

New Designs for Bit-Interleaved Coded Modulation with Hard-Decision Feedback Iterative Decoding

Alireza Kenarsari-Anhari, *Student Member, IEEE*, and Lutz Lampe, *Senior Member, IEEE*

Abstract—Bit-interleaved coded modulation (BICM) is often the method of choice for multilevel transmission over fading channels. If relatively simple forward error-correction (FEC) codes such as convolutional codes are employed, iterative decoding between demapper and FEC decoder can provide significant performance improvements over non-iterative decoding. To keep the complexity of iterative decoding low, the use of hard-decision feedback from FEC decoder to demapper is appealing. However, the price to be paid is a performance degradation due to feedback errors. In this letter, two new demapper designs are developed which are able to strongly mitigate the effect of erroneous feedback. The key ideas are (a) the use of an estimate of the average error rate for the hard-decision feedback and (b) the interpretation of feedback errors as additive impulsive noise. Simulation results show that the proposed designs achieve error rates close to those for iterative decoding with soft feedback, while they maintain the complexity advantage of using hard-decision feedback.

Index Terms—Bit-interleaved coded modulation, iterative decoding, hard-decision feedback.

I. INTRODUCTION

Bit-interleaved coded modulation (BICM) [1], [2] has become a popular coded modulation scheme for transmission over fading channels. BICM connects a binary forward error correction (FEC) encoder to a multilevel modulator and approaches the associated constellation-constrained capacity limit if a capacity-approaching binary code and Gray labeling are used [2]. The BICM structure can also be looked at as a concatenated coding system, with the FEC encoder and the multilevel modulator as outer and inner encoder, respectively. The inner encoder is made “stronger”, if non-Gray labeling is applied, cf. e.g. [3]–[5]. Interestingly, this BICM with non-Gray labeling can achieve excellent error-rate performance with relatively simple outer binary codes, such as for example convolutional codes.

BICM considered as concatenated code is commonly decoded in an iterative fashion, in which the demapper and a soft-input soft-output (SISO) channel decoder for the outer FEC code exchange extrinsic information. We will refer to this structure as BICM with soft-feedback iterative decoding (BICM-SID). An alternative decoder proposed in [6] uses a

soft-input hard-output (SIHO) outer decoder, like the Viterbi decoder for convolutional codes. This BICM with hard-decision feedback iterative decoding (BICM-HID) has two complexity advantages over BICM-SID. First, the outer SIHO decoder is less complex than its SISO counterpart, and second, the demapper using hard-decision feedback needs to consider only two instead of all constellation points for each labeling bit [3]. On the downside, BICM-HID is considerably outperformed by BICM-SID due to the effect of erroneous feedback.

In this letter, we propose two novel demapper designs for BICM-HID which mitigate the effect of feedback errors and thus improve the overall error-rate performance of BICM-HID. We focus on convolutionally coded transmission, for which SIHO FEC decoding, i.e., Viterbi decoding is commonly applied. The first key idea is that the demapper makes use of the error rate of the hard-decision feedback after iteration i , which we denote by P_i . As we will show, this essentially provides the demapper with reliability information and penalizes unreliable feedback. The corresponding demapper is similar to that proposed in [7]. However, different from the scheme in [7], which relies on a SISO outer decoder, our approach retains the outer Viterbi decoder and makes use of a recent analytical result [8] for the error rate performance of coded BICM to estimate P_i . Furthermore, parameter tuning as in [7, Eq.(10)] is unnecessary. The second key idea is the interpretation of the effect of feedback errors as additive impulsive noise, and its statistical approximation through a two-term Gaussian mixture probability density function (PDF). This leads us to the second proposed demapper, whose complexity is practically the same as that of the demapper used in conventional BICM-HID. We provide simulative evidence that BICM-HID with the proposed designs effectively bridges the error-rate gap between conventional BICM-HID and BICM-SID.

The remainder of this letter is organized as follows. In Section II, BICM with iterative decoding and demappers for conventional BICM-HID and BICM-SID are briefly reviewed. In Section III, we derive the two new demapper designs. Numerical results are presented in Section IV. In Section V, concluding remarks are offered.

