IMPROVED EMBEDDED PARAMETER ESTIMATION FOR BIT-INTERLEAVED CODED MODULATION WITH ITERATIVE DECODING

Frederik Simoens and Marc Moeneclaey

Department of Telecommunications and Information Processing
Ghent University, St.-Pietersnieuwstraat 41, B-9000 GENT, Belgium
E-mail: {fsimoens,mm}@telin.UGent.be

ABSTRACT

In this contribution, code-aided parameter estimation is considered for bit-interleaved coded modulation with iterative decoding (BICM-ID). The computational overhead caused by the estimation is minimized by embedding the parameter estimation into the iterative decoding. As will be apparent, the performance of the estimation depends on the availability of accurate *symbol* a posteriori probabilities (SAPP's). Since the decoder computes bit a posteriori probabilities (BAPP's), the SAPP's are traditionally computed through a straightforward bit-to-symbol probability conversion. For advanced mappings which are optimized for BICM-ID, however, the latter method does not result in reliable SAPP's. Therefore a new method is proposed and simulation results verify that the resulting iterative estimator-detector yields a close to perfectly synchronized performance.

1. INTRODUCTION

In Ungerboeck's trellis-coded modulation (TCM) [1], convolutional codes are designed to match the specific modulation scheme. On some channels, however, combining modulation and encoding does not yield the best performance. BICM [2], increases the time diversity of coded modulation compared with TCM, improving performance over fast-fading channels. This idea was later extended to BICM with iterative decoding (BICM-ID), whereby the decoder and the demapper exchange information in order to improve the performance [3]. When perfect channel state information is available, BICM-ID experiences large coding gains and is able to outperform TCM on any channel. The choice of mapping (i.e., how a group of bits is mapped onto a complex constellation symbol) is recognized to be a crucial design parameter for BICM-ID [4, 5].

Despite the potential high performance, BICM-ID schemes turn out to be very sensitive to parameter errors [6]. Estimation based on a limited number of pilot symbols does not provide sufficiently accurate estimates in order to exploit the potential coding gain [6]. Furthermore, existing iterative code-aided estimation schemes ([7, 8]) are so far unsuccessful for BICM-ID when using advanced optimized mappings. These estimation schemes, which are based on the Expectation Maximization (EM) algorithm, require the a posteriori probabilities of the unknown transmitted symbols and the latter are not directly obtained from the decoder. The conventional bit-to-symbol probability conversion (applied in [7, 8]), which computes the SAPP's as the product of the corresponding bit a posteriori probabilities (BAPP), is not very accurate when using advanced mapping schemes. In this contribution, we propose a new method to compute the SAPP's. The improved performance of the estimation scheme is illustrated through computer simulations.

2. SYSTEM MODEL

Let us consider the BICM-ID system depicted in Fig. 1. A sequence of information bits is encoded by a rate R convolutional code and bit-interleaved by a random interleaver, resulting in a sequence of L coded bits, $(c_1, \ldots, c_L)$. The coded bits are grouped in K blocks of m bits; the k-th block is denoted $\mathbf{c}_k = (c_k[1], c_k[2], \ldots, c_k[m])$. Then, $\mathbf{c}_k$ is mapped to a symbol a_k having elements in the M-ary signal set Ω, using a bijective mapping function $\mathcal{M} : \{0,1\}^m \rightarrow \Omega$, where $M = 2^m$. We write

$$a_k = \mathcal{M}(\mathbf{c}_k)\,. \tag{1}$$

We assume that the symbols a_k have a variance equal to $E_s = R.m.E_b$, where E_s and E_b denote the symbol energy and energy per information bit respectively.

