# On Interleaved, Differentially Encoded Convolutional Codes

Michael Peleg, Igal Sason, Shlomo Shamai (Shitz) and Avner Elia

Department of Electrical Engineering

Technion - Israel Institute of Technology

Haifa 32000, Israel

June 1999

**Abstract**

We study a serially interleaved concatenated code construction, where the outer code is a standard convolutional code, and the inner code is a recursive convolutional code of rate 1. Focus is put on the ubiquitous inner differential encoder (used in particular to resolve phase ambiguities), double differential encoder (used to resolve both phase and frequency ambiguities), and another rate 1 recursive convolutional code of memory 2. We substantiate analytically the rather surprising result, that the error probabilities corresponding to a maximum likelihood (ML) coherently detected antipodal modulation over the AWGN channel, for this construction are advantageous as compared to the stand-alone outer convolutional code. This is in-spite of the fact that the inner code is of rate 1. The analysis is based on the tangential sphere upper bound of a ML decoder, incorporating the ensemble weight distribution (WD) of the concatenated code, where the ensemble is generated by all random and uniform interleavers. This surprising result is attributed to the WD thinning observed for the concatenated scheme which shapes the WD of the outer convolutional code to resemble more closely the binomial distribution (typical of a fully random code of the same length and rate). This gain is maintained regardless of a rather dramatic decrease, as demonstrated here, in the minimum distance of the concatenated scheme as compared to the minimum distance of the outer stand-alone convolutional code. The advantage of the examined serially concatenated code given in terms of bit and/or block error probability decoded by a practical suboptimal decoder, over optimally decoded standard convolutional code is demonstrated by simulations, and some insights into the performance of the iterative decoding algorithm are also discussed. Though we have investigated only specific constructions of constituent inner (rate 1) and outer codes, we trust, hinging on the rational of the arguments here, that these results extend to many other constituent convolutional outer codes and rate 1 inner recursive convolutional codes. Union bounds on the performance of serial and hybrid concatenated codes were addressed in [8], where differential encoding was also examined, and shown efficient.

**Keywords:** differential encoding, distance spectrum, error bounds, iterative decoding, serial concatenation.

This work has been supported by the Samuel Neaman Institute for Advanced Studies in Science and Technology of the Technion.

# I. Introduction

Differential encoding (DE) is widely used to enable non-coherent detection when carrier phase acquisition and tracking are impossible, unreliable or prohibitively complex due to excessive phase noise. This situation is commonly encountered where noise and errors of local oscillators are significant and where short messages are transmitted such as typical in packet communication and in frequency hopping systems. DE usually impairs the performance of the system, especially when combined with error correcting codes. Various methods were developed to mitigate this impairment to some degree in [6],[16],[26], and references therein. Interleaving is often used to combat burst noise and fading.

A transmitter of a system utilizing these widely used concepts is depicted in the upper part of Fig. 1. It comprises an error correcting convolutional code, interleaver, inner encoder such as the DE and BPSK modulator. Some recent reports [1],[12],[13],[15],[17],[21],[22], recognized this structure as an interleaved serial concatenation of two codes considering the DE as an inner code, and applied appropriate iterative decoding method developed for the detection of interleaved serially concatenated codes, see [3],[4] and references therein. Some of those works have put in evidence not only mitigation of the impairment caused by DE and non-coherent detection but reported performance exceeding somewhat that of a convolutionally encoded and coherently decoded system.

In this work we examine the capabilities of a serially interleaved concatenation of a convolutional code and DE in a coherent setting. We develop an ensemble upper bound of Block Error Rate (BLOER) using a novel bounding method [27] and demonstrate the superiority of the concatenated scheme over the standard convolutional code when both are maximum likelihood decoded. The ensemble bound comprises averaging over the random uniform interleaver. A less tight union bound for a very similar code was presented in [8]. We show by simulations that the iterative decoding technique, while clearly suboptimal, still outperforms significantly in terms of Bit Error Rate (BER) and BLOER the standard convolutionally coded and optimally (in terms of BER or BLOER) decoded system.

# II. System Description

The system under consideration is depicted in Fig. 1.

**A. Transmitter and channel:** A sequence $\mathbf{b}$ of $N_b$ uniformly and independently distributed information bits $b_i \in (0, 1)$ is encoded by an outer non-recursive convolutional Error Correction Code (ECC) resulting in sequence $\mathbf{c}$ of $N_c$ coded bits $c_i$. We use $rate = 0.5$, 16 states code

with generators (23,35) (octal) and memory $m = 4$ (see Fig. 2e). The code is terminated by appending $m = 4$ zero-bits to the sequence $\mathbf{b}$. The coded bits $c_i$ are interleaved by a bit-wise interleaver generated randomly and uniformly [3] for each sequence of $N_c$ coded bits which produces a permuted sequence of bits $c_j$. We use the same symbols with modified indices for interleaved and non-interleaved signals. This type of random interleaving is a probabilistic structure that takes into consideration all the possible interleaving permutations, with all interleaving configurations equally weighted. The interleaved sequence is encoded by a $rate = 1$ inner encoder which produces the sequence $\mathbf{d}$ of bits $d_j$. The most practical inner code is the standard DE which can be represented as a recursive convolutional code of $rate = 1$ with generators 3 (octal) for feedback and 2 for the output sequence generation, see Fig. 2a. We examine also two similar codes but with generators (7, 4) and (5, 1) denoted respectively D3 and DDE, see Fig. 2. The double differential encoder (DDE) known also as the second order phase-difference modulator is robust to frequency offsets when used in conjunction with an appropriate noncoherent detector [20]. All the inner codes are terminated to the zero state by a one or two additional bits for the DE and D3/DDE codes respectively. A system with no inner code, i.e. using $d_j = c_j$ (the outer stand-alone convolutional code) was also examined for reference. The bits $d_j$ are mapped to BPSK symbols $s_j$, $s_j = 2d_j - 1$ and transmitted over a coherent AWGN channel, the outputs of which (sampled at the output of a matched filter) are:

$$r_j = s_j + n_j \ , \tag{1}$$

where $n_j$ is the in-phase component of AWGN noise the variance of which is $\mathrm{Var}(n_j) = \sigma^2$. Since the overall coding rate $R_c$ is 0.5 we have $E_b/N_o = \frac{1}{R_c} \cdot E_s/N_o = 1/\sigma^2$, where $E_s/N_o = \frac{1}{2\sigma^2}$ is the channel symbol energy over the noise spectral density ratio.

