

Bit-Interleaved Turbo-Coded Noncoherent Orthogonal Modulation with Iterative Demodulation and Decoding: Capacity Limits and Convergence Analysis

Shi Cheng and Matthew C. Valenti
Wireless Communications Research Laboratory
Lane Dept. of Comp. Sci. & Elect. Eng.
West Virginia University
Morgantown, WV 26506-6109
email: [shic,mvalenti]@csee.wvu.edu

Abstract— This paper studies bit interleaved coded modulation (BICM) using noncoherent orthogonal modulation (e.g. NFSK). To improve performance, information is fed from the decoder back to the demodulator, using the recently proposed strategy known as BICM with Iterative Demodulation (BICM-ID). The capacity of noncoherent orthogonal modulation under BICM constraints is compared against that with joint demodulation and decoding (which BICM-ID approximates). Also, the convergence behavior of iterative demodulation with turbo decoding is investigated by using density evolution.

I. INTRODUCTION

Communication receivers must often operate in the presence of phase uncertainty. One popular method to handle the lack of phase information is the combination of orthogonal modulation and noncoherent detection, as typified by noncoherent frequency shift keying (NFSK). A key benefit of using orthogonal modulation is that it allows for a tradeoff between energy-efficiency and bandwidth. By using a higher order modulation, the required \mathcal{E}_b/N_o is decreased. In systems that are limited by energy rather than bandwidth (e.g. many military systems and sensor network applications), larger values of M (the number of orthogonal signals in the signal set) are desired.

The capacity of noncoherent modulation was considered by Stark in [1] and is reviewed in Section III. A pragmatic approach to approaching capacity, known as bit interleaved coded modulation (BICM) [2], uses a capacity-approaching binary code, such as a turbo or LDPC code, followed by a bitwise interleaver and an M -ary modulator. A standard receiver for BICM consists of a demodulator that outputs the binary log-likelihood ratio of each code bit, followed by a bitwise deinterleaver and soft-input decoder. However, by constraining the demodulator to make soft decisions on the code bits (rather than on symbols) some information is lost, and as a consequence of the data processing theorem, the capacity will be lower. Following the methodology in [2], we found the capacity of NFSK under BICM constraints, which is worse than that found by Stark (since Stark's capacity assumes

joint demodulation and decoding).

One way to recover the loss due to the BICM constraint is to feed information from the decoder back to the demodulator. This process approximates joint demodulation and decoding and allows some of the capacity loss due to using BICM to be recovered. This type of receiver processing was proposed by Li and Ritcey [3] for two dimensional signal sets and has been termed *BICM-ID*, for *BICM with Iterative Decoding (or Demodulation)*. In [4] we considered BICM-ID for noncoherent orthogonal modulation. In this paper, we build upon our results in [4] by contributing two new results: (1) The aforementioned comparison of capacity under BICM constraints versus the capacity of joint demodulation and decoding, and (2) The convergence analysis of BICM-ID with noncoherent orthogonal modulation.

The convergence analysis builds upon recent work on the convergence of turbo and LDPC codes. The extrinsic information transfer chart (EXIT chart) was introduced by S. ten Brink [5] to analyze the convergence behavior of the iteratively decoded concatenated codes. Average mutual information is measured for both inputs and outputs of the soft input soft output (SISO) decoder. A similar strategy called density evolution was taken by [6][7], based on a signal to noise ratio (SNR) measure (rather than average mutual information). We will later apply this idea to our system model, and show the convergence behavior of BICM-ID in a unique way.

The remaining part of this paper is organized as follows. In Section II, we first consider the system model for bit interleaved coded noncoherent orthogonal modulation with or without iterative demodulation and decoding. Then in Section III, we show that the channel capacity is improved by iterative demodulation and decoding, for M -ary ($M > 2$) NFSK modulation. Numerical simulation results are given in Section IV. Section V studies the convergence behavior of the turbo decoding with iterative demodulation. Finally, Section VI concludes the paper.

