EXHIBIT 1

Bit-Interleaved Coded Modulation With Iterative Decoding and 8PSK Signaling

Xiaodong Li, Aik Chindapol, *Member, IEEE*, and James A. Ritcey, *Member, IEEE*

***Abstract*—We have suggested bit-interleaved coded modulation with soft decision iterative decoding (BICM-ID) for bandwidth-efficient transmission over Gaussian and fading channels. Unlike trellis coded modulation, BICM-ID has a small free Euclidean distance but large diversity order due to bit interleaving. With iterative decoding, soft bit decisions can be employed to significantly improve the conditional intersignal Euclidean distance. This leads to a large coding gain, comparable to that of turbo TCM, over both Gaussian and Rayleigh fading channels with much less system complexity. We address critical design issues to enhance the decoding performance and provide the analytical bounds on the performance with an ideal feedback assumption. We investigate the performance characteristics of BICM-ID through extensive simulations and show that at high signal to noise ratios, the performance of BICM-ID converges to the performance assuming error-free feedback.**

***Index Terms*—BICM, coded modulation, digital communications, iterative decoding, turbo codes.**

I. Introduction

According to information theory, block code performance can be improved by increasing the codeword length. Yet, for a convolutional code or an equivalent block code formed from a convolutional code, the decoding performance is related to the constraint length of the code [1]. Typically, one can not benefit from using a long input data sequence, because the bits far apart on the trellis do not interact. Increasing the constraint length may bring significant improvement, but at the expense of exponentially increasing complexity in the maximum-likelihood (ML) decoder.

One clever way to circumvent the above dilemma is the recently proposed turbo coding scheme [3], [4], where two or more short-memory convolutional codes are concatenated in parallel or in serial. Due to the pseudorandom interleaving, a "global interaction" is introduced among the bits over an entire block. As a result, error protection is achieved not only through the constraints on the local trellis transitions, but also through the influence of other trellis sections. Although a true ML decoder for such concatenated codes is hard to implement, iterative decoding methods which employ the maximum *a posteriori* probability (MAP) rule for each individual decoder have been shown to provide near-capacity performance [3]–[5]. Compared with convolutional codes, turbo codes effectively take advantage of the potential of large block length but with the reasonable decoding complexity of simple constituent codes.

Another simpler approach is to use iterative decoding with a serial concatenation of encoding, bit-by-bit interleaving and high-order modulation. Unlike turbo codes, this scheme requires *only one* set of encoder/decoder; therefore, the receiver complexity is significantly reduced. At a first glance, the block diagram is no different from that of conventional symbol-interleaved trellis-coded modulation (TCM), a bandwidth-efficient coding approach suggested by Ungerboeck [6]. Indeed, the scheme, called bit-interleaved coded modulation (BICM) [8], was first suggested by Zehavi [7] to increase the time diversity of coded modulation and therefore to improve the performance of TCM over fully interleaved Rayleigh fading channels. However, this improvement is achieved at the expense of reduced free (squared) Euclidean distance (FED), leading to a degradation over nonfading Gaussian channels [7], [8].

In this paper, we show that BICM, a bandwidth-efficient approach primarily considered for fading channels in the past, can in fact be used to provide excellent performance over both Gaussian and fading channels, with *iterative decoding* (ID). To maximize the gain of ID, we make critical changes to traditional Gray labeling used in Zehavi's BICM transmitter design. We call our scheme BICM with iterative decoding (BICM-ID) [9], [10].

The goal of this paper is to give a comprehensive set of performance analysis, simulation results and new labeling maps. In Section II, we first briefly review the scheme of BICM and its conventional decoding [7], [8]. There we expose the reasons for the performance degradation of BICM compared with conventional TCM over Gaussian channels. In Section III, we address system design issues critical to the performance of BICM-ID and give detailed information on our iterative decoding algorithm, signal labeling method and interleaver design. In Section IV, we provide performance analysis and show extensive simulation results for BICM-ID for both AWGN and Rayleigh fading channels. Section V concludes the paper.

Paper approved by R. D. Wesel, the Editor for Coding and Communication Theory of the IEEE Communications Society. Manuscript received December 17, 1999; revised November 19, 2000 and July 20, 2001. This work was supported by the National Science Foundation under Grant CCR-0073391 and a Royalty Research Fund Grant from the University of Washington. This paper was presented in part at the 33rd Annual Conference on Information Sciences and Systems (CISS'99), March 17–19, 1999, Baltimore, MD, and at the 1999 IEEE International Conference on Communications (ICC'99), June 6–10. 1999, Vancouver, BC, Canada.