II. PRELIMINARIES

We consider a convolutionally coded BICM system, whose block diagram is shown in Figure 1. The bit-interleaved output of the FEC encoder is input to a subsequent mapper $\mu : \{0, 1\}^r \rightarrow \mathcal{X}$ assigning r coded bits $[c_1, \dots, c_r]$ to a signal point $x \in \mathcal{X}$. The signal x is transmitted over a flat fading additive white Gaussian noise (AWGN) channel, and

Manuscript received July 25, 2009; revised November 13, 2009. This work was supported by the National Sciences and Engineering Research Council (NSERC) of Canada.

A. Kenarsari-Anhari was with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada. He is now with Dyaptive Systems, Vancouver, BC, Canada. L. Lampe is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada. (e-mail: a.kenarsari@gmail.com, Lampe@ece.ubc.ca).

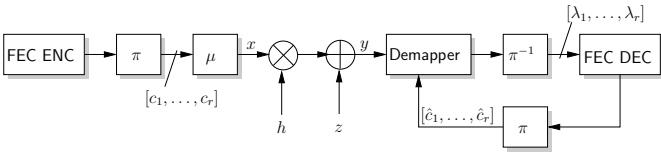


Fig. 1. Block diagram of BICM transmission over a fading channel and iterative decoding. π and π^{-1} denote interleaving and deinterleaving, respectively. r coded binary symbols $[c_1, \dots, c_r]$ are mapped to one transmitted symbol x . The feedback from the FEC decoder (FEC DEC) to the demapper is shown the form of decoded code symbols $[\hat{c}_1, \dots, \hat{c}_r]$, i.e., hard-decision feedback.

assuming a coherent receiver with perfect compensation of the channel phase, the corresponding received sample $y \in \mathbb{C}$ can be written as

$$y = hx + z, \quad (1)$$

where $h \in \mathbb{R}^+$ denotes the channel gain and $z \in \mathbb{C}$ is the AWGN sample. Without loss of generality, we assume $\mathbf{E}\{|x|^2\} = 1$, $\mathbf{E}\{h^2\} = \gamma$, and $\mathbf{E}\{|z|^2\} = 1$, and thus γ is the average signal-to-noise ratio (SNR).

The receiver applies a concatenated demapper-decoder structure (see Figure 1), where the demapper generates r bit-wise decoding metrics

$$\lambda_j = \log \left[\sum_{a \in \mathcal{X}_{j,1}} \exp(f(a, j) - |y - ha|^2) \right] - \log \left[\sum_{a \in \mathcal{X}_{j,0}} \exp(f(a, j) - |y - ha|^2) \right], \quad (2)$$

$1 \leq j \leq r$, for each transmitted symbol x . In the above expression, $\mathcal{X}_{j,b}$ denotes the set of symbols in the constellation with the j th binary label fixed to b , and $f(a, j)$ represents a-priori information or a *bias* provided by the FEC decoder. Of course, $f(a, j) = 0$ for the first iteration, in which no feedback from the decoder is available.

In the case of a BICM-SID with soft feedback $\Pr(c_j = b)$, $1 \leq j \leq r$, we can write the bias-term as

$$f(a, j) = \sum_{\substack{\ell=1 \\ \ell \neq j}}^r \log(\Pr(c_\ell = b_\ell)), \quad (3)$$

where $[b_1, \dots, b_r] = \mu^{-1}(a)$. For BICM-HID with hard-decision feedback $\hat{c}_j \in \{0, 1\}$, $1 \leq j \leq r$, we have

$$f(a, j) = \begin{cases} 0, & \text{if } b_\ell = \hat{c}_\ell, \forall \ell \in \{1, \dots, r\}, \ell \neq j, \\ -\infty, & \text{otherwise} \end{cases}. \quad (4)$$

We observe that BICM-HID has two complexity advantages over BICM-SID. First, a SIHO decoder, typically the Viterbi decoder, can be used instead of a more complex SISO decoder. Second, the summation in (2) and thus the 2^r -times evaluation of the Euclidean distance can be omitted and the computation of the bias term $f(a, j)$ is greatly simplified.