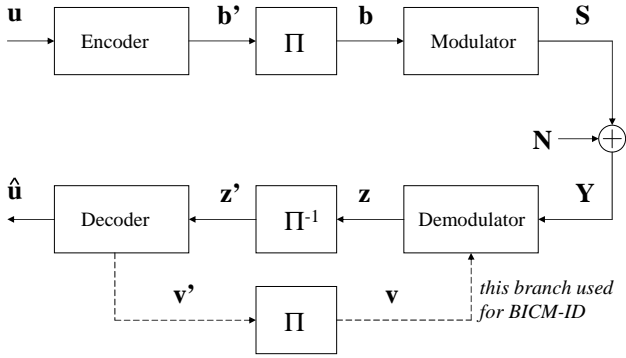


Fig. 1. System model

II. SYSTEM MODEL

Before proceeding further, let us stipulate some notational conventions. Bold lowercase letters will be used to denote vectors, e.g. \mathbf{x} , and bold uppercase will be used for matrices, e.g. \mathbf{X} . All vectors are row-vectors, but can be transposed into column vectors, e.g. \mathbf{x}^T . Vector elements are plain lowercase letters with subscripts beginning at zero, e.g. $\mathbf{x} = [x_0, x_1, \dots, x_{M-1}]$. Matrices are represented as a row of column vectors, e.g. $\mathbf{X} = [\mathbf{x}_0^T, \mathbf{x}_1^T, \dots, \mathbf{x}_{N-1}^T]$. The function $p(\cdot)$ represents the probability of an event, a probability density function, or a probability mass function with the context clearly dependent upon the argument.

A. Transmitter

The discrete-time system model is shown in Fig. 1. A vector $\mathbf{u} \in \{0, 1\}^K$ of message bits is passed through a binary encoder to produce a codeword $\mathbf{b}' \in \{0, 1\}^N$ which is interleaved by a permutation matrix $\mathbf{\Pi}$ to produce the bit-interleaved codeword $\mathbf{b} = \mathbf{b}'\mathbf{\Pi}$. The bit-interleaved codeword is then passed through a M -ary orthogonal modulator to produce the $M \times L$ matrix of $L = \lceil N/\log_2 M \rceil$ symbols $\mathbf{S} = [\mathbf{s}_0^T, \dots, \mathbf{s}_{L-1}^T]$ where the i^{th} symbol $\mathbf{s}_i \in \{\mathbf{e}_0, \dots, \mathbf{e}_{M-1}\}$ is one of M possible M -dimensional elementary vectors \mathbf{e}_m comprised of all zeros except for a one in the m^{th} position. Consider that an arbitrary symbol \mathbf{s} is transmitted. Without loss of generality, assume that the first $\mu = \log_2 M$ bits in \mathbf{b} are gathered to form the symbol, i.e. $\mathbf{s} \Leftrightarrow \{b_0, \dots, b_{\mu-1}\}$. With orthogonal modulation, the mapping of code bits to symbols is unimportant since the symbols are equidistant, and thus natural mapping suffices. In the following, we assume that $\mathbf{s}_0 \Leftrightarrow \{0, \dots, 0\}$.

B. BICM Receiver

The coded symbol stream passes through an AWGN channel, and the input to the demodulator is the matrix of received symbols $\mathbf{Y} = \mathbf{S} + \mathbf{N}$, where $\mathbf{N} = [\mathbf{n}_0^T, \dots, \mathbf{n}_{L-1}^T]$ is a $M \times L$ matrix of i.i.d. complex Gaussian noise samples. The conventional noncoherent BICM demodulator computes the bitwise log-likelihood ratio \mathbf{z} , and the decoder then generates its original information estimation $\hat{\mathbf{u}}$ through the deinterleaved log-likelihood ratio $\mathbf{z}' = \mathbf{z}\mathbf{\Pi}^{-1}$.

C. BICM-ID Receiver

The BICM-ID receiver iterates between demodulation and decoding, with the reliability of the exchanged extrinsic information improved after each half-iteration. The deinterleaved output of the demodulator \mathbf{z}' is passed to the decoder, and the decoder produces extrinsic information \mathbf{v}' of the code bits, which is reinterleaved to form the a priori input \mathbf{v} to the demodulator.