Let us further consider a general frequency-flat channel. For sake of simplicity, we consider no timing or frequency offset, but the derivation including these parameters is very straightforward. The discrete time baseband received signal $\mathbf{y} = [y_1, \ldots, y_K]$ can be written as:

$$y_k = A\, a_k e^{j\theta} + n_k \tag{2}$$

where n_k is a complex-valued AWGN process with variance N_0. The channel parameters A and θ represent the ampli-

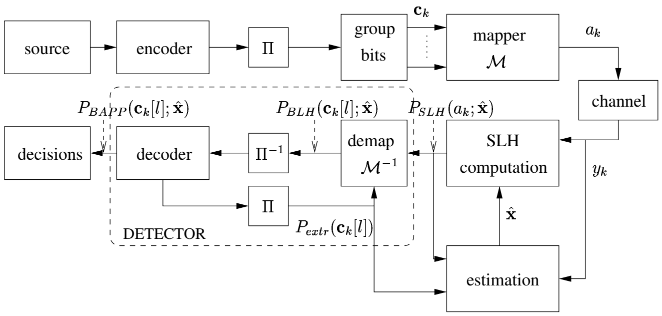

Fig. 1. BICM-ID system with embedded estimation.

tude and carrier phase offset respectively. We denote the vector of parameters $\mathbf{x} = [A, \theta]$.

The operation of the iterative detector is outlined in the next section.

3. DETECTOR OPERATION

The detector consists of two main blocks: a demapper and a decoder. First the symbol likelihoods (SLH) $P_{SLH}(a_k; \mathbf{x})$ of the received signal are computed for all possible a_k. These are easily obtained for a channel model conform with (2)

$$\begin{aligned} P_{SLH}(a_k;\mathbf{x}) &= p\left(y_k \left| \mathbf{c}_k = \mathcal{M}^{-1}(a_k), \mathbf{x}\right.\right) \\ &= C\, e^{-\frac{1}{N_0}\left\| y_k - A a_k e^{j\theta}\right\|^2}. \qquad (3) \end{aligned}$$

The demapper and decoder operate according to the turbo-principle [3]. The bit likelihoods (BLH) $P_{BLH}(c_k[l]; \mathbf{x})$ computed in the demapper (and to be forwarded to the decoder) are given by

$$\begin{aligned} & P_{BLH}(c_k[l] = b; \mathbf{x}) \\ & = p\left(y_k \left| c_k[l] = b, \mathbf{x}\right.\right) \\ & = C \sum_{\mathbf{c}_k:\, c_k[l]=b} p\left(y_k \left| \mathbf{c}_k, \mathbf{x}\right.\right) p\left(c_k[1],..,c_k[l{-}1], c_k[l{+}1],..,c_k[m]\right) \\ & = C \sum_{a_k \in \chi_b^l} P_{SLH}(a_k;\mathbf{x}) \prod_{l' \neq l} P_{extr}(c_k[l']) \qquad (4) \end{aligned}$$

where C is a normalizing constant and χ_b^l denotes the subset of all symbols a_k for which the l-th bit of the inverse mapping $\mathcal{M}^{-1}$ equals b (with $b \in \{0,1\}$); the product in (4) is over the remaining bits of $\mathcal{M}^{-1}(a_k)$. The extrinsic probabilities $P_{extr}(c_k[l])$ are fed back from the decoder (6). Note that we assumed in (4) that the coded bits are independent of the channel parameters $p(c_k[1], \ldots, [m] \,|\mathbf{x}) = p(c_k[1], \ldots, [m])$ and that bits, mapped to the same symbol, are independent because of the interleaver. The bit likelihoods $P_{BLH}(c_k[l])$ computed in the demapper are deinterleaved and applied as a priori information to the MAP decoder. The decoder, operating according to the BCJR algorithm [9], calculates the bit a posteriori probabilities (BAPP) $P_{BAPP}(c_k[l]) = p\left(c_k[l] \left| \mathbf{y}, \mathbf{x}\right.\right)$. According to the turbo principle, these can be factorized up to an irrelevant constant C in an a priori and extrinsic part:

$$\begin{aligned} P_{BAPP}(c_k[l]; \mathbf{b}) &= C.p\left(y_k \left| c_k[l], \mathbf{x}\right.\right) p\left(c_k[l]\right) \\ &= C.P_{BLH}(c_k[l];\mathbf{x})\, P_{extr}(c_k[l]) \qquad (5) \end{aligned}$$

again assuming $p\left(c_k[l] \left| \mathbf{x}\right.\right) = p\left(c_k[l]\right)$. From (5), we see that the extrinsic probabilities of the code bits can be computed as follows