III. NEW BICM-HID SCHEME

We now present two new demapper designs for BICM-HID. The additional information that is used by the proposed

demappers is an estimate of the average bit-error rate (BER) for the coded bits c_j after the i th iteration, P_i . The first proposed demapper has the same computational complexity as the demapper proposed in [7], but does not require outer SISO decoding and parameter tuning. The second demapper design enjoys a very similar complexity advantage over BICM-SID as the conventional BICM-HID, but achieves a greatly improved error-rate performance.

For the moment, let us assume that P_i is known perfectly at the demapper.

A. Demapper Design I

Having available only the hard decision \hat{c}_j and knowledge of P_i , the probability $\Pr(c_j = b_j)$ is given by $P_i^{d_j}(1 - P_i)^{1 - d_j}$, where $d_j = d_H(\hat{c}_j, b_j)$ and $d_H(\cdot, \cdot)$ returns the Hamming distance between its arguments. Hence, the bias $f(a, j)$ from (3) is replaced by

$$\begin{aligned} f(a, j) &= \sum_{\substack{\ell=1 \\ \ell \neq j}}^r \log(P_i^{d_\ell}(1 - P_i)^{1 - d_\ell}) \\ &= \sum_{\substack{\ell=1 \\ \ell \neq j}}^r \left[d_\ell \log\left(\frac{P_i}{1 - P_i}\right) + \log(1 - P_i) \right] \\ &= \beta_i \left(\sum_{\substack{\ell=1 \\ \ell \neq j}}^r d_\ell \right) - \alpha_i \\ &= \beta_i d_H(\hat{\mathbf{c}}_{j,b}, \mu^{-1}(a)) - \alpha_i, \end{aligned} \quad (5)$$

where we defined

$$\alpha_i \triangleq -(r - 1) \log(1 - P_i) \quad (6)$$

$$\beta_i \triangleq \log\left(\frac{P_i}{1 - P_i}\right), \quad (7)$$

$$\hat{\mathbf{c}}_{j,b} \triangleq [\hat{c}_1, \dots, \hat{c}_{j-1}, b, \hat{c}_{j+1}, \dots, \hat{c}_r]. \quad (8)$$

Substituting the bias from (5) in the metric (2) leads to

$$\lambda_j = \log \left[\sum_{a \in \mathcal{X}_{j,1}} \exp(\beta_i d_H(\hat{\mathbf{c}}_{j,1}, \mu^{-1}(a)) - |y - ha|^2) \right] - \log \left[\sum_{a \in \mathcal{X}_{j,0}} \exp(\beta_i d_H(\hat{\mathbf{c}}_{j,0}, \mu^{-1}(a)) - |y - ha|^2) \right], \quad (9)$$

which is our first new demapper design and referred to as “design I”. Note that the constant term α_i cancels out in (9).

In passing, we remark that an efficient implementation of the log-sum of exponentials or the max-log approximation [9] can equally be applied to (2) and (9), if the exp-operation is to be avoided.

B. Demapper Design II

To derive the second demapper design (“design II”), we first provide an interpretation of feedback errors $\hat{c}_j \neq c_j$ as impulsive noise.

1) *Gaussian Mixture Noise Interpretation:* The new metric (9) allows the interpretation of BICM-HID as transmission over a channel affected by additive Gaussian mixture noise [10]. More specifically, the bit-wise metric (9) is also obtained when considering transmission of $\mu(\hat{\mathbf{c}}_{j,b})$, $b \in \{0, 1\}$, over a channel with the effective noise

$$z_\nu^e = z + z_\nu^f, \quad (10)$$

where z_ν^f is due to the hard-decision feedback and thus depends on bit position j , feedback $\hat{\mathbf{c}}_{j,b}$, and fading gain h , which are collected in the parameter vector $\nu \triangleq [j, \hat{\mathbf{c}}_{j,b}, h]$. The feedback noise has a probability mass function $p_{z_\nu^f}(m)$ with mass