X. Li was with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA. He is now with Broadstorm Telecommunications Inc., Bellevue, WA 98004 USA (e-mail: xli@broadstorm.com).

A. Chindapol was with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA. He is now with the Networks Division, Siemens Information and Communication Mobile LLC, San Diego, CA 92127 USA (e-mail: aik.chindapol@icm.siemens.com).

J. A. Ritcey is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA (e-mail: ritcey@ee.washington.edu).

Publisher Item Identifier 10.1109/TCOMM.2002.801524.

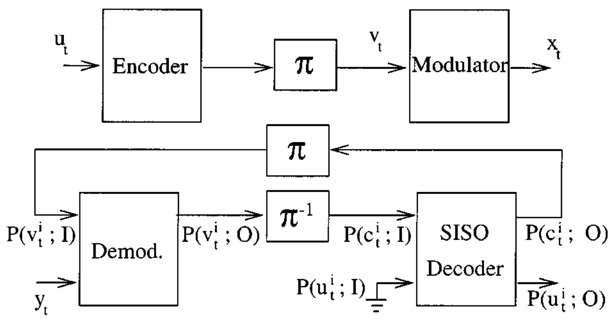

Fig. 1. Block diagram BICM-ID with soft feedback.

II. Review of BICM

A. The BICM Transmitter

The BICM transmitter is a serial concatenation of the encoder, the bit interleaver π and the memoryless modulator as shown in Fig. 1. Note that the pseudorandom interleaver π permutes the encoding output *binary bits*, instead of coded symbols using a conventional symbol-interleaved system. To simplify our discussion, we assume the information transmission rate of 2 bits/s/Hz using a rate-2/3 convolutional code and 8PSK modulation. Extensions to other information rates, code rates or modulation schemes are possible. For example, BICM-ID with 16QAM for fading channels is studied in [20].

Denote the two input bits of a rate 2/3 encoder at time t by $\mathbf{u}_t = [u_t^1, u_t^2]$ and its corresponding three output bits (a code symbol) by $\mathbf{c}_t = [c_t^1, c_t^2, c_t^3]$, where u_t^i or c_t^i is the ith bit. After permutation by a pseudorandom block interleaver, each three binary bits of the interleaver output are grouped together, $\mathbf{v}_t = [v_t^1, v_t^2, v_t^3]$ and are mapped to a complex channel symbol x_t chosen from M-ary constellation χ by a signal label μ

$$x_t = \mu(\mathbf{v}_t),\ x_t \in \chi \tag{1}$$

where the 8PSK signal set is $\chi = \{e^{j\ell 2\pi/8}, \ell = 0, \ldots, 7\}$.

With coherent detection, the received discrete-time baseband signal is

$$y_t = \rho_t \sqrt{E_s} x_t + z_t \tag{2}$$

where ρ_t is the fading coefficient, E_s is the symbol energy, and z_t is complex additive white Gaussian noise with one-sided spectral density N_0. For the AWGN channel, $\rho_t = 1$. For a frequency nonselective Rayleigh fading channel, ρ_t is Rayleigh distributed with $E(\rho_t^2) = 1$. In this paper, we assume perfect channel state information (CSI); hence, ρ_t is perfectly estimated and available to the receiver.

B. Conventional Decoding for BICM

Due to bit-based interleaving, true ML decoding of BICM requires *joint* demodulation and convolutional decoding and is therefore too complex to implement in practice. In [7], Zehavi