$$P_{extr}(c_k[l]) = C \frac{P_{BAPP}(c_k[l];\mathbf{x})}{P_{BLH}(c_k[l];\mathbf{x})}, \qquad (6)$$

which means that $P_{extr}(c_k[l])$ corresponds to the decoder output-input ratio. The extrinsic probabilities are then interleaved and fed back to the demapper (4). After convergence of the iterative algorithm, decisions are made based on the BAPP's (5). Note that during the initial demapping step, we set the extrinsic probabilities equal to $P_{extr}(c_k[l]) = 1/2$. The system's operation is outlined in the probability-domain (for clarity), however, for reasons concerning complexity and stability, the implementation is performed in the log-domain.

4. EMBEDDED ESTIMATION

4.1. Parameter estimation

In this section, we propose a joint estimation and detection scheme for the iterative detector outlined above. The idea is to embed the parameter estimation into the iterative detection in order to minimize the computational overhead caused by the estimation. This means that the parameter estimates should be updated at the end of each detection iteration, using the information available at that time. For this problem, we resort to the iterative EM algorithm [10]. The problem has been addressed for turbo-coded systems with *Gray-mapped* 8-PSK signaling in [8] and *Gray-mapped* 16-QAM signaling in [7]. As will be apparent, the estimation for our BICM-ID set-up with a general signaling is resemblant to these schemes, however calculation of the a posteriori symbol probabilities requires extra attention.

Assume we have somehow obtained parameter estimates $\hat{\mathbf{x}}^{(i-1)}$ (from a previous iteration). The EM algorithm yields the following parameter update equations (see [8] for more details):

$$\hat{\theta}^{(i)} = \arg\left\{\sum \overline{a_k}^* y_k\right\} \qquad (7)$$

$$\hat{A}^{(i)} = \frac{\left|\sum_{k=0}^{K-1} \overline{a_k}^* y_k\right|}{\left(\sum_{k=0}^{K-1} \overline{a_k^* a_k}\right)} \qquad (8)$$

where the first and second order moments are computed as

$$\overline{a_k} = \sum_{\alpha_k \in \Omega} \alpha_k P_{SAPP}\left(\alpha_k; \hat{\mathbf{x}}^{(i-1)}\right)$$

$$\overline{a_k^* a_k} = \sum_{\alpha_k \in \Omega} \alpha_k^* \alpha_k P_{SAPP}\left(\alpha_k; \hat{\mathbf{x}}^{(i-1)}\right)$$

and where $P_{SAPP}(\alpha_k; \hat{\mathbf{x}}^{(i-1)}) = p\left(a_k = \alpha_k \left| \mathbf{y}, \hat{\mathbf{x}}^{(i-1)}\right.\right)$ denotes the SAPP conditioned on the parameter estimates $\hat{\mathbf{x}}^{(i-1)}$. Note that the iterative detector does not compute these *symbol* probabilities. In section 4.2, we will elaborate on how these should be obtained.

At the end of each detection stage, the parameters are updated according to ((7) and (8)). This yields little extra computation since the estimation process is now embedded in the detection scheme and the computation of (7) and (8) is minimal compared to the BCJR decoding operation. Initial parameter estimates $\hat{\mathbf{x}}^{(0)}$ are obtained by means of a short training sequence.

4.2. Symbol a posteriori probabilities

In the previous section, we pointed out that the estimation algorithm requires symbol a posteriori probabilities to compute the first and second order moment of the unknown transmitted symbols. The detector, however, provides us with a posteriori probabilities of the coded *bits*. Having these BAPP's, the SAPP's are traditionally computed based on a *bit-to-symbol* probability conversion [7, 8]:

$$\begin{aligned} P_{SAPP}^{(1)}(a_k; \mathbf{x}) &= p\left(\mathbf{c}_k \left| \mathbf{y}, \mathbf{x}\right.\right) \\ &\cong \prod_{l=1}^{m} p\left(c_k[l] \left| \mathbf{y}, \mathbf{x}\right.\right) \\ &= \prod_{l=1}^{m} P_{BAPP}\left(c_k[l]; \mathbf{x}\right) \end{aligned} \tag{9}$$

