

# Union Bound Analysis of Bit Interleaved Coded Orthogonal Modulation with Differential Precoding

Shi Cheng and Matthew C. Valenti

Lane Department of Computer Sci. and Electrical Eng.

West Virginia University

Morgantown, WV 26505, USA

Email: {shic, mvalenti}@csee.wvu.edu

**Abstract**—A union bound on the performance of bit interleaved coded orthogonal modulation (BICOM) is presented. The bound is derived by representing the system as an equivalent serially concatenated convolutional code (SCCC). Differential precoding prior to orthogonal modulation is shown to provide an interleaver gain, and termination of the differential precoder with a tail bit is taken into account in the bound. The bounds are evaluated for both coherent and noncoherent detection in AWGN and fully interleaved Rayleigh fading and shown to be tight by comparison against simulation results.

## I. INTRODUCTION

When signaling with nonbinary signal sets ( $M > 2$ ), it is often more convenient to use a *binary* channel code, rather than one with an alphabet matched to the signal set. The combination of binary channel code, bitwise interleaver, and M-ary modulator is called *bit interleaved coded modulation* (BICM) [1]. This combination can be viewed as a serially concatenated code [2], with the modulator serving as the inner code. From this perspective, a receiver that iterates between a soft-input, soft-output (SISO) demodulator and a SISO decoder is capable of approximating true ML joint demodulation and decoding. Iterative demodulation and decoding has been previously considered by several authors, including Li and Ritcey [3] who coined the term *bit interleaved coded modulation with iterative decoding* (BICM-ID).

When two dimensional signaling, e.g. QAM or PSK, is used with a capacity approaching code, there is little to gain by using BICM-ID because there is virtually no difference between the capacity of unconstrained *coded modulation* (CM) and gray-labelled BICM [1]. On the other hand, with orthogonal modulation, a gray labelling does not exist since all neighbors are equidistant. Therefore, with orthogonal modulation there will always be a non-negligible loss in capacity when BICM is used instead of CM, and the use of BICM-ID can offer significant performance gains over BICM [4].

In this paper, we consider bit interleaved coded *orthogonal* modulation (BICOM) from the perspective of being a serially concatenated code (SCC). In our baseline system, the inner code is simply a memoryless mapper that transforms groups of  $\log_2 M$  bits to orthogonal symbols. This system is identical to the one considered in [4], but now the emphasis is on its performance at high SNR through union bounding techniques instead of its performance at low SNR through EXIT analysis (which was already considered in [4]).

When BICOM is viewed as a SCC, the inner encoder is nonrecursive. However, a well known result from [2] is that an interleaver gain is only possible if the inner code is recursive. Likewise, an interleaving gain is achieved with iterative demodulation and decoding only if the *modulator* is recursive [5]. One way to guarantee this condition is to use modulation that is inherently recursive, such as continuous phase modulation (CPM) [6] or differential phase shift keying (DPSK) [7]. Another way is to precede an otherwise memoryless modulator with a recursive *precoder*, which could simply be a differential encoder. Henceforth, we will use the term BICOM with *differential precoding* (BICOM-DP) to describe a serially concatenated system whose inner code is a (binary) differential encoder followed by an orthogonal modulator.

In this paper, we derive union bounds on the performance of both BICOM and BICOM-DP. Similar to [8], we explicitly take into account the tail bits used to terminate the trellis of the recursive inner code, which makes the bound more accurate. Since one of the most appealing properties of orthogonal modulation is that it can be noncoherently demodulated, we present bounds for both coherent and noncoherent reception in AWGN and fully-interleaved Rayleigh fading (both with and without knowledge of the fading amplitudes). While the bounds presented in this paper assume ML joint demodulation and decoding, we show that the bounds are a good prediction of the performance of iterative demodulation and decoding by comparing the bounds against simulation results and the BICM-ID error free (EF) bound of [9].