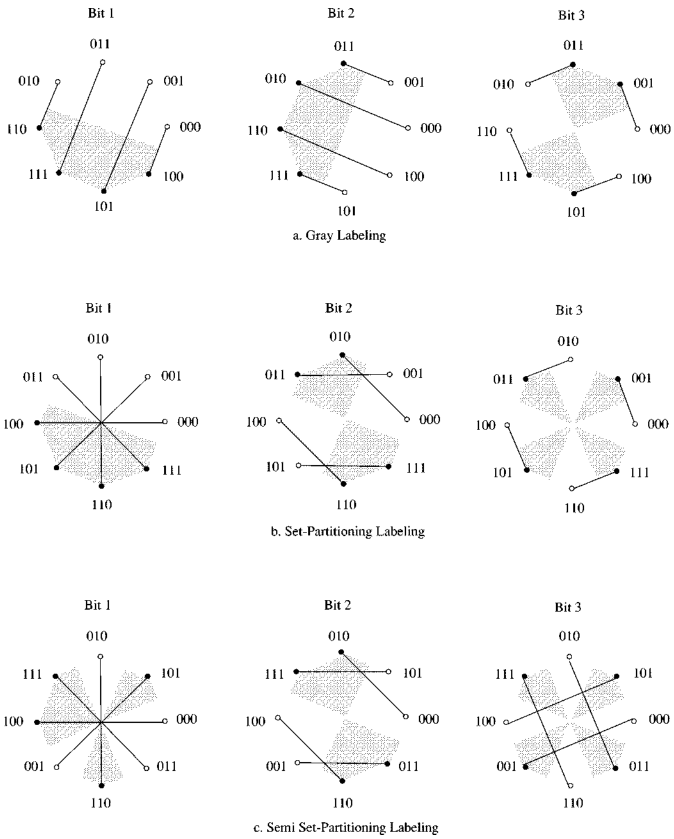

Fig. 2. For various labeling maps, the shaded regions correspond to the decision regions for each bit taking the value of 1. (a) Gray. (b) Set partitioning. (c) Semi set partitioning.

signal y_t, six bit metrics are generated, using the ML rule. For the three binary bits and 8PSK symbols

$$\begin{aligned}\lambda(v_t^i = b) &= \log P(y_t | v_t^i = b)\\ &\sim \log \sum_{x \in \chi_b^i} P(y_t|x), \qquad i = 1,2,3;\ b = 0,1\end{aligned} \tag{3}$$

where the signal subsets are $\chi_b^i = \{\mu([v^1, v^2, v^3]) | v^i = b\}$. The notation $\sim$ indicates replacement by an equivalent statistic. For 8PSK, the size of each subset χ_b^i is 4.

In practice, the log-sum calculation in (3) is computed either by approximation

$$\log \sum_j a_j \approx \max_j \log a_j \tag{4}$$

or by table lookup for better accuracy. Finally, $\log P(y_t|x)$ is replaced by the squared Euclidean distance $\| y_t - \rho_t x \|^2$.

C. Degradation of BICM Over Gaussian Channels

Although BICM performs well over fading channels because of an increase in diversity order, one pitfall of BICM is the degradation over Gaussian channels due to the "random modulation" caused by bit interleaving [7]. For example, referring to Fig. 2, where the shaded regions correspond to all received symbols for which bit i =1, 2, 3 takes on the value "1". With bit interleaving and suboptimal decoding, the symbol may originate

It can be shown that the FED of BICM is $d_E^2 = d_H d_0^2$ [7], [11], where d_H is the free Hamming distance of a code and d_0 is the smallest Euclidean distance between the modulation constellation points. For 8PSK modulation, $d_0 = 2\sqrt{E_s}\sin(\pi/8)$, where E_s is the energy of a channel symbol. In general, the FED of BICM is a few dB below its counterpart TCM [7]. Therefore, conventional BICM is less efficient than TCM for Gaussian channels.

III. BICM-ID

Bit interleaving connects the coded bits, originally far apart in the sequence, to the same channel symbol. With ideal interleaving, coded bits forming a channel symbol are independent; therefore, the feedback from strong data sections (with less influence of channel noise) can remove the ambiguity in the high-order demodulation and enhance the decoding of weak data sections (those subject to undesirable noise patterns). With the perfect knowledge of the other two bits, which are provided by the decoding feedback, 8PSK modulation is effectively reduced to binary modulation for each bit position. Hence, the interconstellation distance for the binary modulation can be significantly increased.

Of course, if the feedback contains errors, we have picked a wrong binary constellation. Therefore, it is also important to reduce the effect of feedback errors and to control error propagation. These factors are considered in system design by using soft-decision feedback and well-designed interleavers. While more complex than our hard-decision feedback [10], soft feedback is the key to realizing the inherent gains in BICM while mitigating error propagation.

A. Iterative Decoding Using Soft Feedback

The recent success of turbo codes has demonstrated the advantages of iterative processing in the decoding of concatenated schemes. A good introduction by Hagenauer can be found in [12], where the method is called the "turbo principle." Note that iterative decoding was also considered by Seshadri and Sunderberg for multilevel coded modulation [13]. In [14], Woerz and Hagenauer also used the reliabilities of the decoding results to control the feedback.

As shown in Fig. 1, our receiver uses a suboptimal, iterative method through individually optimal, but separate demodulation and convolutional decoding steps. The *a posteriori* probabilities for the coded bits can be calculated as

$$P\left(v_t^i = b|y_t\right) \sim \sum_{x_t \in \chi_b^i} P(x_t|y_t) \sim \sum_{x_t \in \chi_b^i} P(y_t|x_t)P(x_t). \quad (5)$$

Note that, compared with (3), (5) considers the *a priori* probability $P(x_t)$.