In our analysis and simulations we used the following sequence lengths: $N_c = 1000$, $N_b = \frac{N_c}{2} - m = 496$, $N_d$, the number of bits $d_j$, is 1001 and 1002 for the DE and D3/DDE codes respectively.

**B. Optimal receivers:** The optimal receiver, minimizing the probability of bit error, computes for each bit $b_i$ the a posteriori probability

$$p(b_i = 1|\mathbf{r}) = \sum_{\mathbf{b}, b_i=1} p[\mathbf{r}|\mathbf{s}(\mathbf{b})] \Big/ \sum_{\mathbf{b}} p[\mathbf{r}|\mathbf{s}(\mathbf{b})] \tag{2}$$

where $\mathbf{r}$ is the sequence of $N_d$ channel outputs $r_j$. The memoryless channel model yields:

$$p[\mathbf{r}|\mathbf{s}(\mathbf{b})] = \prod_{i=1}^{N_d} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-(s_j - r_j)^2 / 2\sigma^2\right] \tag{3}$$

A different maximum likelihood optimal receiver which minimizes the BLOER, chooses $\mathbf{b}$ to maximize $p[\mathbf{r}|\mathbf{s}(\mathbf{b})]$ in (3).

The optimal receivers for the concatenated codes are prohibitively complex due to the large number of the possible sequences **b** and of the corresponding codewords **s** for which eq. (3) needs to be evaluated, giving thus motivation for suboptimal reduced complexity efficient decoders.

**C. A practical receiver:** A high performance suboptimal practical receiver for two serially concatenated convolutional codes separated by an interleaver was presented in [3]. It utilizes the iterative decoding procedure introduced in [5] for turbo codes. We resort to the iterative receiver of [3], changing only the code generators and the code-rate, see Fig. 1, operating as follows:

The channel outputs are fed into the inner decoder which produces soft metrics $\lambda_j$ as the Aposteriory Probabilities (AP) of the coded bits $c_j$ based on the structure of the inner code and on the side information $\hat{P}(c_j)$ on the same bits provided by the outer decoder on previous iterations. The outer decoder uses the deinterleaved sequence of $\lambda_i$ to produce information bit estimates $\hat{b}_i$ and the soft outputs $\hat{P}(c_i)$, which are the AP of the bits $c_i$ based on the sequence $\lambda_i$ and on the structure of the outer code, to be used by the inner decoder on the next iteration. Both the inner and the outer decoders are Soft Input Soft Output (SISO) decoders, see [3], [4] for more details.

## III. Analysis

In this section, we state the underlying assumptions on which our bounds are based, introduce notations and basic relations from [3],[18][24],[27],[28],[30], which apply to serially concatenated codes.

**A. Assumptions:** We assume here ML decoding and that the results for large values of interleaver length ($N_c \gg m$) are insensitive to a termination method of the inner code.

**B. Weight distribution:** The number of codewords of a serially concatenated code $c_s$ that are encoded by information bits of weight $w$ and having a total Hamming weight of $h$ is designated by $A_{w,h}^{c_s}$.

For a serially concatenated code $c_s$ with a uniform interleaving situated between the outer and inner codes ($c_{out}$ and $c_{in}$ respectively) and operating on $N_c$ bits, the following equation holds ([3],[30]) under the termination assumptions mentioned above:

$$A_{w,h}^{c_s} = \sum_{\ell=0}^{N_c} \frac{A_{w,\ell}^{c_{out}} A_{\ell,h}^{c_{in}}}{\binom{N_c}{\ell}} \quad 0 \le w \le \frac{N_c}{2} - m\,, \quad 0 \le h \le N_c\,, \tag{4}$$

where $m$ is the memory length of the outer code ($m = 4$ in the considered case) and $N_c/2 - m$ is the length of the information sequence in our case (since the outer code is of rate 1/2). Clearly,

by linearity, a zero input results in a zero output and therefore $A_{0,0}^{c_s} = 1$. As indicated, the block length of the serial concatenation $c_s$ is $N_c$ bits + termination (the same as the interleaver length, since the inner codes are assumed to have a unity rate). If we look at the set of codewords produced by the outer encoder $c_{out}$ as a block code of length $N_c$ and designating its weight distribution by $\{S_\ell^{c_{out}}\}_{\ell=0}^{N_c}$, then the following equation holds:

$$S_\ell^{c_{out}} = \sum_{w=0}^{\frac{N_c}{2}-m} A_{w,\ell}^{c_{out}} \qquad 0 \le \ell \le N_c \; . \tag{5}$$

The Hamming weight of the termination bits is ignored and the number of all the codewords with an overall Hamming weight of $\ell$ are counted. By linearity, the summation over $w$ begins with $w = 1$ for any non-zero value of $\ell$ and also $S_o^{c_{out}} = 1$. Also, a similar relation as in equation (5) holds for the weight distribution of the code $c_s$:

$$S_h^{c_s} = \sum_{w=0}^{\frac{N_c}{2}-m} A_{w,h}^{c_s} \qquad 0 \le h \le N_c \; . \tag{6}$$

By combining equations (4),(5),(6), the following equation is derived for the weight distribution of the serially concatenated code $c_s$:

$$S_h^{c_s} = \sum_{\ell=0}^{N_c} \frac{S_\ell^{c_{out}} A_{\ell,h}^{c_{in}}}{\binom{N_c}{\ell}} \quad 0 \le h \le N_c \tag{7}$$

Equation (7) was used for the calculation of the weight distribution of the two options of serially concatenated codes (with the aid of the calculations presented in appendices I,II,III). For evaluating equation (7), it is required first to calculate the output-weight distribution of the outer code and also the input-output weight distribution of the inner code. Then the weight distribution of the code (treated as a block code of length $N_d$) is used to calculate the tangential sphere upper bound on the block error probability with ML decoding.

In our setting the ratio of the BER to BLOER designated by RE = BER/BLOER is rather high (see section IV-B and Fig. 5) for low and moderate $E_b/N_o$, implying that an erroneous block inflicts many bit errors and therefore the upper bound on BLOER serves as a reasonable upper bound on BER. A tighter bound on BER can be developed as in [27], but this requires the input-output weight distribution of the considered outer code which make this evaluation rather hard in our case. For the BLOER, only the output weight distribution of the outer code is needed and that is calculated (see Appendix I) by using equation (7) yields with the tangential sphere bounding technique [27] an upper bound on the BLOER of the considered, serially concatenated scheme.