The demodulator then updates the output extrinsic information after every iteration [4],

$$z_k = \log \frac{\sum_{i \in S_k^{(1)}} p(\mathbf{y}|\mathbf{s}_i) p(\mathbf{s}_i|\tilde{\mathbf{v}}, b_k = 1)}{\sum_{i \in S_k^{(0)}} p(\mathbf{y}|\mathbf{s}_i) p(\mathbf{s}_i|\tilde{\mathbf{v}}, b_k = 0)} \quad (1)$$

where $\tilde{\mathbf{v}} = [v_0, \dots, v_{\mu-1}]$ is the portion of \mathbf{v} that corresponds to this symbol, $S_k^{(1)}$ contains the indices of all symbols labelled with $b_k = 1$, and $S_k^{(0)}$ contains the indices of all symbols labelled with $b_k = 0$. After some manipulation, (1) can be expressed as

$$z_k = \max_{i \in S_k^{(1)}}^* \left[\log I_0 \left(\frac{2\mathcal{E}_s |y_i|}{N_o} \right) + \sum_{\substack{j=0 \\ j \neq k}}^{\mu-1} b_j^{(i)} v_j \right] - \max_{i \in S_k^{(0)}}^* \left[\log I_0 \left(\frac{2\mathcal{E}_s |y_i|}{N_o} \right) + \sum_{\substack{j=0 \\ j \neq k}}^{\mu-1} b_j^{(i)} v_j \right]. \quad (2)$$

where $I_0(\cdot)$ is the zeroth order modified Bessel function of the first kind, and the max-star operator is defined in [8],

$$\max_i^* \{x_i\} = \log \left\{ \sum_i e^{x_i} \right\} \quad (3)$$

III. CHANNEL CAPACITY

When normalized to units of M -ary symbols per channel use, the capacity of orthogonal modulation is [1]

$$C = 1 - \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{s}_0) \log_M \left(1 + \sum_{i=1}^{M-1} \Lambda_i(\mathbf{y}) \right) d\mathbf{y}, \quad (4)$$

where \mathbf{y} and \mathbf{s} are as defined in Section II and

$$\Lambda_i(\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{s}_i)}{p(\mathbf{y}|\mathbf{s}_0)} \quad (5)$$

which for noncoherent orthogonal modulation in AWGN is

$$\Lambda_i(\mathbf{y}) = \frac{I_0 \left(\frac{2\mathcal{E}_s |y_i|}{N_o} \right)}{I_0 \left(\frac{2\mathcal{E}_s |y_0|}{N_o} \right)} \quad (6)$$

For $M > 2$, the multidimensional integration in (4) quickly becomes intractable and thus the capacity cannot be readily found using traditional numerical integration techniques. However, it can be found using Monte Carlo integration, as suggested in [2].

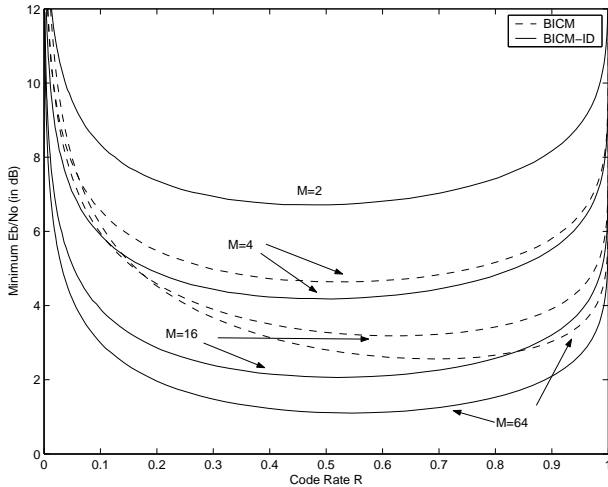


Fig. 2. The information theoretical minimum \mathcal{E}_b/N_o required to achieve arbitrarily low BER for noncoherent orthogonal modulation in AWGN. From top to bottom, results are shown for $M = \{2, 4, 16, 64\}$. For each $M > 2$, two curves are shown: The upper curve constrains the receiver to first make soft decisions on the binary code symbols and then decode using the binary LLR values produced by the demodulator; the lower curve places no constraints on the receiver. The upper curve is the limit for BICM, while the lower curve is the limit for BICM-ID.

The capacity results given above place no constraints upon the receiver design, and thus assume joint demodulation and decoding. The results are an appropriate bound for BICM-ID, which approximates joint demodulation and decoding. On the other hand, the standard BICM receiver takes a segregated approach to demodulation and decoding, and hence has a different capacity. The capacity of BICM is found from [2]

$$\begin{aligned}
 C &= 1 - \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{s}_0) \sum_{k=0}^{\mu-1} \log_M \left(1 + \frac{p(\mathbf{y}|b_k=1)}{p(\mathbf{y}|b_k=0)} \right) d\mathbf{y}, \\
 &= 1 - \log_2(e) \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{s}_0) \max^*(0, z_0) d\mathbf{y}, \quad (7)
 \end{aligned}$$

The above expression exploits the symmetry of orthogonal modulation, and uses (2) to compute z_0 (with all v_j set to zero).