One argues that because of the interleaver, the bits are independent and the joint probability $p\left(\mathbf{c}_k \left| \mathbf{y}, \mathbf{x}\right.\right)$ can be factorized. However, this approach is not completely correct, because of the conditioning on the observation $\mathbf{y}$. Even if the bits belonging to the same symbol are a priori independent (e.g. uncoded), they will in general not be independent when conditioned on $\mathbf{y}$. Therefore we present a new method avoiding the conditioning on $\mathbf{y}$:

$$\begin{aligned} P_{SAPP}^{(2)}(a_k; \mathbf{x}) &= p\left(\mathbf{c}_k \left| \mathbf{y}, \mathbf{x}\right.\right) \\ &= C.p\left(\mathbf{y} \left| \mathbf{c}_k, \mathbf{x}\right.\right) p\left(\mathbf{c}_k\right) \\ &\cong C.P_{SLH}\left(a_k; \mathbf{x}\right) \prod_{l=1}^{m} P_{extr}\left(c_k[l]\right) \end{aligned} \tag{10}$$

Again, we assumed that bits belonging to the same symbol are independent, but in contrast with (9), we do not assume independence conditioned on the observation. Consequently, soft-symbol computation according to (10) should be more accurate.

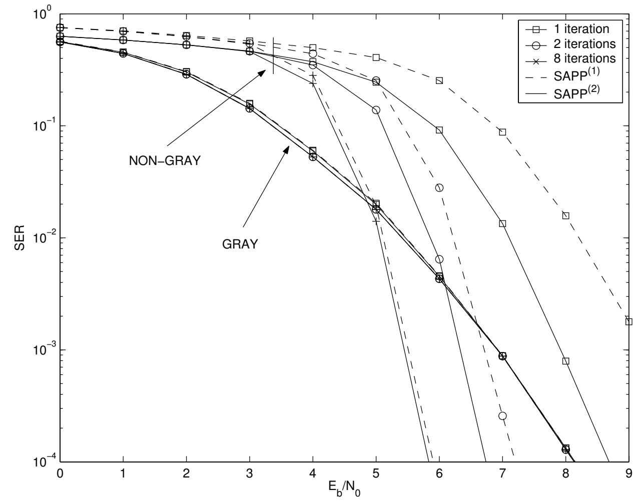

Fig. 2. Symbol error rate based on the a posteriori symbol probabilities for a perfectly synchronized system.

5. NUMERICAL RESULTS

To illustrate the performance of the iterative detector with embedded estimation (as depicted in Fig. 1), we consider the joint estimation and detection of the parameters A and θ. We further assume 16QAM signaling, along with a rate $\frac{1}{2}$ convolutional code with an interleaver size of 1024 bits. We consider both Gray-mapping and an optimized non-Gray mapping [5]. First we evaluate the accuracy of the two algorithms outlined above for the computation of the SAPP's, by means of a symbol error rate (SER) analysis. The SER is determined based on the SAPP's: the symbol with the largest a posteriori probability is compared to the actual transmitted symbol. Fig. 2 compares the SER resulting from the two SAPP computation methods. For Gray mappings, hardly any performance difference is observed between the two SAPP algorithms. There is also little gain experienced by iterating between decoder and demapper for Gray mappings; this is consistent with [3].

For an optimized non-Gray mapping, however, there is a substantial difference between the two methods. Particularly, in early iterations the new method outperforms the traditional bit-to-symbol probability conversion. In our combined estimation and detection scheme, the reliability of the SAPP in early iterations is important, because the accuracy of the parameter estimates from early iterations strictly depends on these. As will be apparent, poor estimates in early iterations have a detrimental impact on the convergence.