**C. Union bound:** The most familiar upper bound that is based on the weight distribution of the code is the union bound ([10],[24],[28]). However, for long and complex codes this bound usually

diverges in the rate region above the cutoff rate, and therefore yields useless results for efficient code configurations like the serial or parallel concatenated turbo schemes [28].

**D. Tangential sphere bound:** The tangential sphere bound applied in the analysis here, is an improved weight distribution based upper bound on the ML decoding block error probability ([24],[27]). The mathematical derivation of this bound is omitted here, see the references mentioned for details. For a variety of block codes, this upper bound is tighter than some other known upper bounds, especially for moderate and low values of $\left(\frac{E_b}{N_o}\right)$, see [10].

**E. Double differential encoder (DDE):** In this subsection we shall present a certain equivalence of the DDE to DE in our concatenated uniformly interleaved system using BPSK modulation and coherent detection.

For DDE the following relation holds [20]:

$$\begin{aligned} \pi \cdot c_n &= ((\phi_n - \phi_{n-1}) - (\phi_{n-1} - \phi_{n-2})) \bmod (2\pi) \\ &= (\phi_n - 2\phi_{n-1} + \phi_{n-2}) \bmod (2\pi) \ , \end{aligned}$$

when $c$ and $\phi$ are the input bit and the corresponding phase of the BPSK modulated signal respectively. Since the phase $\phi$ may be 0 or $\pi$ for a BPSK modulated signal, then the term '$2\phi_{n-1}$' has no influence, $((2\phi_{n-1} \bmod (2\pi) \equiv 0)$, and therefore

$$\pi \cdot c_n = (\phi_n + \phi_{n-2}) \bmod (2\pi) \ .$$

For a BPSK signal, it is equivalent to the relation $c_n = d_n \oplus d_{n-2}$. To reduce the sensitivity of a receiver to a shift in the carrier frequency, the demodulator should take into account $r_{n-1}$ and not only $r_n, r_{n-2}$ [20], however, this problem does not rise in coherent communication. Therefore, for a BPSK modulated signal, the DDE is equivalent to the encoder of the code D31 (see Fig. 2d) described by $d_n = c_n \oplus d_{n-2}$ .

We denote by $\mathbf{c}_o, \mathbf{c}_e, \mathbf{d}_o, \mathbf{d}_e$, the sequences comprising the odd and even indexed bits $c_i$ and $d_i$ respectively. Then it is clear from the last equation that $\mathbf{d}_o$ and $\mathbf{d}_e$ are the standard differentially encoded $\mathbf{c}_o$ and $\mathbf{c}_e$ respectively. To demonstrate exact equivalence, we initialize the encoder used to encode $\mathbf{c}_e$ to the last state of the encoder used to encode $\mathbf{c}_o$. This data dependent initialization, equivalent to a similar initialization of the DDE encoder, is insignificant for large block lengths and enables encoding the whole sequence $[\mathbf{c}_o, \mathbf{c}_e]$ into the sequence $[\mathbf{d}_o, \mathbf{d}_e]$ by a single path through a DE. A simple permutation of $[\mathbf{d}_o, \mathbf{d}_e]$ yields $\mathbf{d}$, the DDE output. Thus the DE and DDE encoders are equivalent except for the mapping $\mathbf{c} \rightarrow [\mathbf{c}_o, \mathbf{c}_e]$ which is absorbed by the uniform interleaver and the mapping $[\mathbf{d}_o, \mathbf{d}_e] \rightarrow \mathbf{d}$, which is a weight conserving permutation. Therefore the two codes have identical weight distribution under the data dependent initialization.

To demonstrate that the performance of the DE and DDE coded systems are similar also when decoded iteratively, we will show the performance is identical with the data dependent initialization of the DDE described above. It is easy to construct a trellis based optimal MAP decoder for the DDE inner code by the standard method, however we shall examine other equivalent form as follows:

a. The received sequence $\mathbf{r}$ is divided into two subsequences of odd and even indexed symbols and a new received sequence is formed by appending the even indexed subsequence to the odd one. This invertible operation preserves all the information available in the received signal.

b. The resulting received sequence is the sequence $[\mathbf{d}_o, \mathbf{d}_e]$ of interleaved bits $d_n$ described above, at the channel output, therefore it may be optimally detected by a DE trellis based decoder. Since the inner decoder, which is now identical to that of DE, operates on data obtained identically as in the DE case (i.e. ECC encoding, uniform interleaving and DE) and the outer encoder and decoder are identical too, the DE and the DDE systems will achieve the same performance with the initialization described above.

We conclude thus that all results, bounds and simulations, for DE and DDE are equivalent and therefore only results for DE are presented.

## IV. Results

**A. Upper bounds:** We report here the results of the upper bounds on the block error probabilities for two serially concatenated codes and compare them to the upper bound on the block error probability of the outer stand-alone convolutional codes. For the examined options, the concatenated codes are of rate $\frac{1}{2}$ and a block consists $N_c = 1000$ coded bits. The weight distribution of all the considered concatenated codes are calculated (appendices I,II,III) and then the tangential sphere bound is applied to provide upper bounds on the ML block error probability. The weight distribution of the considered concatenated codes are also compared to the weight distribution of a random block code (with the same code length and rate) which is binomially distributed.

The reduction of the minimal distance of the serially concatenated codes due to the existence of the random interleaver (of length $N_c$) and the corresponding inner code for each option is demonstrated in Fig. 3. For the outer stand-alone code and a block of $N_c = 1000$ bits, $d_{min} = 6$ and for the two serially concatenated codes: $d_{min} = 3$ (where the inner code is DE) and $d_{min} = 2$ (where the inner code is D3). This decrease in the minimum distance is evident as there exist interleavers which group the 6 'ones' at the output of the convolutional code into three pairs of ones, yielding at the DE output the minimum distance of 3. However, in spite of the fact that the minimal distances of the two serial concatenated codes are smaller than the minimal distance of the

outer stand-alone code (of the same block length), the multiplicity (number of codewords with the same Hamming weight) of the near neighbors is decreased considerably and even more pronounced, each of the multiplicities associated with Hamming weights between $d_{min}$ and $\frac{N_c}{2}$ (half of the code length) are decreased, as a consequence of the random interleaver and the serial concatenation with an inner recursive convolutional code. This phenomena named 'spectral thinning' was reported also in [23] as a weight distribution interpretation for the outstanding ML performance of turbo codes at moderate to low values of $E_b/N_o$. The irrelevance of the minimum distance for good low $E_b/N_o$ performance was advocated by Battail [2], and references therein, and this feature is confirmed also in our case.