Capacity results are shown in Fig. 2 for four values of M and an AWGN channel. The figure shows the information theoretical minimum \mathcal{E}_b/N_o required to achieve arbitrarily low BER. For each $M > 2$, two curves are shown: One for the BICM constrained receiver and the other for the unconstrained receiver. Note that performance improves with increasing M , but does not necessarily improve with decreasing code rate. This is due to the noncoherent combining penalty and was also observed in [1]. Note the performance gap between BICM and unconstrained reception. The gap increases with increasing M and decreasing code rate.

IV. NUMERICAL RESULTS

To illustrate the effectiveness of the proposed BICM-ID technique for M -ary NFSK, we conducted an extensive set of simulations. For the channel code, the turbo code from the cdma2000 specification was used [9]. We investigated all

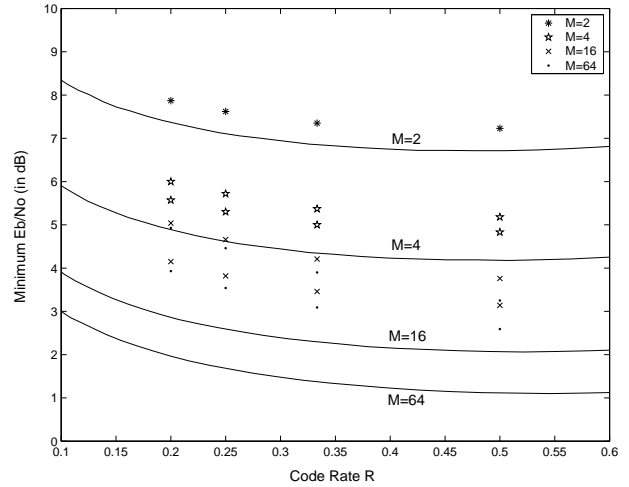


Fig. 3. Minimum \mathcal{E}_b/N_o required to achieve $\text{BER} = 10^{-4}$ as a function of code rate R over an AWGN channel using M -ary NFSK modulation and the $K = 6138$ bit cdma2000 turbo code. For each value of $M > 2$, two points are shown. The upper point is for BICM, while the lower point is for BICM-ID. The capacity curve for unconstrained reception is also shown.

four code rates supported by cdma2000, specifically $R = 1/2, 1/3, 1/4$, and $1/5$. While cdma2000 supports 12 distinct frame sizes, we focused on frames created using $K = 6138$ message bits (we also tested the three larger frame sizes of 9210, 12282, and 20730, but found that their performance was not significantly better than the 6138 bit frame size). The BICM interleaver Π was implemented as a μ by L block interleaver, with bits written into the interleaver row-wise and read out column-wise. We also tried some other interleaver designs, including s-random interleavers and interleavers designed according to the three rules in [10]. However we found that performance was not significantly influenced by interleaver design, presumably due to the fact that the turbo code already contains its own internal interleaver.

In our simulation, we assumed that the average value of \mathcal{E}_b/N_o is known at the receiver. Four values of the modulation order M were considered, $M = 2, 4, 16$, and 64 . For $M > 2$, both BICM and BICM-ID were considered (for $M = 2$, BICM-ID degenerates into BICM and thus separate results are not necessary). In each case, 16 iterations of BICM-ID decoding were performed (with a single local iteration of turbo decoding for each global iteration of BICM-ID). For every data point, the simulation ran until at least 30 frame errors were recorded.

For each case, we found the value of \mathcal{E}_b/N_o for which the $\text{BER} = 10^{-4}$. These values are indicated in Fig. 3. For each value of $M > 2$, two points are shown. The upper point is for BICM, while the lower point is for BICM-ID. In Table I, we list the value of \mathcal{E}_b/N_o required to achieve a BER of 10^{-5} for each modulation order using the rate $1/2$ cdma2000 turbo code (which has better performance than the other code rates).

Now consider how performance improves as a function of the number of iterations. BER curves are shown in Fig. 7 and Fig. 8 for $M = 16$ and $M = 64$, respectively, for code rate $R = 1/2$. Each plot shows curves for both

TABLE I

MINIMUM \mathcal{E}_b/N_o REQUIRED IN AWGN TO ACHIEVE A BER OF 10^{-5}
USING THE 1/2 RATE 6138 BIT CDMA2000 TURBO CODE, M-ARY
NONCOHERENT FSK, AND EITHER BICM OR BICM-ID.