The bounds in this paper are new and distinct from [2] because we allow the inner code to produce M-ary orthogonal outputs rather than mere BPSK symbols. The bounds are also distinct from [6] because the modulator does not have a continuous phase constraint, making it applicable to other forms of orthogonal modulation, such as Walsh-modulation. While the differential encoder imbedded in a CPM modulator operates on M-ary *symbols* [6], the differential encoder considered here operates on the interleaved *bits*, making for a more convenient and less complex demodulator implementation. Also, by removing the constant phase property, we are able to derive a noncoherent demodulator that operates on the trellis of the equivalent inner code, which is in contrast to the suboptimal block-based detector of [6].

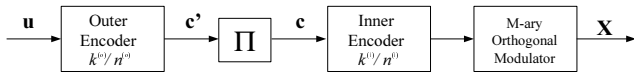


Fig. 1. System model.

## II. SYSTEM MODEL

The system model is as shown in Fig. 1. A binary message sequence  $\mathbf{u}$  of length  $K$  is passed into a rate  $k^{(o)}/n^{(o)}$  outer convolutional encoder, generating a binary sequence  $\mathbf{c}'$  of length  $N$ .  $\mathbf{c}'$  is then bit-wise interleaved, and passed into a rate  $k^{(i)}/n^{(i)}$  inner convolutional encoder. The output of the inner encoder is forwarded to a memoryless orthogonal modulator, which maps each group of  $m = \log_2 M$  consecutive bits to a symbol. The modulated symbols form the matrix  $\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N_x-1}]$ , with each symbol picked from the orthogonal set  $\mathcal{X} = \{\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_{M-1}\}$  of elementary column vectors.

The received signal matrix can then be written as,

$$\mathbf{Y} = \sqrt{\mathcal{E}_s} \mathbf{X} \mathbf{A} + \mathbf{N}, \quad (1)$$

where  $\mathbf{N}$  is a  $M \times N_x$  complex Gaussian noise matrix containing independent zero-mean complex Gaussian noise samples with variance  $N_0/2$  in each complex direction.  $\mathbf{A}$  is a  $N_x \times N_x$  diagonal matrix, whose  $i^{\text{th}}$  diagonal component  $a_i$  is the fading amplitude for symbol  $\mathbf{x}_i$ . In this paper, two types of channel model are considered, (1) AWGN:  $a_i = e^{j\theta_i}$  and the  $\theta_i$ 's are independent and identically distributed (i.i.d.) uniform over  $[0, 2\pi)$ ; and (2) Fully Interleaved Rayleigh Fading: the  $a_i$ 's are i.i.d. zero-mean complex Gaussian with a variance of  $1/2$  in each of the real and imaginary directions. For both channel models, we discuss coherent and noncoherent detection. In Rayleigh fading, noncoherent reception can be either: (1) *With channel state information (CSI)*: the fading amplitude is known to the receiver; or (2) *Without CSI*: the fading amplitude is not known to the receiver.

Because the orthogonal signals are equidistant, performance does not depend on labelling and a gray mapping cannot be used to recover BICOM's capacity loss relative to CM. Instead we propose an iterative receiver, similar to the one proposed in [5], that passes extrinsic information between two SISO processors. The inner SISO processor operates over a merged trellis which describes the inner code and orthogonal modulator, while the outer SISO processor performs soft-output decoding of just the outer code. While the implementation of the outer SISO decoder is quite straightforward, the inner SISO decoder requires that the inner encoder and modulator be merged into a single trellis. When  $m/n^{(i)}$  is an integer, as is the case for a differential inner code, each state-transition in the merged trellis corresponds to one output symbol. The number of states in the merged trellis is the same as the number of states of the inner code, while the number of branches leaving or entering any state is equal to  $M$ . Two examples of merged trellises are shown for  $M = 4$  in Fig. 2, (a) BICOM (inner encoder  $g^{(i)} = 1$ ), and (b) BICOM-DP (inner encoder  $g^{(i)} = 1/(1+D)$ ). For each case, the diagram on the left