The second effect of the interleaved serially concatenated codes is the increase of $d_{max}$ of the corresponding block codes. For the case of the outer stand-alone code, $d_{max} = 813$ (for a block of length $N_c = 1000$ bits). However, when the outer code is randomly interleaved and then is serially concatenated with the differential encoder, $d_{max}$ is increased to the value of $N_c - 1 = 999$ (because the initial state of the encoder is '0' and therefore the first output is also zero). When the outer code is randomly interleaved and then is serially concatenated with the inner code D3, $d_{max}$ is increased from 813 to 965. Therefore, the random interleaver combined with any of the recursive rate 1 convolutional codes as an inner code, decreases $d_{min}$ and increases $d_{max}$ of the outer code and also causes the phenomena mentioned above.

These observations based on the weight distribution of the examined codes provide the insight for the reason that *sufficiently large random interleaver* combined with a serial concatenation of a recursive *rate 1* inner code improves the overall performance of the coded system even in *coherent detection* and ML decoding (and this is in spite of the rather dramatic decrease of $d_{min}$). The improvement is attributed to the closer resemblance of the weight distribution of the serially concatenated code to the binomial distribution [2], as described in Fig. 3. After introducing the weight distribution of the examined codes in Fig. 3, the corresponding upper bounds on the block error probability are presented in Fig. 4. As can be observed from the solid lines of Fig. 4, the improvement of the upper bounds on the block error probability of the serially concatenated codes (as compared to the outer stand-alone code) is rather dramatic. For a block error probability of $10^{-2}$, the gain of $E_b/N_o$ achieved by the interleaved serially concatenated code with DE as an inner code is 3.8dB, and if the inner code is replaced by D3 (Fig. 2b), there is another slight improvement with ML decoding.

The effect of the decrease in $d_{min}$ as a consequence of the random interleaver and the serial concatenation can also be verified from the slopes of the curves in Fig. 4 for $\frac{E_b}{N_o} > 3$dB. Curves 1,2,3 correspond to the cases of $d_{min} = 2, 3, 6$ respectively, and it is verified that the larger the value of $d_{min}$ is, the steeper the curves are for sufficiently high values of $\frac{E_b}{N_o}$ ($\frac{E_b}{N_o} > 3$dB is the relevant range

of $\frac{E_b}{N_o}$ values, see Fig. 4). Therefore, for sufficiently high values of $\frac{E_b}{N_o}$, the outer stand-alone code (Fig. 2c) is superior. It is depicted in Fig. 4 that the upper bounds on the block error probabilities of all the three examined codes are approximately the same for $\frac{E_b}{N_o} = 10.3$dB (these upper bounds are approximately also the corresponding exact block error probabilities for such values of $\frac{E_b}{N_o}$, since the union bound is asymptotically tight for large values of $\frac{E_b}{N_o}$).

However, the solid curves of Fig. 4 (ML decoding) present the effect of the random interleaving and the serial concatenation on the weight distribution of the code in the rate region that is of practical interest and on the other hand show the effect of the minimal distance on the performance of the coded systems for sufficiently high values of $E_b/N_o$ considerably below the cutoff rate. It is demonstrated that the concatenated systems are advantageous over a broad region of $\frac{E_b}{N_o}$. That tangential sphere bounds yield significant results also at rates exceeding the cutoff rate [27].

**B. Simulation results:** The system comprising the transmitter described in section II-A and the iterative decoder of section II-C performing 10 iterations was simulated while generating an independent random interleaver for each simulated block. The performance in terms of BER and BLOER (i.e. the probability of at least one error in a decoded block of 496 information bits $b_i$) was estimated (see Fig. 5). We examined the system with the two inner codes DE and D3 described in section II-A and compared them to the standard convolutionally coded system without inner code, comprising the convolutional code only, and decoded by an optimal BER minimizing receiver, see section II-B, which is equivalent to omitting the inner code and performing only one iteration at the receiver described in section II-C.

The most remarkable result is the good performance of the differential interleaved coded system which needs 0.8dB less $E_b/N_o$ to achieve BER of $10^{-3}$ than the outer convolutional code alone demonstrating thus the superiority of the concatenated scheme predicted by the upper bounds above. This is caused by the good WD of the code, as discussed above, and by the high performance of the iterative receiver which approaches in this case, within 1dB, the upper bound on BLOER above. Also worth noting is the high ratio RE of bit error rate over the block error rate which is easily shown to be equal to the average of BER in the erroneous blocks. For the iteratively decoded systems RE is in the range of 0.18 to 0.25 with DE and 0.25 to 0.4 with D3 for BER above $10^{-3}$ as noticed before in some iteratively decoded systems at low $E_b/N_o$ [14] in contrast to much lower values with no inner code (about $10^{-2}$ for low BER which is determined for very low BER solely by the number of errors of the minimum distance error event and block length). High RE yields a low BLOER for a given BER which is a substantial advantage in some communication systems. Indeed the BLOER performance of the system with DE is better at BLOER $= 10^{-2}$ by 2dB than that of the coherent standard convolutional coded system.

To compare meaningfully the simulation results to the above bounds, we examine the $E_b/N_o$ required to achieve BLOER $= 10^{-2}$, see Fig. 4. Without an inner code, 4.3dB are needed for the simulated system versus 5dB predicted by the tangential sphere based upper bound. This is a good fit, since the simulated system performs an optimal detection minimizing BER which approximates closely a BLOER minimizing detector (such as an MLSE Viterbi decoder). The simulated system with DE inner code requires $E_b/N_o = 2.25$dB for BLOER $= 10^{-2}$, while the upper bounds predicts that $E_b/N_o = 1.1$dB is possible. This indicates the suboptimality of the iterative decoder which performs worse by at least 1.15dB than a global MLSE decoder. This suboptimality is even more pronounced with the D3 inner code where the simulated system requires 3.25dB, a 2.25dB more than the 1dB upper bound on block error rate.

The discrepancy in the simulated performance of the DE and D3 codes as compared to the bounds (see Figs. 4,5) is attributed to the inherent limitations of the suboptimal iterative decoding. In appendix IV, we provide some intuition to this phenomenon based on assessing the associated average mutual information. The iterative decoder is more complex than a non-iterative standard decoder with the same convolutional code. Thus from a practical point of view the performance of the iterative system with 16 states outer code should be compared to a non-concatenated and non-iterative one based on a more complex error correcting code. Indeed the iterative system described here outperforms, at BER below $10^{-3}$, the industry standard 128 states rate $= 0.5$ convolutional code over a coherent channel, see for example [21].