At the initial demodulation, we assume the equally likely prior $P(x_t)$. Then, the soft-input–soft-output (SISO) module [5] is used for convolutional decoding and to generate the *a posteriori* bit probabilities for the information and coded bits. Following the notation of Benedetto *et al.* [5], we denote by $P(q;I)$ the *a priori* probability for a random variable q and dition, $P(u_t^i;O)$ and $P(c_t^i;O)$ are the *extrinsic information*, a term well explained in the literature of turbo codes [3], [5].

On the second pass, $P(c_t^i;O)$ is interleaved and fedback, as $P(v_t^i;I)$, to the demodulator. Assuming $P(v_t^1;I)$, $P(v_t^2;I)$ and $P(v_t^3;I)$ are independent (a good interleaver assures near independence), we obtain, for each $x_t \in \chi$,

$$\begin{aligned} P(x_t) &= P\left(\mu\left(\left[v^1(x_t), v^2(x_t), v^3(x_t)\right]\right)\right) \\ &= \prod_{j=1}^{3} P(v_t^j = v^j(x_t); I) \end{aligned} \quad (6)$$

where $v^j(x_t) \in \{0,1\}$ is the value of the jth bit of the label for x_t. Using (5) and (6), we derive the *extrinsic a posteriori* bit probabilities for the second-pass demodulation

$$\begin{aligned} P\left(v_t^i = b; O\right) =& \frac{P\left(v_t^i = b|y_t\right)}{P\left(v_t^i = b; I\right)} \\ =& \frac{\left(\sum\limits_{x_t \in \chi_b^i} P\left(y_t|x_t\right) P\left(x_t\right)\right)}{P\left(v_t^i = b; I\right)} \\ =& \sum_{x_t \in \chi_b^i} \left(P(y_t|x_t) \prod_{j \neq i} P\left(v_t^j = v^j(x_t); I\right)\right), \\ i =& 1,2,3;\ b = 0,1. \end{aligned} \quad (7)$$

Therefore, when recalculating the bit metrics for one bit, we only need to use the *a priori* probabilities of the other bits in the same channel symbol. The regenerated bit metrics are put into the decoder and we iterate demodulation and decoding. The final decoded output is the hard decision on the extrinsic bit probability $P(u_t^i;O)$, which is also the *total a posteriori* probability since $P(u_t^i;I)$ is unused.

In our implementation, the SISO decoder uses an additive "log-map" algorithm [5]. Also, the log-sum in (5) is approximated by max operations, aided by table lookups. These approaches greatly reduce the system complexity.

B. Signal Labeling

In our design it is critical to note that different decoding methods are optimized with different signal constellation labels. In this paper, we consider Gray, set-partitioning (SP), and semi set-partitioning (SSP) as examples. A comparison of these labeling schemes for 8PSK is shown in Fig. 2. The decision regions for each bit in χ_1^i are shown in the shaded areas (only shown inside the unit circle) while the unshaded regions correspond to χ_0^i. It can be seen that all labeling schemes have the same minimum Euclidean distance between subsets of χ_1^i and χ_0^i but a different number of nearest neighbors. Therefore, for conventional BICM, Gray labeling has been considered to be optimal [7], [8] due to the smallest number of nearest neighbors.

With perfect knowledge of all other bits, 8PSK modulation is translated to binary modulation selected from four possible sets of binary modulation. It can be seen that iterative decoding of BICM not only increases the intersubset Euclidean distance, channels. Fig. 2 also illustrates the increase in the minimum Euclidean distance between subsets. It is obvious that Gray labeling is not the preferred choice since the minimum distance between subsets is not increased. More detailed analysis on the effect of labeling schemes is given in the next section where the analytical bound for BICM-ID is derived.

C. Interleaver Design

The interleaver design is critical to the high performance of BICM-ID. We use pseudorandom interleavers with the following design objectives: 1) to increase FEDC and 2) to mitigate error propagation during iterative decoding. Readers familiar with turbo codes can see that some of our ideas are inspired by the spread-random interleavers suggested in [16]. Here are our design rules.

Rule 1: Modularity: The bit positions before and after interleaving must have the same modulo-m value, i.e., for 8PSK and $m = 3$, Bit 1 in the encoder output bit stream can only be mapped to one of the positions at Bit 1, 4, 7,… in the interleaver output. Essentially, the entire interleaver is composed of m subinterleavers. This ensures that the coded bits with different protection, due to their different positions at the channel-symbol labels, are distributed uniformly along the trellis.