$$p(a, \hat{\mathbf{c}}_{j,b}) \triangleq P_i^{d_H(\hat{\mathbf{c}}_{j,b}, \mu^{-1}(a))} (1 - P_i)^{[r-1-d_H(\hat{\mathbf{c}}_{j,b}, \mu^{-1}(a))]} \quad (11)$$

at location $m = h(\mu(\hat{\mathbf{c}}_{j,b}) - a)$, $a \in \mathcal{X}_{j,b}$. Hence, the PDF of z_ν^e is the Gaussian mixture density

$$p_{z_\nu^e}(m) = \frac{1}{\pi} \sum_{a \in \mathcal{X}_{j,b}} p(a, \hat{\mathbf{c}}_{j,b}) \exp\left(-|m - [h(\mu(\hat{\mathbf{c}}_{j,b}) - a)]|^2\right). \quad (12)$$

2) *Simplified Demapper:* For the second, simplified demapper, we propose the approximation of (12) by a single two-term Gaussian mixture function for all j , $\hat{\mathbf{c}}_{j,b}$, and h . That is, we apply the two-term PDF

$$\tilde{p}_{z^e}(m) = \frac{(1 - P_i)^{r-1}}{\pi} \exp(-|m|^2) + \frac{1 - (1 - P_i)^{r-1}}{\pi \sigma^2} \exp\left(-\frac{|m|^2}{\sigma^2}\right), \quad (13)$$

where the first term indicates the noise when there is no error in the feedback from the FEC decoder, an event which occurs with probability $(1 - P_i)^{r-1}$, and the second term represents the equivalent noise in the presence of feedback errors, which occur with probability $1 - (1 - P_i)^{r-1}$. We obtain the variance σ^2 for the second term in (13) by making sure that the average variance according to the exact PDF in (12) is maintained. Hence, we have $(p_h(\eta))$ denotes the PDF of the channel gain h)

$$\begin{aligned} & (1 - P_i)^{r-1} + (1 - (1 - P_i)^{r-1})\sigma^2 \\ &= \frac{1}{2^r r} \sum_{\substack{j \in \{1, \dots, r\} \\ b \in \{0, 1\} \\ \mathbf{e}_{j,b} \in \{0, 1\}^{(r-1)}}} \int_0^\infty p_h(\eta) \int_{m \in \mathcal{C}} |m|^2 \\ & \quad \times \frac{1}{\pi} \sum_{a \in \mathcal{X}_{j,b}} p(a, \hat{\mathbf{c}}_{j,b}) \exp\left(-|m - [\eta(\mu(\hat{\mathbf{c}}_{j,b}) - a)]|^2\right) dm d\eta \\ &= 1 + \frac{1}{2^r r} \sum_{\substack{j \in \{1, \dots, r\} \\ b \in \{0, 1\} \\ \mathbf{e}_{j,b} \in \{0, 1\}^{(r-1)}}} \left[\gamma \sum_{a \in \mathcal{X}_{j,b}} p(a, \hat{\mathbf{c}}_{j,b}) |\mu(\hat{\mathbf{c}}_{j,b}) - a|^2 \right], \end{aligned} \quad (14)$$

from which we obtain σ^2 .

Using the noise PDF given in (13), defining

$$\delta_i \triangleq \log\left(\frac{\sigma^2}{1 - (1 - P_i)^{r-1}}\right), \quad (15)$$

and applying the max-log approximation [9], we can re-write the bit-metric (2) as

$$\lambda_j = -\min \left\{ \alpha_i + |y - h\mu(\hat{\mathbf{c}}_{j,1})|^2, \delta_i + \frac{|y - h\mu(\hat{\mathbf{c}}_{j,1})|^2}{\sigma^2} \right\} + \min \left\{ \alpha_i + |y - h\mu(\hat{\mathbf{c}}_{j,0})|^2, \delta_i + \frac{|y - h\mu(\hat{\mathbf{c}}_{j,0})|^2}{\sigma^2} \right\}. \quad (16)$$