## V. Summary and Conclusions

A family of serially concatenated codes comprising of outer convolutional code, a random and uniform interleaver and a rate 1 recursive convolutional inner code (such as differential encoding) is investigated for use over a coherent AWGN channel. This study is motivated by the excellent performance of those codes, iterative decoded over a noncoherent channel. Here we invoke the improved tangential sphere bounding technique [27], which is not subjected to the harsh threshold effect of the union bound for relatively low signal to noise ratio values. Further we examine a longer interleaver and an outer code of prolonged memory, as compared to [8] and achieve therefore, a considerable improvement in the performance bounds. The superiority of the serially concatenated codes over the stand alone outer convolutional code as demonstrated by the bounds is a rather surprising result, since the inner code is of rate 1 and is known to degrade performance in conjunction with conventional decoders [6] where no interleaving is introduced. We attribute the improved performance which is demonstrated here by analytical bounds to the distance spectral thinning effect, which counter balances (for low and moderate $\frac{E_b}{N_o}$) the minimum Hamming weight is usually

degraded by the inner differential encoder.

We consider the performance of the sub-optimal (practical) iterative decoding by simulations, and show that also this method enjoys the advantage of this special simple concatenation. By providing examples where the simulated block error rate of iterative decoding exceeds the upper bound on the ML decoder, we demonstrate the mild suboptimality of the iterative decoding as compared to the optimal ML decoding. By comparing the performance with different rate 1 inner codes we have argued that in this setting the average mutual information (AMI) between the received sequence of symbols and an inner code input bit at the first iteration is a significant criterion in determining the performance of the iterative algorithm, while the average sequence-wise AMI determines to ultimate capacity.

## Appendix I: The calculation of the weight enumerator of the outer code

Our paper assumes an outer code that is a rate $\frac{1}{2}$ convolutional code of memory length $m = 4$ and generators that are represented as (23, 35) in octal form (it was taken from the list of table 8-2-1, [29]). The structure of the encoder is shown in Fig. 2e.

The state equations of this encoder are easily derived:

$$\begin{cases} c_{1,n} = b_n \oplus \sigma_{3,n} \oplus \sigma_{4,n} \\ c_{2,n} = b_n \oplus \sigma_{1,n} \oplus \sigma_{2,n} \oplus \sigma_{4,n} \\ \sigma_{2,n} = \sigma_{1,n-1} \\ \sigma_{3,n} = \sigma_{2,n-1} \\ \sigma_{4,n} = \sigma_{3,n-1} \end{cases}$$

Following the derivation technique used for the calculation of the output weight enumerator of a convolutional code [12], the $xy$-incidence matrix corresponding to the state diagram of the encoder in Fig. 6 is the following:

$A(x,y) =$

| | 0000 | 0001 | 0010 | 0011 | 0100 | 0101 | 0110 | 0111 | 1000 | 1001 | 1010 | 1011 | 1100 | 1101 | 1110 | 1111 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0000 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0001 | $y^2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $x$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0010 | 0 | $y$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ | 0 | 0 | 0 | 0 | 0 | 0 |
| 0011 | 0 | $y$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ | 0 | 0 | 0 | 0 | 0 | 0 |
| 0100 | 0 | 0 | $y$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ | 0 | 0 | 0 | 0 | 0 |
| 0101 | 0 | 0 | $y$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ | 0 | 0 | 0 | 0 | 0 |
| 0110 | 0 | 0 | 0 | $y^2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $x$ | 0 | 0 | 0 | 0 |
| 0111 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy^2$ | 0 | 0 | 0 | 0 |
| 1000 | 0 | 0 | 0 | 0 | $y$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ | 0 | 0 | 0 |
| 1001 | 0 | 0 | 0 | 0 | $y$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ | 0 | 0 | 0 |
| 1010 | 0 | 0 | 0 | 0 | 0 | $y^2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $x$ | 0 | 0 |
| 1011 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy^2$ | 0 | 0 |
| 1100 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy^2$ | 0 |
| 1101 | 0 | 0 | 0 | 0 | 0 | 0 | $y^2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $x$ | 0 |
| 1110 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $y$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ |
| 1111 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $y$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $xy$ |

and the matrix $A(y)$ that corresponds to the output weight only follows by substituting $x = 1$ in the matrix $A(x,y)$ above.

Based on the technique presented in [18], its output weight distribution is given by

$$W(y,z) = z^m \left\{ (I - zA(y))^{-1} \; A(y)^m \right\}_{(0,0)}$$

where in our case $m = 4$ is the memory length of the convolutional encoder, $I$ is the identity matrix of order 16 ($2^m$, the number of possible states of the shift register) and $\left\{ T \right\}_{(0,0)}$ represents the term in the first row and first column of the matrix $T$. It is assumed in this derivation that $m$ bits are added for the termination of the input sequence (for returning the shift register to the zero state).

For the case of an interleaver of length $N$ in the serially concatenated code, since the rate of the code is $\frac{1}{2}$, we are interested in the output weight distribution that corresponds to an information sequence of length $\frac{N}{2} - m$ bits and then a termination of additional $m$ bits. The corresponding output weight distribution is written as $W_{N/2}(y)$.

As in [18], by computing the characteristic polynomial $det\,(I - z\,A(y))$, a recursive equation for the output weight distribution is found:

$$W_k(y) \;\; = \;\; (y+1)\,W_{k-1}(y) + (y^3 - y)\,W_{k-2}(y) + (y^6 - y^7)\,W_{k-5}(y) +$$

$$+(y^{13}-y^{10}-2y^9+y^7+y^6+y^5-y^3)\,W_{k-8}(y)+$$
$$+(-y^{13}+y^{12}+2y^9-2y^8-y^5+y^4)\,W_{k-10}(y)+$$
$$+(-y^{19}+y^{16}+3y^{15}-2y^{12}-3y^{11}+3y^8+y^7-y^4)\,W_{k-12}(y),\quad k\le 13\,.$$

since the memory length is $m=4$:

$$W_1(y)=W_2(y)=W_3(y)=W_4(y)=1$$

and also by the equation in [18] $W_k(y)=\left\{A(y)^k\right\}_{(0,0)}$ $k=1,2,3\ldots$ The additional required initial functions needed for the recursion are then calculated:

$$\begin{aligned}
W_5(y) &= y^6+1\\
W_6(y) &= y^7+2y^6+1\\
W_7(y) &= y^9+y^8+2y^7+3y^6+1\\
W_8(y) &= y^{10}+4y^9+2y^8+3y^7+5y^6+1\\
W_9(y) &= y^{12}+2y^{11}+5y^{10}+8y^9+3y^8+5y^7+7y^6+1\\
W_{10}(y) &= y^{16}+y^{13}+6y^{12}+8y^{11}+11y^{10}+14y^9+6y^8+7y^7+9y^6+1\\
W_{11}(y) &= y^{17}+2y^{16}+y^{15}+2y^{14}+10y^{13}+20y^{12}+17y^{11}+\\
&\quad 21y^{10}+24y^9+9y^8+9y^7+11y^6+1\\
W_{12}(y) &= y^{19}+2y^{18}+4y^{16}+12y^{15}+18y^{14}+34y^{13}+44y^{12}+\\
&\quad 30y^{11}+37y^{10}+36y^9+12y^8+12y^7+13y^6+1
\end{aligned}$$