M	Coded Performance		Capacity	
	BICM	BICM-ID	BICM	BICM-ID
2	7.27 dB	7.27 dB	6.71 dB	6.71 dB
4	5.23 dB	4.89 dB	4.64 dB	4.18 dB
16	3.82 dB	3.18 dB	3.27 dB	2.07 dB
64	3.32 dB	2.64 dB	2.81 dB	1.11 dB

BICM (dashed lines) and BICM-ID (solid lines). From right to left, the performance after iterations 1, 2, 3, 4, 5, 10, and 16 are shown. The convergence thresholds determined through density evolution are plotted in the same figures (the derivation of these thresholds is discussed below). The curves indicate that the performance of BICM-ID after 4 iterations is always better than the performance of BICM after all 16 iterations. This implies that, although BICM-ID is marginally more complex per iteration than BICM, a system using BICM-ID can actually be much less complex than BICM because it can achieve the same performance by running one-fourth as many iterations.

V. CONVERGENCE ANALYSIS

To illustrate the gain achieved by BICM-ID, we use density evolution to analyze the convergence behavior of BICM-ID compared with BICM. The density evolution method has been previously applied by [6], [7] to predict the waterfall region of turbo codes. [6] and [7] both apply a Gaussian assumption to the log-likelihood ratios or extrinsic information flowing out the upper and lower SISO decoders. The Gaussian distributions are assumed to obey the consistency condition which assures that the Gaussian distribution with mean μ has variance $\sigma^2 = 2\mu$. In this case, $\text{SNR} = \mu^2/\sigma^2 = \mu/2$ is a good measure of the quality of the decoder input or output, since a high SNR implies the two peaks of the log-likelihood ratio can be easily discriminated.

The SNR statistics of the extrinsic information flowing between the SISOs can be tracked to analyze the convergence behavior of the turbo decoder. However, since the Gaussian assumption does not match the true distribution exactly, the variance σ^2 is not exactly twice the mean μ . Thus, there are three different ways to measure the SNR: $\text{SNR} = \mu^2/\sigma^2$, $\text{SNR} = \mu/2$ and $\text{SNR} = [Q^{-1}(P_b)]^2$, where P_b is the error rate of the hard decision of the log-likelihood ratios. [7] indicates that the second way is a good measure for the log-likelihood ratios or extrinsic information generated by the SISO decoders, and usually gives a better prediction. We will use this measurement for the analysis in this paper, except that we observed the first way, $\text{SNR} = \mu^2/\sigma^2$, characterizes the approximately Gaussian output of the demodulator quite well.

Consider the input and output SNRs for each SISO at each iteration, as shown in Fig. 4. A particular \mathcal{E}_b/N_o from the channel enables the SISO demodulator to produce an output

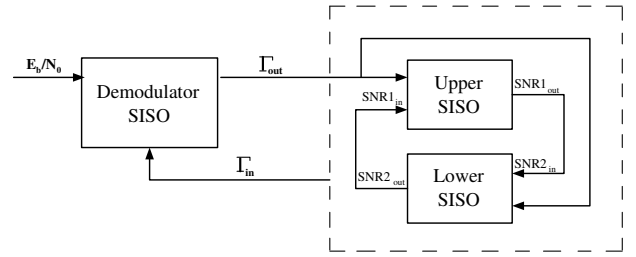


Fig. 4. Receiver model used by the density evolution analysis.

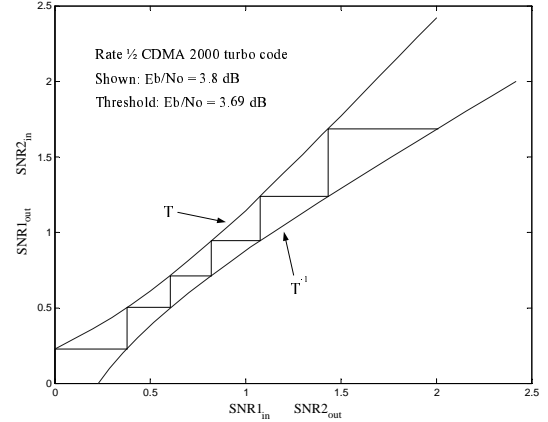


Fig. 5. Convergence of the BICM receiver, 16-ary NFSK, rate 1/2 cdma2000 turbo code