C. Comments on Complexity

As mentioned in Section I, BICM-HID has two distinct complexity advantages over BICM-SID. The proposed demappers work with a SIHO decoder and thus both make use of the first complexity advantage of SIHO versus SISO decoding, which amounts to a gain by about a factor of three [11, Appendix F2]. Demapper II also realizes the second complexity advantage in that symbol metrics for only two signal points, instead of all signal points in the case of BICM-SID, need to be computed per bit metric. Thus, different from BICM-SID, the complexity of Demapper II is independent of the constellation size. While the corresponding gain factor is implementation dependent, we note that (16) requires only two additional comparisons relative to conventional BICM-HID using (4), and thus further complexity savings are difficult to obtain. We note that σ^2 in (14) only depends on the signal constellation \mathcal{X} , the SNR γ , and the BER P_i , and thus it is computed once per iteration and the additional number of computations per bit are negligible.

D. Low-Complexity Estimation of P_i

We now consider the estimation of P_i , which is required for the new demapper designs (9) and (16). It should be noted that the overall error-rate performance is relatively robust with respect to the estimation accuracy, and thus we are content with a simple method that provides a coarse estimate for P_i .

For the error rate P_1 after the first iteration, we consider the dominant error events of the Viterbi decoder and thus use the estimate

$$\tilde{P}_1 = \frac{1}{n_c} w_{d_{\text{free}}} \text{PEP}(d_{\text{free}}), \quad (17)$$

where $\text{PEP}(d_{\text{free}})$ is the pairwise error probability between two codewords with Hamming distance d_{free} , d_{free} denotes the free distance of the convolutional code of rate $R_c = k_c/n_c$, and $w_{d_{\text{free}}}$ is the total weight of *output* bits in error events with Hamming weight d_{free} . Using the saddlepoint approximation [12, Eq.(12)], closed-form expressions for this PEP using elementary functions have been provided in [8] for the rich class of Nakagami- m fading channels and arbitrary parameter $m \geq 1/2$, and thus (17) can be easily evaluated. Next, for the BER P_n after the final iteration $i = n$, we use the perfect-feedback lower bound from [3] together with the saddlepoint approximation, which again results in a closed-form PEP expression and thus estimate \tilde{P}_n . Finally, we estimate the BER for iteration i using the interpolation

$$\log(\tilde{P}_i) = \frac{n-i}{n-1} \log(\tilde{P}_1) + \frac{i-1}{n-1} \log(\tilde{P}_n). \quad (18)$$

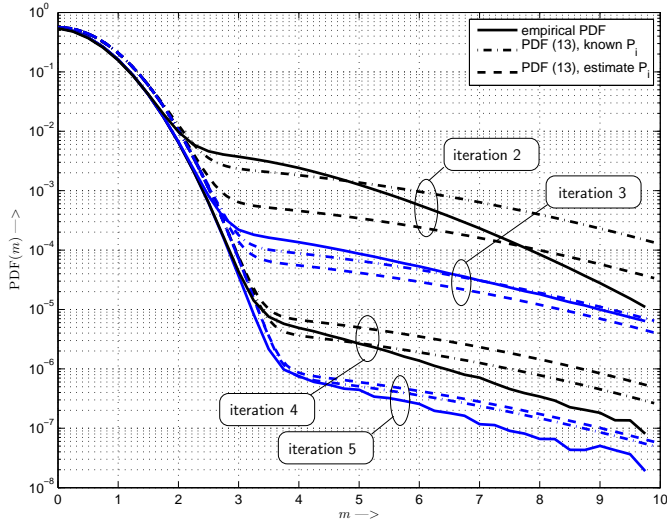


Fig. 2. Two-term PDF approximation from (13) and empirical PDF from simulations during the $(i + 1)$ st iteration, $i = 1, \dots, 4$, at SNR $\gamma_b = 10$ dB. The PDF approximation is shown for the cases of known P_i and using the estimate \tilde{P}_i from (18). A total of $n = 5$ iterations is assumed.

We remark that (18) is a pragmatic choice that has been proven adequate in numerical experiments (see also the results in the next section). The value of n should be chosen such that BICM-SID converges to the perfect-feedback lower bound.