Note that the sum of the coefficients of the output weight distribution $W_k(y)$ is $2^{k-m}$ $(k\ge m)$, because the weight of the last $m$ symbols at the input of the encoder is not taken into account (used only for termination to the zero state).

The initial functions and the recursive equation derived above are used for the calculation of $W_{N/2}(y)$. In the examined case, we are interested in the case of an interleaver length $N=1000$ bits, and therefore this recursive equation was used $\frac{N}{2}-12=488$ times.

## Appendix II: The calculation of the input-output weight distribution of the differential encoder

The differential encoder is a recursive convolutional code of rate 1 and a memory length of $m=1$. The state equations of a differential encoder are:

$$\begin{cases} d_n &= \sigma_{1,n-1}\\ \sigma_{1,n} &= c_n\oplus\sigma_{1,n-1}\end{cases}.$$

Its state diagram is presented in Fig. 2a.

The input-output weight distribution of a differential encoder is presented in [9].

## Appendix III: The calculation of the input-output weight distribution of the codes DDE and D3

The alternative inner code that is discussed in our paper is a recursive convolutional code of rate 1 and a memory length of $m = 2$, DDE and D3.

The encoders of these codes are shown in Fig. 2b. As explained, the WD of the DDE code is exactly the same as that of the DE. The state equations of the D3 code are the following:

$$\begin{aligned} \sigma_{1,n} &= c_n \oplus \sigma_{1,n-1} \oplus \sigma_{2,n-1} \\ \sigma_{2,n} &= \sigma_{1,n-1} \\ d_n &= \sigma_{2,n-1} \quad , \end{aligned}$$

and its state diagram is presented in Fig. 2b (the powers of $x$ and $y$ there are the Hamming weight of the input and output respectively). As before, following the derivation technique for the input-output weight distribution or a convolutional code [18], the $xy$-incidence matrix derived from the state diagram in Fig. 2b is

$$A(x,y) = \begin{array}{c} \\ 00 \\ 01 \\ 10 \\ 11 \end{array} \begin{array}{c} \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{pmatrix} 1 & 0 & x & 0 \\ xy & 0 & y & 0 \\ 0 & x & 0 & 1 \\ 0 & y & 0 & xy \end{pmatrix} \end{array}$$

and also $A(y) \stackrel{\triangle}{=} A(1,y)$.

Based on [18], the input-output weight distribution is

$$W(x,y,z) = z^m \left\{ (I - z\,A(x,y))^{-1}\; A\,(y)^m \right\}_{(0,0)} \quad ,$$

where $T_{(0,0)}$ denotes the term of the first column and first row of the matrix $T$ and $m = 2$ is the memory length of the inner code in this case.

The denominator of $W(x,y,z)$ is the characteristic polynomial

$$det(I - z\,A(x,y)) = 1 - z(1+xy) - z^3(y^2 - xy)(1-x^2) + z^4(1-x^2)^2y^2$$

and therefore the recursive equation for the input-output weight distribution $W_k(x,y)$ is:

$$W_k(x,y) = (1+xy)W_{k-1}(x,y) + (y^2 - xy)(1-x^2)W_{k-3}(x,y) - y^2(1-x^2)^2W_{k-4}(x,y)\ .$$

It follows that,

$$W(x,y,z) = \sum_{k=2}^{\infty} W_k(x,y)z^k \text{ when } W_k(x,y) = \left[A(x,y)^{k-2}A(y)^2\right]_{(0,0)} \quad ,$$

which by a direct computation of $W_k(x,y)$, is used to calculate the initial weight distributions needed for implementing the recursive equation:

$$\begin{aligned} W_2(x,y) &= 1 \text{ (as expected since } W_k(x,y)=1 \text{ for values of } k \text{ such that } k \le m) \\ W_3(x,y) &= 1+xy \\ W_4(x,y) &= 1+xy+x^2y+xy^2 \\ W_5(x,y) &= 1+xy+x^2y+2xy^2+x^3y+x^2y^2+x^2y^3 \ . \end{aligned}$$

Since a serially concatenated code is considered with an interleaver length of $N$ bits and 2 additional cycles are needed for termination [25], the computation of the input-output weight distribution $W_{N_c+2}(x,y)$ is required (therefore, the recursive equation derived above is used $N_c-2$ times).

## Appendix IV: Average Mutual Information (AMI)

The performance of the system with the D3 inner code, and to a lesser degree the system with a DE inner code are limited by the capability of the iterative decoding algorithm which is suboptimal. The capability of the iterative algorithm to achieve near optimal performance in various systems is an open research subject, see for example [31],[19]. To gain a partial and incremental understanding of the issues governing the behavior of the iterative decoder in our serially concatenated system, we examined the flow of information to the outer code at the first iteration. In particular we computed the Average Mutual Information (AMI) between the Inner Decoder (ID) input $\mathbf{r}$ and a one (any) bit $c_i$ coded by the outer encoder while considering the bits $c_i$ as independent and ignoring the outer code as does the ID. Since at the first iteration $\mathbf{r}$ is the ID's only input and the ID is an optimal MAP decoder, this AMI is an indication of the ID information content. An AMI too low relatively to the outer code rate at first iteration, may prevent the convergence of the iterative algorithm to a near optimal solution. We use

$$AMI = I(c_j;\mathbf{r}) = E\left[\log \frac{p(\mathbf{r}|c_j)}{p(\mathbf{r})}\right] \ ,$$

