac{\Pr(\hat{c}=1|y,h)}{\Pr(\hat{c}=0|y,h)}}\right] \tag{12}$$

denotes the Bhattacharyya factor. Notice that B can be easily evaluated by numerical integration using the Gauss-Hermite (for the AWGN channel) and a combination of the Gauss-Hermite and Gauss-Laguerre (for the fading channel) quadrature rules, which are tabulated in [8].

3.4. Saddlepoint Approximation

A more accurate approximation to (7) for the error probability is a saddlepoint approximation. A simple version of can be found in [7] and is given by

$$\Pr\left(\sum_{j=1}^{d}\Lambda_j > 0\right) \simeq \frac{1}{\sqrt{2\pi d}\lambda}\, e^{d\kappa(\hat{s})}, \qquad (13)$$

where $\lambda = \sqrt{\kappa''(\hat{s})}\hat{s}$. The exponent is the same as for the Chernoff bound, in accordance with its asymptotic optimality. Note that the second derivative $\kappa''(\hat{s})$,

$$\kappa''(\hat{s}) = \frac{\mathrm{E}\left[\Lambda^2 e^{\hat{s}\Lambda}\right]}{\mathrm{E}\left[e^{\hat{s}\Lambda}\right]}$$
$$= \frac{1}{B}\,\mathrm{E}_{m,w,h,n}\left[\left(\log\frac{\Pr(\hat{c}=1|y,h)}{\Pr(\hat{c}=0|y,h)}\right)^2\sqrt{\frac{\Pr(\hat{c}=1|y,h)}{\Pr(\hat{c}=0|y,h)}}\right],$$

can again be efficiently computed using Gaussian quadrature rules.

Although of no importance for this case, as $\hat{s} = 1/2$, the saddlepoint approximation (13) fails for $\hat{s}$ in the vicinity of 0. Another saddlepoint approximation, called uniform due to its validity for the whole range of $\hat{s}$, is the Lugannani-Rice formula [6], given by

$$\Pr\left(\sum_{j=1}^{d}\Lambda_j > 0\right) \simeq Q\big(\sqrt{-2d\kappa(\hat{s})}\big) + \frac{1}{\sqrt{2\pi d}}e^{d\kappa(\hat{s})}\left(\frac{1}{\lambda}-\frac{1}{r}\right),$$

where $|r| = \sqrt{-2\kappa(\hat{s})}$. The sign of r is the sign of $\hat{s}$. Due to its wider validity, we shall use this approximation to the tail probability.

In [3] we used a similar approximation to this, where the only term was

$$Q\big(\sqrt{-2d\kappa(\hat{s})}\big) = Q\big(\sqrt{-2d\log B}\big). \qquad (14)$$

The key idea in [3] was to approximate the binary input BICM channel as a binary-input AWGN (BI-AWGN) with SNR $\ell = \kappa(\hat{s}) = -\log B$, the same approximation that links the exact formula and the Chernoff bound for the BI-AWGN channel [9]. This paper provides thus further justification to the accuracy of the results obtained in [3], and also suggests how to extend its validity with the general framework of saddlepoint approximations. In this line, Figure 2 also depicts the distributions of the LLR for a BI-AWGN with SNR $\ell = -\log B$, known to be $\sim \mathcal{N}(-4\ell, 8\ell)$. Remark the extreme closeness of the distributions in the tails.