IV. SIMULATION RESULTS

In this section, we present selected simulation results to illustrate the performance of BICM-HID with the proposed demappers. We use the example of 16-ary quadrature amplitude modulation (QAM) BICM with modified set-partitioning labeling over a Rayleigh fading channel, cf. [3], and employ the maximum-free distance 4-state rate-1/2 convolutional code (very similar results have been obtained for codes with larger memory). BER and frame-error rate (FER) results are shown as function of the average bit-wise SNR $\gamma_b = \gamma/(R_c r) = \gamma/2$.

First, we take a look at the goodness of the proposed PDF approximation (13). For this purpose, Figure 2 shows the analytical two-term PDF $\tilde{p}_{z^e}(m)$ and the empirical PDF obtained from Monte Carlo simulation for an SNR of $\gamma_b = 10$ dB. The two-term PDF is plotted for known P_i and using the estimate \tilde{P}_i from (18), respectively. A total of $n = 5$ iterations is assumed, so four sets of curves are plotted in Figure 2. The interleaver length is set to 10000 binary symbols. It can be seen that the empirical PDF displays a markedly non-Gaussian shape, which is a result of erroneous feedback and explains the relatively poor performance of conventional BICM-HID that is based on the Euclidean distance metric (see also results below). The two-term approximation seems to mimic the non-Gaussian behaviour fairly well, which corroborates the proposed demapper design approach. The difference between P_i and its estimate \tilde{P}_i reflects mainly in a vertical shift of the second term of the mixture PDF. This can be also seen from (13) when approximating $(1 - P_i)^{r-1}$ by $(1 - rP_i)$, which is valid for small P_i .

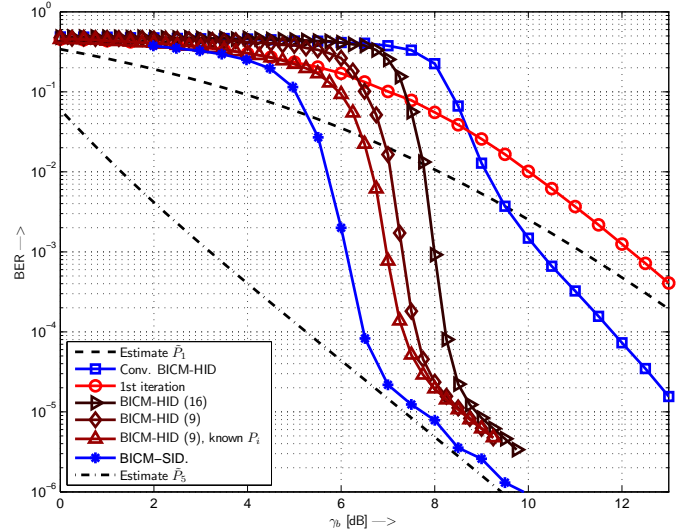


Fig. 3. BER for BICM-HID with conventional and proposed demappers. As reference, BER for BICM-SID, and BER estimates \tilde{P}_1 and \tilde{P}_5 are also included.

Figure 3 presents the simulated BER¹ after the first and fifth iteration for BICM-HID with (a) the conventional demapper ((2) with bias (4)), (b) demapper design I (9), and (c) demapper design II (16). For design I we also include the case of perfect BER estimation, i.e., $\tilde{P}_i = P_i$. Furthermore, as a reference, the curves for BICM-SID and the estimated BERs \tilde{P}_1 and $\tilde{P}_n = \tilde{P}_5$ after the first and fifth iteration are shown (see Footnote 1). The total number of iterations and interleaver length are the same as for the results in Figure 2.

We observe that the proposed demapper designs significantly improve the BER performance relative to the conventional demapper for BICM-HID. In particular, the considerable gap between conventional BICM-HID and BICM-SID is largely closed through the modified demappers. Demapper design II achieves a performance close to that of demapper design I, which renders it the perhaps preferred choice considering its low computational complexity. It can also be seen that the feedback-BER approximation \tilde{P}_i of P_i is sufficiently accurate for the purposes of BICM-HID, since the two BER curves for estimated and known P_i almost coincide. The quality of the estimate \tilde{P}_i is also confirmed by the relatively good approximation of the actually BER after one and five iteration through \tilde{P}_1 and \tilde{P}_5 , respectively.