(where $E$ denotes statistical expectation) which is easy to estimate by the Monte Carlo method, since only 2 or 3 symbols $r_j$ are statistically dependent on $c_j$ for the DE and D3 codes respectively due to the uniform, independent and identical distribution (u.i.i.d.) of $c_j$, (see the assertion in Appendix V). The resulting AMI are presented versus $E_s/N_o$ in Fig. 6. The AMI curve of the D3 code is inferior by 1.15dB at AMI = 0.5 to that of the DE code as compared to a similar difference of 1 and 1.05dB in the BLOER and BER performance of those codes listed above. Also the $E_b/N_o$ required to achieve BER of $10^{-2}$ occurs when $RA \stackrel{\triangle}{=} AMI/\text{code rate}$ is 0.8 to 0.9 for both codes. A similar value of RA for the component codes occurs at BER = $10^{-2}$ for $R=0.5$ parallel concatenated turbo codes [5]. We do not claim to have provided a sound theoretical basis

of the dependence of iterative algorithm convergence on RA of the component codes and counter examples can be traced in [22]. There the performance of a similar interleaved iteratively decoded system, which used turbo code instead of the convolutional code here and operated in a noncoherent setting, improved significantly when 'RDE' code replaced DE in spite of identical input and output alphabets, rates and AMI at first iteration of both RDE and DE. (RDE was chosen to exploit more effectively the side information on further iterations and we do not expect it to perform well when concatenated with a convolutional code). Furthermore AMI, even at first iteration, is hard to evaluate analytically for serially concatenated codes with complex inner codes, and therefore bounds [29] and Monte Carlo methods are employed.

## Appendix V:

**Definition:** Binary linear rate 1 code with constraint length $k$ and length $N$ is a mapping of any sequence $\mathbf{b}(b_1 \dots b_N)$ of bits $b_i$ to a sequence $\mathbf{s} = (s_1 \dots s_N)$ of bits $s_i$

$$s_i = b_i \oplus f\left(\mathbf{s}_{i-k+1}^{i-1}\right) \tag{V.1}$$

where $f$ is a mapping to (1,0), $\mathbf{s}_j^k = (s_j \dots s_k)$ and $s_i|_{i\le 0} = 0$ is an initial state known to the receiver and the transmitter.

**Property 1:** from (V.1):

$$b_i = s_i \oplus f\left(\mathbf{s}_{i-k+1}^{i-1}\right) \tag{V.2}$$

Thus the code is uniquely decodable and the mapping $\mathbf{b} \to \mathbf{s}$ is one to one.

**Assertion:** When a sequence of $N$ uniformly independently and identically distributed (u.i.i.d.) bits $b_i$ is encoded by binary linear rate 1 code to a sequence of $s_i$ and transmitted over a memoryless channel the output of which is $\mathbf{r} = (r_1 \dots r_N)$ then:

$$I(b_i; \mathbf{r}) = I\left(b_i; \mathbf{r}_{i-k+1}^{i}\right) \tag{V.3}$$

**Proof:** Define:

$$\begin{aligned} \mathbf{s}_- &= \mathbf{s}_1^{i-k} \quad , \quad \mathbf{r}_- = \mathbf{r}_1^{i-k} \\ \mathbf{s}_+ &= \mathbf{s}_{i+1}^{N} \quad , \quad \mathbf{r}_+ = \mathbf{r}_{i+1}^{N} \\ \mathbf{s}_* &= \mathbf{s}_{i-k+1}^{i} \quad , \quad \mathbf{r}_* = \mathbf{r}_{i-k+1}^{i} \end{aligned}$$

then:

$$I(b_i; \mathbf{r}) = I(b_i; \mathbf{r}_-, \mathbf{r}_*, \mathbf{r}_+) \le I(b_i; \mathbf{r}_-, \mathbf{r}_*, \mathbf{r}_+, \mathbf{s}_-, \mathbf{s}_+) = I(b_i; \mathbf{r}_*, \mathbf{s}_-, \mathbf{s}_+) \stackrel{\triangle}{=} I_1 \tag{V.4}$$

by standard decomposition:

$$I_1 = I(b_i; \mathbf{r}_* | \mathbf{s}_-, \mathbf{s}_+) + I(b_i; \mathbf{s}_-, \mathbf{s}_+)$$

Now,

$$I(b_i; \mathbf{s}_-, \mathbf{s}_+) = I(b_i; \mathbf{s}_-) + I(\mathbf{s}_+; b_i|\mathbf{s}_-)$$

The first summand is zero since nothing about $b_i$ is revealed by $\mathbf{s}_-$ alone and so is the second summand since when $\mathbf{s}_-$ and $b_i$ are known each bit $s_j$ of $\mathbf{s}_+$ is xored in eq. (V.1) with a new u.i.i.d. $b_j$ yielding a u.i.i.d. $\mathbf{s}_+$ independent of $b_i$ and $\mathbf{s}_-$. Thus:

$$I_1 = I(b_i; \mathbf{r}_*|\mathbf{s}_-, \mathbf{s}_+) \tag{V.5}$$

Now,

$$p_1 \stackrel{\triangle}{=} p(b_i, \mathbf{r}_*|\mathbf{s}_+, \mathbf{s}_-) = \sum_{\mathbf{s}_*} p(b_i, \mathbf{r}_*|\mathbf{s}_*, \mathbf{s}_+, \mathbf{s}_-)p(\mathbf{s}_*|\mathbf{s}_+, \mathbf{s}_-)$$

$b_i$ is uniquely determined from $\mathbf{s}_*$ by eq. (V.2) and $p(\mathbf{r}_*)$ is determined by $\mathbf{s}_*$ and the memoryless channel, thus $\mathbf{s}_+$, $\mathbf{s}_-$ may be deleted from the first probability term. Also $\mathbf{s}_*$ is u.i.i.d. regardless of $\mathbf{s}_+$, $\mathbf{s}_-$ due to the xor by $b_j$ in eq. (V.1), thus $\mathbf{s}_+$, $\mathbf{s}_-$ may be deleted for the second term, yielding,

$$p_1 = \sum_{\mathbf{s}_*} p(b_i, \mathbf{r}_*|\mathbf{s}_*)p(\mathbf{s}_*) = p(b_i, \mathbf{r}_*)$$

Thus from (V.5):

$$I_1 = I(b_i; \mathbf{r}_*) \ .$$

Combining this with (V.4) and using the obvious relation

$$I(b_i; \mathbf{r}) \geq I(b_i; \mathbf{r}_*)$$

yields the assertion (V.3).

## References


[1] I. Bar-David, and A. Elia, "Incremental Frequency, Amplitude and Phase Tracker (IFAPT) for Coherent Demodulation over Fast Flat Fading Channels", *Proc., Inform. Theory Workshop*, Killarney, Ireland, June 1998.