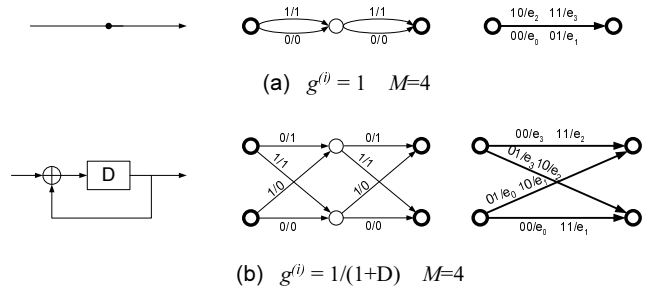


Fig. 2. Trellis merging for the inner code.

depicts the encoder, the diagram in the middle shows the trellis clocked at the bit rate, and the diagram on the right shows the merged trellis clocked at the symbol rate and labeled by  $M$ -ary symbols. Note that in the merged trellis, there are parallel transitions.

## III. UNION BOUND

Because BICOM is uniform, the pairwise error probability (PEP) of the pair  $(\mathbf{X}, \hat{\mathbf{X}})$  is only a function of the number of symbols that  $\mathbf{X}$  and  $\hat{\mathbf{X}}$  differ. Therefore, the union bound can be evaluated using all-zeros codeword as a reference. Under uniform interleaving, the bit error rate (BER) is bounded as [2],

$$P_b \leq \sum_{d=d_{min}}^K \sum_{h=1}^{N_x} \sum_{l=d_{free}^{(o)}}^N \frac{d}{K} \frac{W_{d,l}^{(o)} W_{l,h}^{(i)}}{\binom{N}{l}} P(h), \quad (2)$$

and the frame error rate (FER) as

$$P_f \leq \sum_{d=d_{min}}^K \sum_{h=1}^{N_x} \sum_{l=d_{free}^{(o)}}^N \frac{W_{d,l}^{(o)} W_{l,h}^{(i)}}{\binom{N}{l}} P(h), \quad (3)$$

where  $d_{min}$  is the minimum input weight of the outer code that can generate an error event, and  $d_{free}^{(o)}$  is the minimum free distance of the outer code.  $h$  is the Hamming weight of  $\hat{\mathbf{X}}$ , i.e. the number of symbols  $\hat{\mathbf{X}}$  contains other than  $\mathbf{e}_0$ .  $P(h)$  is PEP of hamming distance  $h$ .  $W_{l,h}^{(i)}$  is the number of inner trellis sequences with input weight  $l$  and output weight  $h$ , and  $W_{d,l}^{(o)}$  is similarly defined for the outer code.

We assume that the outer code is nonrecursive, since it performs no worse than a recursive code. An upper bound on  $W_{d,l}^{(o)}$  is [2],

$$W_{d,l}^{(o)} \leq \sum_{j=1}^{t_{max}^{(o)}} \binom{K}{j} W_{d,l,j}^{(o)}, \quad (4)$$

where  $W_{d,l,j}^{(o)}$  is the number of outer code sequences of input weight  $d$ , output weight  $l$ , and  $j$  adjacent error events [2], and  $t_{max}^{(o)}$  is the maximum number of possible error events. When the inner code is also nonrecursive, the same bound can be applied on the merged inner trellis,

$$W_{l,h}^{(i)} \leq \sum_{j=1}^{t_{max}^{(i)}} \binom{N_x/p}{j} W_{l,h,j}^{(i)}. \quad (5)$$

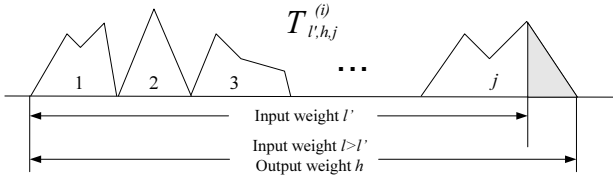


Fig. 3. The calculation of tail terminated error events

The definitions of  $W_{l,h,j}^{(i)}$  and  $t_{max}^{(i)}$  are similar to the ones of outer code.