with SNR $\Gamma_{out} = D(\Gamma_{in}, \mathcal{E}_b/N_o)$, where Γ_{in} denotes the SNR at the output of the turbo decoder. Similarly, we define the nonlinear functions T_1 and T_2 for the upper and lower SISO decoders. Thus, we have $\text{SNR1}_{out} = T_1(\text{SNR1}_{in}, \Gamma_{out})$ and $\text{SNR2}_{out} = T_2(\text{SNR2}_{in}, \Gamma_{out})$. Also, since $\text{SNR1}_{out} = \text{SNR2}_{in}$, $\text{SNR2}_{out} = T_2(T_1(\text{SNR1}_{in}, \Gamma_{out}), \Gamma_{out})$. Note that D , T_1 and T_2 are all nondecreasing functions with respect to either argument [6].

In a conventional BICM receiver, the SISO demodulator only computes the bitwise log-likelihood ratio once. Thus, the SNR input to the SISO decoders Γ_{out} is fixed for the whole decoding process, no matter how many local iterations the turbo decoder executes. Since D is nondecreasing, it is equivalent to write the SISO decoders as functions of the channel SNR, i.e. $\text{SNR1}_{out} = \tilde{T}_1(\text{SNR1}_{in}, \mathcal{E}_b/N_o)$ and $\text{SNR2}_{out} = \tilde{T}_2(\text{SNR2}_{in}, \mathcal{E}_b/N_o)$. In this paper, $T_1 = T_2$ and $\tilde{T}_1 = \tilde{T}_2$ since cdma2000 code uses identical constituent RSC encoders. We evaluated the functions D and T ($T = T_1 = T_2$) through Monte Carlo simulations. Fig. 5 plots the function T and T^{-1} for 16-ary NFSK BICM at $\mathcal{E}_b/N_o = 3.8$ dB. The trajectory shows the progress of the decoder's iterations. The SNR keeps rising and the decoder successfully converges as long as there is no early intersection between T and T^{-1} . The threshold for this case is found at $\mathcal{E}_b/N_o = 3.69$ dB.

The BICM-ID receiver acts just like the BICM receiver during the first pass. However, after the first pass, nonzero extrinsic information is generated to improve performance of the demodulator, which can be viewed as the enhancement for

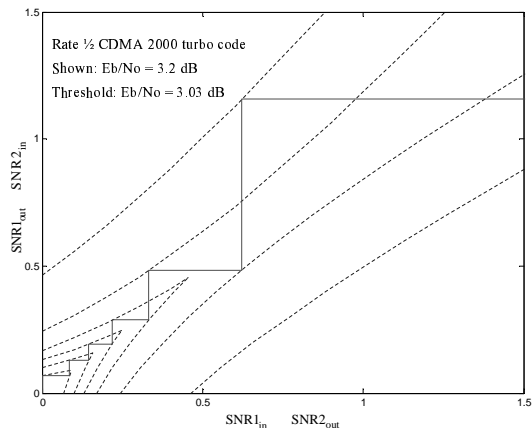


Fig. 6. Convergence of the BICM-ID receiver, 16-ary NFSK, rate 1/2 cdma2000 turbo code.

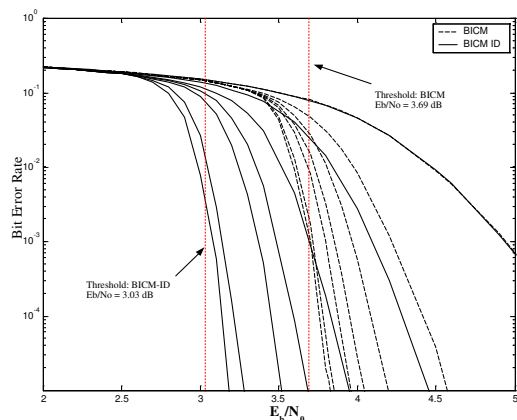


Fig. 7. BER performance in AWGN of the $R = 1/2$ input-length $K = 6138$ bit cdma2000 turbo code using 16-ary NFSK and both BICM (dashed line) and BICM-ID (solid line). From right to left, the curves show performance after 1, 2, 3, 4, 5, 10, and 16 iterations.

the output SNR of the demodulator SISO. Thus, the Γ_{out} is improved after every iteration. Fig. 6 shows the trajectory of 16-ary NFSK BICM-ID at $\mathcal{E}_b/N_o = 3.2$ dB. Because \mathcal{E}_b/N_o is much lower than the BICM threshold, the SNR would get stuck at $\text{SNR}_{1,in} \approx 0.1$ dB when using a BICM receiver. However, the BICM-ID receiver keeps on improving the input SNR Γ_{out} to the decoder. The trajectory, which is reflected by an higher T or lower T^{-1} curve, survives through the narrow area at low SNR, and the turbo decoder finally converges. We found the threshold for 16-ary NFSK BICM-ID is at $\mathcal{E}_b/N_o = 3.03$ dB, 0.66 dB smaller than the BICM threshold.