4. NUMERICAL RESULTS

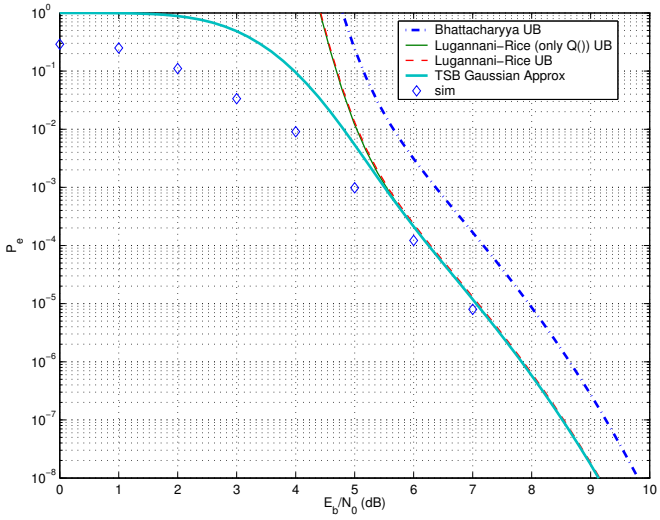
Figure 3(a) shows the performance of the rate-1/2, 64-state, convolutional code over 16-QAM with Gray mapping. We depict four approximations: the Chernoff (or Bhattacharyya) bound [2], the Lugannani-Rice formula, the $Q()$ terms in the Lugannani-Rice formula which correspond to the Gaussian approximation in [3], and the tangential-sphere bound based on the Gaussian approximation [3]. Note that the effect of the exponential term in the Lugannani-Rice formula is negligible which shows that the Gaussian approximation is very accurate, and that it is much tighter than the Bhattacharyya bound.

Figure 3(b) depicts the bit error probability of coded 8-PSK with Gray mapping, and an optimal 8-state rate-2/3 convolutional code. As before, we consider four cases, the Chernoff bound, the complete Lugannani-Rice formula, the Lugannani-Rice formula without exponential term, and the tangential-sphere bound using the Gaussian approximation. For large SNR, the behaviour is approximately linear in SNR (code diversity), as is typical of Rayleigh fading channels. In this case we observe a difference when using the complete Lugannani-Rice formula with the exp() function, which implies that in this case, the BICM equivalent channel is not Gaussian with parameter ℓ.

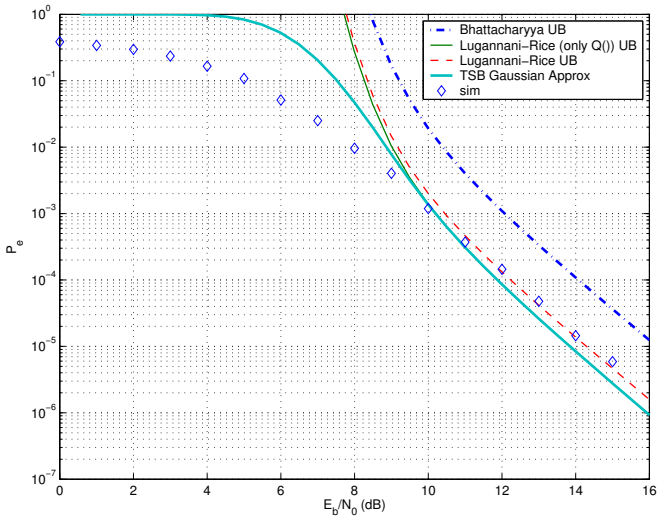
Similar comments apply to Figures 4(a) and 4(b), where we show the performance of the repeat-and-accumulate code ensemble [10] of rate 1/4 with 16-QAM over AWGN and Rayleigh fading channels; the overall spectral efficiency is 1 bps/Hz. In the case of the AWGN channels $K = 1024$ and in the fully-interleaved Rayleigh fading channel $K = 512$. For the sake of comparison, we also show the simulation with 20 iterations of belief-propagation decoding. The bounds seem to follow closely the error rates of the code, even though the error floor region is not reached. Again, the bounds with only the $Q()$ terms are very accurate for the AWGN channel; in the case of the Rayleigh fading channel, the complete bound is needed. Note also that, the tangential-sphere with the Gaussian approximation bound may also be somehow optimistic in the waterfall region. However, the Gaussian approximation still yields fairly accurate results.