*Example 1* The conventional BICOM system has a trivial one state rate 1 inner code as shown in Fig. 2(a) for  $M = 4$ . All the edges except the all-zeros edge contain errors. Let  $W^{(i)}(L, H, j)$  be the input-output weight enumerating function (IOWEF) with  $j$  adjacent error events. It can be calculated as

$$W^{(i)}(L, H, 1) = ((L+1)^m - 1)H, \quad (6)$$

$$W^{(i)}(L, H, j) = W^{(i)}(L, H, 1)^j. \quad (7)$$

It is obvious that the maximum number of error events is  $t_{max}^{(i)} = N/m$ , which is the total number of inner trellis stages. Therefore,

$$\begin{aligned} W^{(i)}(L, H) &= \sum_{j=1}^{N/m} \binom{N/m}{j} W(L, H, j) \\ &= [((L+1)^m - 1)H + 1]^{N/m} - 1. \end{aligned} \quad (8)$$

Note that (5) is satisfied with equality in this case, and  $W_{l,h}^{(i)}$  is just the coefficient of the term  $L^l H^h$ . ■

While a nonrecursive, constraint length  $\nu$  encoder can always be terminated with tail of  $\nu - 1$  zeros, the same cannot be said for a recursive encoder. Rather, the tail bits produced in a recursive encoder have a positive weight and complement the final error event. When the inner code is recursive, we define the term  $T_{l',h,j}^{(i)}$  as the number of sequences with the following properties: (1) The input weight is  $l'$  up to the last  $\nu - 1$  sub-trellis stages (where the sub-trellis is that of the inner code by itself); (2) The output weight is  $h$ ; (3) The number of adjacent error events is  $j$ ; and (4) The total input weight is  $l > l'$ . The last property implies that the tail bits have a positive weight. Fig. 3 illustrates the difference between  $T_{l',h,j}^{(i)}$  and  $W_{l,h,j}^{(i)}$ . While the  $j$  error events of  $W_{l,h,j}^{(i)}$  are all allowed to be arranged among all the trellis stages, only  $j - 1$  error events of  $T_{l',h,j}^{(i)}$  are free to move, since the last error event is fixed at the tail. As a result, for the recursive inner code, (5) is revised as <sup>1</sup>

$$W_{l,h}^{(i)} \leq \sum_{j=1}^{t_{max}^{(i)}} \binom{N_x/p}{j} W_{l,h,j}^{(i)} + \sum_{j=0}^{t_{max}^{(i)}-1} \binom{N_x/p}{j} T_{l,h,j+1}^{(i)}. \quad (9)$$

*Example 2* A BICOM-DP system has a differential inner encoder  $g^{(i)} = 1/(1+D)$ . As shown in Fig.2(b) for  $M = 4$ , the merged inner trellis has two states and  $M/2$  parallel edges

<sup>1</sup>For simplicity, we neglect the fractional rate loss due to tail bits.

between every distinct starting and ending state. Without tail bits, only even weight inputs can generate error events, i.e.  $W_{l,h,j}^{(i)} = 0$  for all odd  $l$ . If terminated by a tail bit, the tail is non-zero only for odd weight inputs, i.e.  $T_{l,h,j}^{(i)} = 0$  for all even  $l$ . Let  $T^{(i)}(L, H, j)$  be the tail termination IOWEF with  $j$  adjacent error events. When  $M = 4$ ,

$$\begin{aligned} W^{(i)}(L, H, 1) &= 3L^2H + 4L^2H^2 + 4L^2H^3 + 4L^2H^4 \\ &\quad + 2L^4H^2 + 6L^4H^3 + 10L^4H^4 \dots (10) \end{aligned}$$

$$\begin{aligned} T^{(i)}(L, H, 1) &= 2LH + 2LH^2 + 2LH^3 + 2LH^4 \\ &\quad + L^3H^2 + 3L^3H^3 + 5L^3H^4 \dots (11) \end{aligned}$$

and the concatenation of the error events is equivalent to the product of the IOWEFs,