The accuracy of the above convergence analysis can be verified by indicating the thresholds on the BER curves. Fig. 7 and Fig. 8 are the BER curve for 16-ary NFSK and 64-ary NFSK respectively. Fig. 7 shows the 0.66 dB gain acquired by density evolution is very close to the gain in the BER simulation, which is about 0.74 dB at BER 10^{-5} after 16 iterations. Fig. 8 shows the convergence thresholds determined by density evolution is at $\mathcal{E}_b/N_o = 3.3$ dB and 2.6 dB for BICM and BICM-ID respectively. The gain predicted by

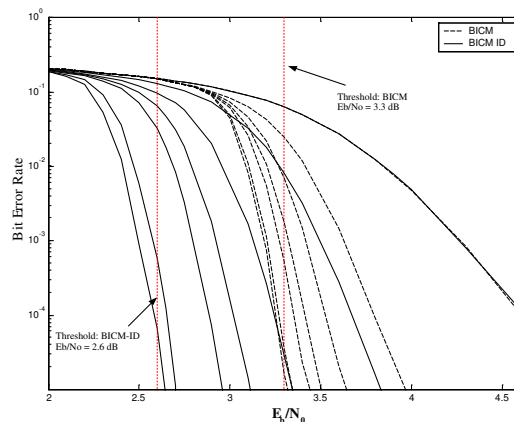


Fig. 8. BER performance in AWGN of the $R = 1/2$ input-length $K = 6138$ bit cdma2000 turbo code using 64-ary NFSK and both BICM (dashed line) and BICM-ID (solid line). From right to left, the curves show performance after 1, 2, 3, 4, 5, 10, and 16 iterations.

density evolution is about 0.7 dB, very close to the 0.68 dB at BER 10^{-5} after 16 iterations.

VI. CONCLUSIONS

The performance of BICM can be improved by feeding soft decisions on the code symbols from the decoder back to the demodulator. Such a process allows some of the loss due to using BICM reception to be recovered. A modified density evolution analysis can be used to predict the threshold for BICM with iterative demodulation. Both the density evolution analysis and the simulation results indicate a gain of approximately 0.7 dB for $M=16$ and $M=64$ when using the proposed iterative reception strategy.

REFERENCES

- [1] W. E. Stark, "Capacity and cutoff rate of noncoherent FSK with nonselective Rician fading," *IEEE Trans. Commun.*, vol. 33, pp. 1153–1159, Nov. 1985.
- [2] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.
- [3] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Letters*, vol. 1, pp. 169–171, Nov. 1997.
- [4] M. Valenti, E. Hueffmeier, B. Bogusch, and J. Fryer, "Towards the capacity of noncoherent orthogonal modulation: BICM-ID for turbo coded NFSK," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, (Monterey, CA), Nov. 2004.
- [5] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, pp. 1727–1737, Oct. 2001.
- [6] H. E. Gamal and A. R. Hammons, "Analyzing the turbo decoder using the gaussian approximation," *IEEE Trans. Inform. Theory*, vol. 47, pp. 671–686, Feb. 2001.
- [7] D. Divsalar, S. Dolinar, and F. Pollara, "Iterative turbo decoder analysis based on density evolution," *The Telecommunications and Mission Operations Progress Report 42-144*, Feb. 2001.
- [8] A. Viterbi, "An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 260–264, Feb. 1998.
- [9] Third Generation Partnership Project 2 (3GPP2), "Physical layer standard for cdma2000 spread spectrum systems, release C," *3GPP2 C.S0002-C Version 1.0*, pp. 115–122, May 28 2002.
- [10] X. Li, A. Chindapol, and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding and 8-PSK signaling," *IEEE Trans. Commun.*, vol. 50, pp. 1250–1257, Aug. 2002.