Figure 4 shows the FER for the same scenario as above, but a relatively short interleaver of length 2000. This allows a comparison with the results from [7, Fig. 3]. We observe that the proposed designs achieve a performance very close to that from [7], despite the use of a SIHO decoder (designs I and II) and a less complex demapper (design II). The small consistent gap between design I and the demapper from [7] could likely be removed through scaling of β_i in (9), which however requires a simulation-based tuning of the scaling parameter as was done in [7]. Since design II performs close

¹As usual, we plot the BER for binary *information* symbols. To enable a direct comparison, the shown analytical results for \tilde{P}_1 and \tilde{P}_5 are also determined with respect to information symbols.

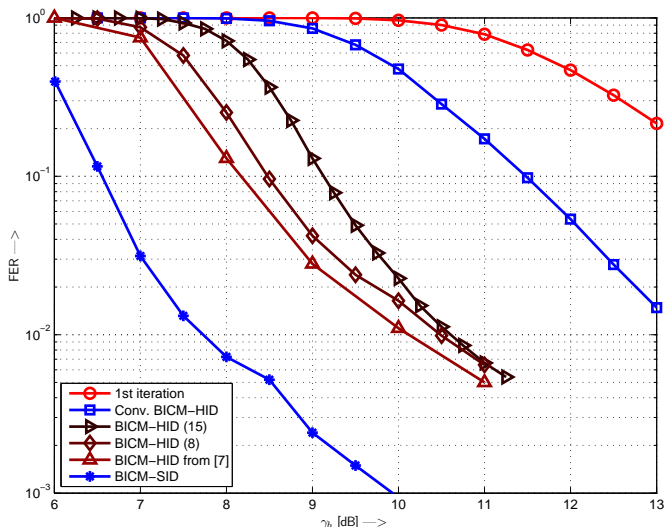


Fig. 4. FER for BICM-HID with conventional and proposed demappers, and demapper from [7]. As reference, FER for BICM-SID.

to design I also in terms of FER and for this relatively short interleaver length, it is an attractive choice for low-complexity BICM-HID.

V. CONCLUSION

In this letter, we have presented two novel demapper designs for BICM-HID. The key ideas are the use of a simple approximation of the feedback error-rate and the interpretation of feedback errors as additive impulsive noise. The first demapper design makes optimal use of error-rate information, while the second design is suboptimal but enjoys the low complexity of conventional BICM-HID. Our simulation results have shown that the proposed schemes significantly outperform conventional BICM-HID and approach the performance of BICM-SID, which is the ultimate performance limit for BICM with iterative decoding.

REFERENCES

- [1] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873–884, May 1992.
- [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. Chindapol and J. Ritcey, "Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 5, pp. 944–957, May 2001.
- [4] J. Tan and G. Stüber, "Analysis and design of symbol mappers for iteratively decoded BICM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 662–672, Mar. 2005.
- [5] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 7, no. 12, pp. 593–595, Dec. 2003.
- [6] X. Li and J. Ritcey, "Bit-interleaved coded modulation with iterative decoding," in *IEEE International Conference on Communications (ICC)*, vol. 2, Vancouver, BC, Canada, Jun. 1999, pp. 858–863.
- [7] T. Li, W. Mow, and K. Letaief, "Low complexity iterative decoding for bit-interleaved coded modulation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 1966–1970, Aug. 2006.
- [8] A. Kenarsari-Anhari and L. Lampe, "An analytical approach for performance evaluation of BICM transmission over Nakagami-m fading channels," *IEEE Trans. Commun.*, 2010.

- [9] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding," *European Transactions on Telecommunication*, vol. 2, no. 8, pp. 119–125, Mar./Apr. 1997.
- [10] D. Stein, "Detection of Random Signals in Gaussian Mixture Noise," *IEEE Trans. Inform. Theory*, vol. 41, no. 6, pp. 1788–1801, Nov. 1995.
- [11] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: with Wireless Applications*. New York: Kluwer Academic / Plenum Publishers, 1999.
- [12] A. Martinez, A. Guillén i Fabrègas, and G. Caire, "Error probability analysis of bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 262–271, Jan. 2006.