$$W^{(i)}(L, H, j) = W^{(i)}(L, H, 1)^j, \quad (12)$$

$$T^{(i)}(L, H, j) = W^{(i)}(L, H, 1)^{j-1} T^{(i)}(L, H, 1). \quad (13)$$

Using the above results,  $W_{l,h}^{(i)}$  can be easily found from (9). ■

It is easy to determine the interleaver gain with a recursive inner code by looking into the exponents of  $N$  on the FER and BER bounds [2],

$$\alpha_f = t^{(o)} + t^{(i)} - l, \quad (14)$$

$$\alpha_b = t^{(o)} + t^{(i)} - l - 1, \quad (15)$$

where  $t^{(o)}$  and  $t^{(i)}$  are the number of error events for the outer and inner trellises, respectively. For a recursive inner code,  $\alpha_{f,max} = -\lfloor \frac{d_{free}^{(o)} - 1}{2} \rfloor$ , and there is an FER interleaver gain in most cases. Here, we look into the minimum output weight associated with the maximum exponent,  $h_{min}(\alpha_{f,max})$ . When  $N$  is large, it can be proved that this is the diversity gain in fully interleaved Rayleigh fading channel. When  $d_{free}^{(o)}$  is even,  $h_{min}(\alpha_{f,max})$  is written as

$$h_{min}(\alpha_{f,max}) = \frac{d_{free}^{(o)} d_{2,free}^{(i)}}{2}, \quad (16)$$

where  $d_{l,free}^{(i)}$  is the minimum output weight of the merged inner trellis with input weight  $l$  (without tail termination). If  $d_{free}^{(o)}$  is odd, the maximum exponent of  $\alpha_{f,max} = -\frac{d_{free}^{(o)} - 1}{2}$  is achieved when one of the following three situations occurs: (1)  $d_{free}^{(o)} - 1$  bits generate  $(d_{free}^{(o)} - 1)/2$  error events each with input weight two, while the combination of the remaining bit and the tail bit generates one more error event; (2) There are  $(d_{free}^{(o)} - 3)/2$  error events with input weight 2 and one event with input weight 3; or (3) The minimum even Hamming distance of the outer code is  $d_{even}^{(o)} = d_{free}^{(o)} + 1$ , and there are  $d_{even}^{(o)}/2$  error events with input weight 2. We have,

$$\begin{aligned} h_{min}(\alpha_{f,max}) &= \min \left\{ \frac{(d_{free}^{(o)} - 1)d_{2,free}^{(i)}}{2} + d_{1,t}^{(i)}, \right. \\ &\quad \left. \frac{(d_{free}^{(o)} - 3)d_{2,free}^{(i)}}{2} + d_{3,free}^{(i)}, \right. \\ &\quad \left. \frac{(d_{free}^{(o)} + 1)d_{2,free}^{(i)}}{2} \Big|_{d_{even}^{(o)} = d_{free}^{(o)} + 1} \right\}, \end{aligned} \quad (17)$$

where  $d_{1,t}^{(i)}$  is the minimum weight generated by a weight 1 input together with the tail. The notation  $A|_B$  means  $A$  takes its value only if  $B$  is satisfied, otherwise  $A = +\infty$ .

#### IV. PAIRWISE ERROR PROBABILITY

In this section, we evaluate the PEP of both coherent and noncoherent detection in AWGN and fully interleaved Rayleigh fading. It is obvious that the PEP is independent of the modulation alphabet size  $M$ , as long as the symbol-wise SNR  $\gamma = \mathcal{E}_s/N_0$  or  $\bar{\gamma}$  keeps constant. In the following, we calculate the PEP over two length  $h$  sequences  $\mathbf{X}$  and  $\bar{\mathbf{X}}$ , which are composed of all  $\mathbf{e}_0$ 's and all  $\mathbf{e}_1$ 's respectively.