5. CONCLUSIONS

In this paper, we have presented a very accurate and simple to compute approximation to the error probability of BICM using the saddlepoint approximation. We have verified the validity of the approximation for both, convolutional and turbo-like code ensembles with BICM, over AWGN and fully-interleaved Rayleigh fad-

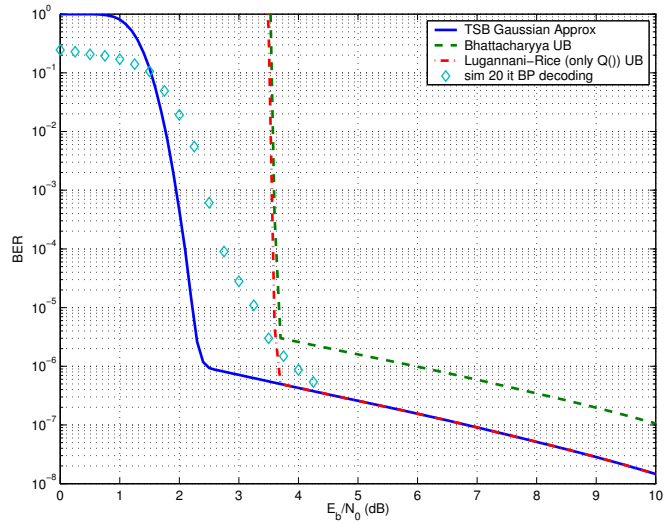


(a) 16-QAM Gray in AWGN (Convolutional Code with rate 1/2, 64 states, $K = 128$ information bits).

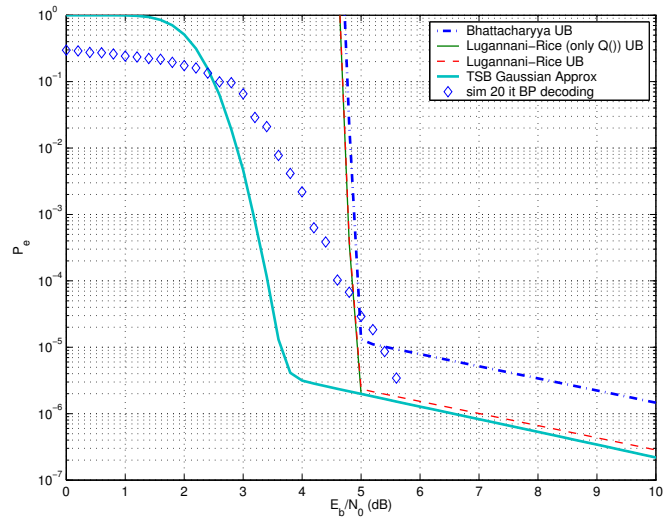


(b) 8-PSK Gray in fully-interleaved Rayleigh fading (Convolutional Code with rate 2/3, 8 states, $K = 128$ information bits).

Figure 3: Comparison of simulation and saddlepoint approximations on the bit error rate of BICM with convolutional codes in AWGN and fully-interleaved Rayleigh fading channels.



(a) 16-QAM Gray in AWGN (Repeat-and-Accumulate with rate 1/4, $K = 1024$ information bits and 20 iterations of belief propagation decoding).



(b) 16-QAM Gray in fully-interleaved Rayleigh Fading (Repeat-and-Accumulate with rate 1/4, $K = 512$ information bits and 20 iterations of belief propagation decoding).

Figure 4: Comparison of simulation and saddlepoint approximations on the bit error rate of BICM with repeat-and-accumulate in AWGN and fully-interleaved Rayleigh fading channels.

ing channels. The proposed method benefits from simple numerical integration using Gaussian quadratures for noise and fading averaging. This simple technique constitutes a powerful tool to the analysis of finite-length BICM; furthermore, being simpler and tighter than the original bounds in [2], it shows a wide range of practical applications.

A. DERIVATION OF THE SADDLEPOINT APPROXIMATION

The saddlepoint approximation exploits the link between the probability density and the cumulant transform via a Fourier-like transform. This means that we can freely move from one domain to the other and deal with the same random variable. As the cumulant transform is a convex function of its argument, it is in some sense easier to characterize than the density.