Fig. 3 illustrates the performance of the embedded EM-

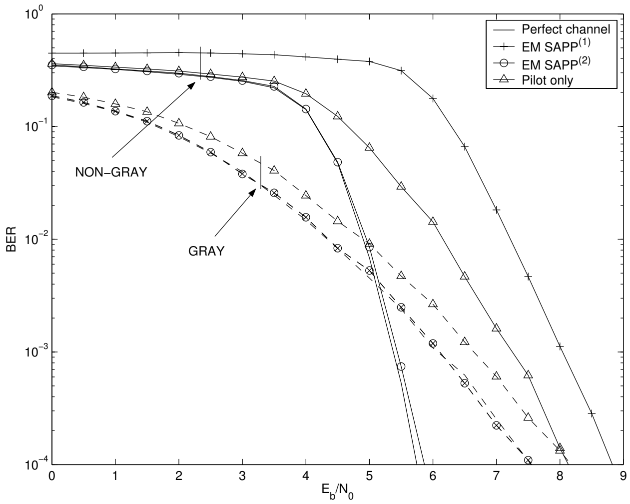

Fig. 3. Bit error rate performance of the iterative detector with embedded estimation of A and θ.

based estimation scheme, compared to detection with perfect channel knowledge and estimation based on 8 pilot symbols only (15 detection iterations in all scenarios). We note that for Gray mapping the use of the *new* SAPP computation method yield no significant performance gain. Hence, the iterative estimation schemes proposed in [7,8] are close to optimum when applied to Gray mapping. For non-Gray mappings, however, the performance gain of the new method is huge. In fact, SAPP computation according to (9) does not converge at all, while computation according to (10) gives excellent results (only 0.1dB degradation compared to perfect synchronization). Note also that the benefit from choosing an optimized mapping is lost when estimating the parameters based on pilot symbols only. This again emphasizes the importance of code-aided estimation.

6. CONCLUSIONS

In this contribution, we have presented a code-aided estimation scheme for BICM-ID. To limit the overhead caused by the estimation, the estimation stages are embedded in the detection stages. A new method is proposed to compute the symbol a posteriori probabilities, which are required for the parameter estimation. In the case of Gray mapping, the new method and the traditional bit-to-symbol probability conversion yield essentially the same accuracy. However, for optimized non-Gray mappings, symbol error rate analysis showed that the new method yields more accurate symbol a posteriori probabilities, especially in early iterations.

The resulting detection with embedded estimation yields a close-to-perfectly synchronized performance, while the estimation causes little extra computational overhead.

7. ACKNOWLEDGMENT

This work has been supported by the Interuniversity Attraction Poles Program P5/11- Belgian Science Policy. The first author also gratefully acknowledges the support from the Fund for Scientific Research in Flanders (FWO-Vlaanderen).

8. REFERENCES

[1] G. Ungerboeck, "Channel Coding with Multi-level/Phase Signals," *IEEE Trans. Inform. Theory*, vol. 1, no. 28, pp. 55–67, Jan. 1982.

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

[3] X. Li and J.A. Ritcey, "Trellis-coded modulation with bit interleaving and iterative decoding," *IEEE Journal on Selected Areas in Comm.*, vol. 17, no. 4, Apr. 1999.

[4] S.Y. Le Goff, "Signal Constellations for Bit-Interleaved Coded Modulation," *IEEE Trans. Inform. Theory*, vol. 49, no. 1, pp. 307–313, Jan. 2003.

[5] F. Schreckenbach, N. Görtz, J. Hagenauer and G. Bauch, "Optimization of Symbol Mappings for Bit-Interleaved Coded Modulation with Iterative Decoding," *IEEE Comm. Letters*, vol. 7, no. 12, pp. 593–595, Dec. 2003.

[6] Y. Huang and J.A. Ritcey, "16-QAM BICM-ID in Fading Channels With Imperfect Channel State Information," *IEEE Trans. Wireless. Comm.*, vol. 2, no. 5, pp. 1000–1007, Sept. 2003.

[7] V. Lottici and M. Luise, "Embedding Carrier Phase Recovery Into Iterative Decoding of Turbo-Coded Linear Modulations," *IEEE Trans. Comm.*, vol. 52, no. 4, pp. 661–669, Apr. 2004.

[8] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey, M Luise and L. Vandendorpe, "Turbo-synchronization: an EM algorithm approach," in *Proc. IEEE ICC*, Anchorage, May 2003.

[9] L.R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal decoding of linear codes for minimising symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, pp. pp. 284–287, Mar. 1974.

[10] A.P. Dempster, N.M. Laird and D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. pp. 1–38, 1977, Series B.