##### A. Coherent Detection

1) *AWGN Channel*: In AWGN, the PEP is just,

$$P(h) = Q\left(\sqrt{h\gamma}\right) \leq e^{-h\gamma/2}. \quad (18)$$

2) *Rayleigh Channel*: When the channel is fully interleaved Rayleigh fading, the PEP is the expectation of the conditional PEP taken over the fading coefficients  $a_i$ 's. The closed form solution to this can be found in [10](14.4-15)(14.4-21),

$$P(h) = \left(\frac{1-\mu}{2}\right)^h \sum_{i=0}^{h-1} \binom{h-1+i}{i} \left(\frac{1+\mu}{2}\right)^i, \quad (19)$$

where

$$\mu = \sqrt{\frac{\bar{\gamma}}{2+\bar{\gamma}}}. \quad (20)$$

##### B. Noncoherent Detection

1) *AWGN Channel*: We first form the LLR  $\Delta$ , which can be written in the summation of individual LLRs,

$$\Delta = \sum_{i=0}^{h-1} [\log p(\mathbf{y}_i|\mathbf{e}_0) - \log p(\mathbf{y}_i|\mathbf{e}_1)]. \quad (21)$$

The PEP is just the probability that the LLRs sum to less than zero. Unfortunately, there is no closed form solution due to the nonlinear form of the noncoherent metric. In order to find a numerical solution, we first take Laplace transform over the probability density function (pdf) of  $\Delta$  [1],

$$\begin{aligned} \Phi_{\Delta}(s) &= E_{\mathbf{y}} \left[ \left( \frac{p(\mathbf{y}|\mathbf{e}_0)}{p(\mathbf{y}|\mathbf{e}_1)} \right)^{-s} \right]^h \\ &= \left( E_{\mathbf{y}[0]} \left[ I_0 \left( \frac{2\sqrt{\mathcal{E}_s} |\mathbf{y}[0]|}{N_0} \right)^{-s} \right] \right)^h \\ &\quad \cdot \left( E_{\mathbf{y}[1]} \left[ I_0 \left( \frac{2\sqrt{\mathcal{E}_s} |\mathbf{y}[1]|}{N_0} \right)^s \right] \right)^h, \end{aligned} \quad (22)$$

where  $\mathbf{y}[i]$  represents the  $i^{\text{th}}$  element in vector  $\mathbf{y}$ . Then the PEP is found using

$$\begin{aligned} P(h) &= P(\Delta \leq 0) \\ &= \frac{1}{2\pi j} \int_{\delta-j\infty}^{\delta+j\infty} \frac{\Phi_{\Delta}(s)}{s} ds. \end{aligned} \quad (23)$$

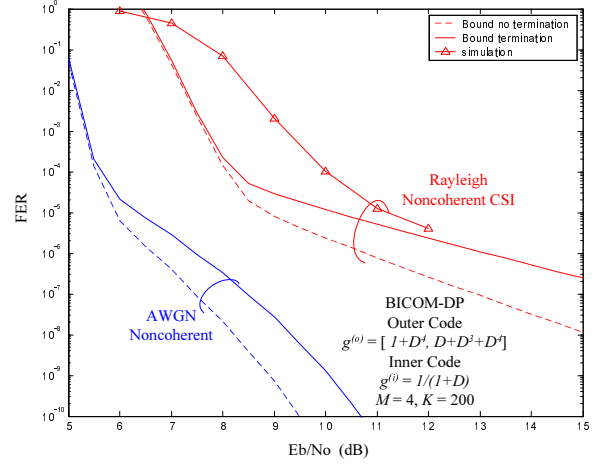


Fig. 4. The effect of tail termination on the bound.

The integral (23) can be evaluated using Gauss-Chebyshev quadratures, suggested in [1] [11]. An easy proof (not included in this paper) shows that  $\Phi_{\Delta}(\delta)$  achieves its minimum value at  $\delta = 1/2$ , which offers the best convergence rate [11]. Although Monte Carlo simulation is needed to evaluate the two expectations in the final step of (22), it is much easier than simulating the whole system.