The cumulant transform of a random variable with density $f_x(x)$ is defined as $\kappa(s) = \log E(e^{sX})$, with $s \in \mathbf{C}$, and defined in a strip of the complex plane $\gamma_1 < \operatorname{Re} s < \gamma_2$ [6]. We will need the property that for a sum of independent random variables $\sum_{i=1}^{\ell} X_i$, the cumulant transform is the sum of the transforms for each variable; if they are identically distributed, we have $\ell \log E(e^{sX_1})$.

The tail probability $\int_x^{+\infty} f_x(t)\,dt$ can be calculated with the following inversion formula[1] as

$$\int_x^{+\infty} f_x(t)\,dt = \frac{1}{2\pi j}\int_C e^{\kappa(s)-sx}\frac{ds}{s}, \qquad (15)$$

which forms the starting point of the saddlepoint approximation. It is convenient to choose the contour is a straight line $(\hat{-}j\infty, \hat{s}+j\infty)$ passing through the saddlepoint, the value of s for which $\kappa'(\hat{s}) - x = 0$. Using a Taylor expansion around the maximum we obtain

$$\begin{aligned} g(s) &= g(\hat{s}) + g'(\hat{s})(s-\hat{s}) + \frac{1}{2}g''(\hat{s})(s-\hat{s})^2 + O\big((s-\hat{s}^3)\big) \\ &= g(\hat{s}) + \frac{1}{2}g''(\hat{s})(s-\hat{s})^2 + O\big((s-\hat{s}^3)\big). \qquad (16) \end{aligned}$$

Analogously, we expand $\frac{1}{s}$ around $\hat{s}$, and we obtain an expansion of (15) whose first term is given by

$$\frac{e^{\kappa(\hat{s})-\hat{s}x}}{2\pi j}\int_C e^{\frac{1}{2}\kappa''(\hat{s})(s-\hat{s})^2}\frac{ds}{\hat{s}}$$

Changing the contour of integration to pass through the saddlepoint, and with the change of variable $s = \hat{s} + js_i$ we can rewrite the term as

$$\frac{e^{\kappa(\hat{s})-\hat{s}x}}{2\pi}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\kappa''(\hat{s})s_i^2}\frac{ds_i}{\hat{s}} = \frac{e^{\kappa(\hat{s})-\hat{s}x}}{\sqrt{2\pi\kappa''(\hat{s})}\hat{s}}. \qquad (17)$$

[1] Here there is a minor technical assumption of the contour.

It is clear that this equation ceases to be valid when $\hat{s}$ is close to zero. This is due to the pole at $s = 0$ in (15), whose effect cannot be neglected when the saddlepoint is in its neighbourhood. The approximations valid also in this range are called uniform, and the formula by Lugannani&Rice is the simplest example of such an approximation.

References

[1] G. D. Forney, Jr., "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, no. 5, pp. 1241–1260, September 1991.

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

[3] A. Guillén i Fàbregas, A. Martinez, and G. Caire, "Error probability of bit-interleaved coded modulation using the gaussian approximation," in *Proceedings of the Conf. on Inform. Sciences and Systems (CISS-04)*, Princeton, USA, March 2004.

[4] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, February 2001.

[5] X. Li and J. A. Ritcey, "Trellis-coded modulation with bit interleaving and iterative decoding," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 715–724, April 1999.

[6] J. L. Jensen, *Saddlepoint Approximations*. Oxford, UK: Clarendon Press, 1995.

[7] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley and Sons, 1968.

[8] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables*, reprint ed. John Wiley & Sons Inc, 1993.

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

[10] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for "turbo-like" codes," in *Proceedings on the Thirty-Sixth Annual Allerton Conference on Communication, Control, and Computing*, Allerton House, Monticello (USA), September 1998, pp. 201–210.