Rule 2: Reverse Spread: The m bits going to the same channel symbol must be at least S_0 trellis stages apart from each other. This ensures feedback independence in bit metric recalculation and mitigates the error propagation through iterative decoding. S_0 much larger than the code constraint length is easily achievable. For a block containing $N = 2000$ information bits, a typical S_0 is 50.

Rule 3: Forward Spread: The bits co-channel-symboled with the bits from a trellis segment of S_1 stages should be spread at least S_2 stages far from each other. This ensures that a burst of decoding errors spread evenly over the entire trellis and does not heavily affect, through bit-metric recalculation using the feedback, another short trellis segment. It is usually difficult to enforce Rule 3 for a short block, even though the window sizes S_1 and S_2 are chosen very small. Therefore, we only try to minimize the number of violations. For $N = 2000$, the typical values of S_1 and S_2 we use are 3 and 6.

The design rules form a multicriterion objective function, of which each component can only be partially optimized in practice. Our interleaver design algorithm uses these design rules as heuristics that guide iterative changes to an initial pseudorandomly drawn permutation.

IV. Performance Evaluation

A. Performance Bound for AWGN Channels

We first derive a BER upper bound for an idealized situation assuming error-free feedback (EFF). With ideal feedback, the 8-PSK channel is transformed into 3 independent BPSK channels. Normalizing $E_s = 1$, the minimum intersignal Euclidean distance for the three BPSK channels are $d_1 = 2$, $d_2 = \sqrt{2}$ and $d_3 = 2\sin(\pi/8)$ for set-partitioning labeling while $d_1 = d_2 = d_3 = 2\sin(\pi/8)$ for Gray labeling.

TABLE I
Comparison Between FED of TCM and BICM and FEDC of BICM-ID. Rate-2/3 Codes and 8PSK Modulation With $E_s = 1$. Punctured Codes for BICM and BICM-ID

# of States	TCM	BICM	BICM-ID		
			Gray	SP	SSP
8	4.59	2.34	2.34	7.17	12.00
16	5.17	2.34	2.34	8.34	13.41
32	5.76	3.51	3.51	8.93	15.41
64	6.34	3.51	3.51	10.10	20.00

quence $\underline{\hat{\mathbf{c}}}$ is selected at the decoder. Denote $w_i, i = 1, 2, 3$, the Hamming weight of the error pattern corresponding to the ith bit position of the encoder output. The total Hamming weight of the error pattern $w = w_1 + w_2 + w_3$. The squared Euclidean distance between the modulated sequences $\underline{\mathbf{c}}$ and $\underline{\hat{\mathbf{c}}}$ is

$$d_E^2(w_1, w_2, w_3) \ge w_1 d_1^2 + w_2 d_2^2 + w_3 d_3^2. \tag{8}$$

Therefore, the PEP is given by

$$p(w_1, w_2, w_3) = Q\left(\sqrt{\frac{d_E^2(w_1, w_2, w_3)}{2N_0}}\right) \tag{9}$$

where $Q(x) = \int_x^\infty e^{-t^2/2} dt/\sqrt{2\pi}$. Finally, we obtain the upper (union) bound on the bit error probability for a rate-2/3 code as

$$P_b \le \frac{1}{2} \sum_{w_1, w_2, w_3} a(w_1, w_2, w_3) p(w_1, w_2, w_3) \tag{10}$$

where $a(w_1, w_2, w_3)$ is the total information weight corresponding to all the error events with coded output weight (w_1, w_2, w_3). With ideal feedback, 8PSK modulation is translated to binary modulation regardless of the labeling map. Therefore, from (10) it can be seen that only the FED conditioned on the ideal feedback (FEDC), which is defined as $\min\{d_E^2(w_1, w_2, w_3); w_1, w_2, w_3\}$ dominates the asymptotic performance of BICM-ID.

In Table I, we compare the FED of TCM and BICM and the FEDC of BICM-ID. The large increase in FEDC over FED shows the potential of BICM-ID. Our extensive simulation results confirm that soft iterative decoding mitigates error propagation and practically realizes the potential of the conditional free Euclidean distance—FEDC. It can be seen that Gray labeling, a signal mapping optimized for conventional BICM [8], shows no improvement due to iterative decoding. Although SSP labeling has the largest FEDC, it has the largest number of nearest neighbors, which affect the first round performance, among all labeling maps considered as shown in Fig. 2. Therefore, an iterative decoding gain may not be evident at BER values of interest and SSP labeling is not further

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase® Smarter legal research.