Another option here is to use an upper bound on the PEP, which comes from the suboptimal square law detection. A closed form of error probability can be found in [12],

$$P(h) \leq \frac{1}{2^{(2h-1)}} e^{-\frac{\gamma}{2}} \sum_{i=0}^{h-1} \left[ \frac{1}{i!} \left( \frac{\gamma}{2} \right)^i \sum_{j=0}^{h-1-i} \binom{2h-1}{j} \right]. \quad (24)$$

2) *Rayleigh Channel with CSI*: When the fading amplitude is known to the receiver, we can still use the method in (22) and (23) to find the PEP. The only difference lies in (22), where  $\mathbf{y}[0]$  and  $\mathbf{y}[1]$  are both dependent on the fading coefficient, and thus are not independent any more. The Laplace transform is written as

$$\Phi_{\Delta}(s) = \left( E_{a,\mathbf{y}} \left[ \left( \frac{I_0 \left( \frac{2\sqrt{\mathcal{E}_s} |a| |\mathbf{y}[1]|}{N_0} \right)^s}{I_0 \left( \frac{2\sqrt{\mathcal{E}_s} |a| |\mathbf{y}[0]|}{N_0} \right)^s} \right) \right] \right)^h, \quad (25)$$

and the same step in (23) can be applied with  $\delta = 1/2$ .

3) *Rayleigh Channel with no CSI*: When neither the fading amplitude nor phase is known, the optimal metric is [4],

$$\log p(\mathbf{y}|\mathbf{e}_i) = \frac{\mathcal{E}_s |\mathbf{y}[i]|^2}{N_0(N_0 + \mathcal{E}_s)} \quad i = 0, 1 \quad (26)$$

This is equivalent to a square law demodulator, and the PEP has the same form as in (19) [10](14.4-30), except that

$$\mu = \frac{\bar{\gamma}}{2 + \bar{\gamma}}. \quad (27)$$

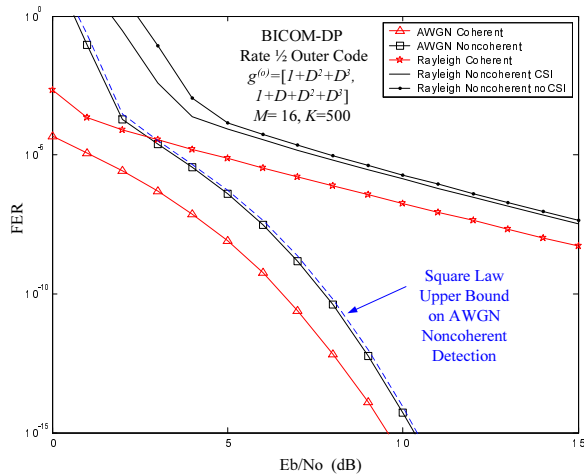


Fig. 6. Bounds for all five receiver/channel combinations.

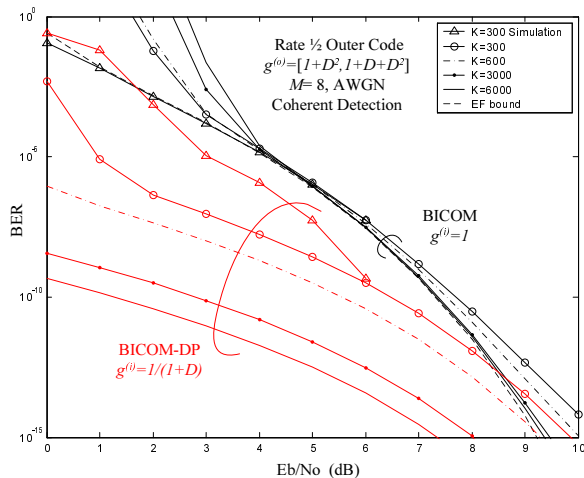


Fig. 5. The interleaver gain of BICOM-DP compared with BICOM.

## V. RESULTS AND ANALYSIS

This section compares simulation results (with 20 iterations of decoding) against the bounds. Fig. 4 shows results for BICOM-DP with noncoherent reception. The outer code is rate 1/2 with  $g^{(o)} = [1 + D^4, D^2 + D^3 + D^4]$  and  $M = 4$  orthogonal modulation. The outer code has odd minimum distance  $d_{odd}^{(o)} = 7$ , and even minimum distance  $d_{even}^{(o)} = 10$ . Also, from Example 2,  $d_{1,t}^{(i)} = d_{2,free}^{(i)} = 1$  and  $d_{3,free}^{(i)} = 0$ . Without tail bits, no odd weight input to the inner code will generate an error event. We get  $h_{min}(\alpha_{f,max}) = 5$ , which comes from (16). However, using a tail can add an additional 1 to the minimum odd input, so  $h_{min}(\alpha_{f,max})$  is actually 4 instead of 5, which is the diversity gain in fading.

Fig. 5 compares BICOM with BICOM-DP, both using outer code  $g^{(o)} = [1 + D^2, 1 + D + D^2]$ . For BICOM,  $\alpha_{b,max} = t_{max}^{(o)} - 1$  from (15). When the SNR is reasonably high, the sequence with only 1 error event in the outer code occurs

with much greater frequency than the other error events. Thus, the effective  $\alpha_b$  is close to zero, which causes BICOM's BER bound to converge to the the EF bound [9]. However, for BICOM-DP, the BER decreases by about  $10^{-3}$  when  $K$  increases 10 times. This verifies the maximum exponent of  $N$  on BER  $-\lfloor \frac{d_{free}^{(o)} + 1}{2} \rfloor = -3$ , where  $d_{free}^{(o)} = 5$  for the 4-state outer code.

Fig. 6 shows the bounds on BICOM-DP for all five channel detection types with outer code  $g^{(o)} = [1 + D^2 + D^3, 1 + D + D^2 + D^3]$ . It is seen that the bounds for AWGN channel go down exponentially, while the bounds in fully independent Rayleigh channel are asymptotically straight, with diversity gain  $h_{min}(\alpha_{f,max}) = 3$ , since the minimum free distance of the outer code is 6.

## VI. CONCLUSIONS

This paper presents a union bound on bit interleaved coded orthogonal modulation. The bound includes the effect of terminating the differential encoder with a tail. The recursive inner code is shown to have the interleaver gain against the nonrecursive code. Both coherent and noncoherent reception are evaluated, in AWGN and fully interleaved Rayleigh fading channel. The diversity gain in the Rayleigh fading channel is also evaluated.

## REFERENCES

- [1] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.
- [2] S. Benedetto, D. Divsalar, D. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 909–926, 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. C. Valenti and S. Cheng, "Iterative demodulation and decoding of turbo coded  $M$ -ary noncoherent orthogonal modulation," *IEEE J. Select. Areas Commun.*, vol. 23, Sept. 2005.
- [5] K. R. Narayanan and G. L. Stüber, "A serial concatenation approach to iterative demodulation and decoding," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 956–961, Jul. 1999.
- [6] K. Narayanan and G. Stuber, "Performance of trellis-coded CPM with iterative demodulation and decoding," *IEEE Trans. Commun.*, vol. 49, pp. 676–687, Apr. 2001.
- [7] P. Hoehner and J. Lodge, "Turbo DPSK: Iterative differential PSK demodulation and channel decoding," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 837–843, June 1999.
- [8] F. Fagnani and F. Garin, "Analysis of serial concatenation schemes for non-binary modulations," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, Adelaide, Australia, Sept. 2005, pp. 745–749.
- [9] A. Chindapol and J. A. Ritcey, "Design, analysis, and performance evaluation of BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 944–957, May 2001.
- [10] J. Proakis, *Digital Communications*, 4th ed. New York, NY: McGraw-Hill, Inc., 2001.
- [11] E. Biglieri, G. Caire, G. Taricco, and J. Ventura, "Simple method for evaluating error probabilities," *IEE Electronics Letters*, vol. 32, pp. 191–192, Feb. 1996.
- [12] M. K. Simon and M. S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. IEEE*, vol. 86, no. 9, pp. 1860–1877, Sept. 1998.