[2] G. Battail, "A conceptual framework for understanding turbo-codes", *IEEE Journal Selected Areas in Commun.*, Vol. 16, No. 2, pp. 245-254, Feb. 1998.

[3] S. Benedetto, G. Montorsi, D. Divsalar, F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding", *JPL-TDA Progress Report*, vol. 42-126, August 15, 1996. See also the Proc., 1997 IEEE Int. Symp. Inform. Theory (ISIT'97), p. 106, Ulm, Germany, July 1997.

[4] S. Benedetto, D. Divsalar, G. Montorsi, F. Pollara, "A soft-input soft-output APP module for iterative decoding of concatenated codes", *IEEE Communication Letters*, vol. 1, No. 1, pp. 22-24, Jan. 1997.

[5] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes", *IEEE Trans. Commun.*, vol. 44, No. 10, pp. 1261-1271, Oct. 1996.

[6] D. Divsalar and M.K. Simon, "Maximum-likelihood differential detection of uncoded and trellis coded amplitude phase modulation over AWGN and fading channels-metrics and performance", *IEEE Trans. Commun.*, vol. 42, No. 1, pp. 76-89, Jan. 1994.

[7] D. Divsalar and F. Pollara, "Turbo codes for PCS applications", *Proc., IEEE Int. Conf. Comm. ICC'95*, pp. 54-59, Seattle, Washington, June 1995.

[8] D. Divsalar and F. Pollara, "Serial and hybrid concatenated codes with applications", *Proc., of the Int. Symp. on Turbo Codes and Related Topics*, pp. 80-87, Brest, France, Sept. 1997.

[9] D. Divsalar, H. Jin and R.J. McEliece, "Coding theorems for 'turbo-like' codes", *Proc., Allerton Conference*, Sept. 1998.

[10] T.M. Duman and M. Salehi, "New performance bounds for turbo codes", *IEEE Trans. Commun.*, Vol. 46, No. 6, pp. 717-723, June 1998. See also: T.M. Duman, "Turbo codes and turbo coded modulation systems: Analysis and performance bounds", Ph.D. Dissertation, Elect. Comput. Eng. Dept., Northeastern University, Boston, MA, May 1998.

[11] R.F.H. Fischer, S. Calabro, L.H.-J. Lampe and S.H. Muller-Weinfurtner, "Performance of coded modulation employing differential encoding over Rayleigh fading channels", *Electronic Letters*, vol. 35, No. 2, pp. 122-123, Jan. 1999.

[12] S. Galán, M. Peleg and S. Shamai (Shitz), 'On iterative phase trellis based noncoherent detection of coded MPSK in a noisy phase regime", *Proc., MELECON'98*, Tel Aviv, Israel, pp. 834-838, May 1998.

[13] M.J. Gertsman, J.H. Lodge, "Symbol by symbol MAP demodulation of CPM and PSK signals on Rayleigh flat fading channels", *IEEE Trans. Commun.*, vol. 45, No. 7, pp. 788-799, July 1997.

[14] J. Hamorsky and U. Wachsmann, "Criteria for minimum bit error rate decoding of turbo-codes", *Proc., Kleinheubacher Tagung*, pp. 607-616, Oct. 1995.

[15] P. Hoeher and J. Lodge, "Iterative differential PSK demodulation and channel decoding", *Proc., 1998 Int. Symp. Inform. Theory (ISIT'98)*, pp. 425, MIT, Cambridge, MA, USA, Aug. 1998.

[16] D. Makrakis, D.P. Bouras and P.T. Mathiopoulos, "Performance analysis of asymptotically optimal noncoherent detection of trellis-coded multi-amplitude-phase modulation signals in Gaussian noise and ISI channels", *IEEE Journal Selected Commun.*, vol. 13, No. 2, pp. 354-370, Feb. 1995.

[17] I.D. Marshland, P.T. Mathiopoulos, S. Kallel, "Noncoherent turbo-equalization for frequency selective Rayleigh fast fading channels", *Proc., Int. Symp. on Turbo Codes & Related Topics*, pp. 196-199, Brest, France, Sept. 1997.

[18] R.J. McEliece, "How to compute weight enumerators for convolutional codes", *Communications and Coding*, M. Darnel and B. Honary, eds., Taunton, Somerset, England Research Studies Press Ltd. 1998, pp. 121–141, see also: http://www.systems.caltech.edu/EE/Faculty/rjm.

[19] A. Moher and A. Gulliver, "Cross-entropy and iterative decoding", *IEEE Trans. Inform. Theory*, vol. 44, No. 7, pp. 3097-3104, Nov. 1998.

[20] Y. Okunev, *Phase and phase-difference modulation digital communications*, 1997, Artech House.

[21] M. Peleg and S. Shamai (Shitz), "Iterative decoding of coded and interleaved noncoherent multiple symbol detected DPSK", *Electronic Letters*, vol. 33, No. 12, pp. 1018-1020, June 1997.

[22] M. Peleg, S. Shamai (Shitz) and S. Galán, "On iterative decoding for coded noncoherent MPSK communications over block-noncoherent AWGN channel", *Proc., Int. Conf. Telecommun., ICT'98*, Halkidiki, Greece, pp. 109–114, June 1998. See also [12].

[23] L.C. Perez, J. Seghers and D.J. Costello, "A distance spectrum interpretation of turbo codes", *IEEE Trans. Inform. Theory*, Vol. 42, No. 6, pp. 1698-1709, Nov. 1996.

[24] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra", *IEEE Trans. on Inform. Theory*, Vol. 40, No. 4, pp. 1284-1292, July 1994.

[25] J.G. Proakis, *Digital Communications*, McGraw-Hill International Editions, Third Edition, 1995.

[26] D. Raphaeli, "Decoding algorithms for noncoherent trellis coded modulation", *IEEE Trans. Commun.*, vol. 44, No. 3, pp. 312-323, March 1996.

[27] I. Sason and S. Shamai (Shitz), "Improved upper bounds on the performance of parallel and serial concatenated turbo codes via their ensemble distance spectrum", *Proc., IEEE Int. Symp.*

*Inform. Theory, (ISIT'98)*, 16–21 August 1998, pp. 30, Boston, USA. To appear in IEEE Trans. Inf. Theory.

[28] I. Sason and S. Shamai (Shitz), "On union bounds for random serially concatenated codes with maximum likelihood decoding", *Proc., French-Israel Workshop in Coding Inform. Integrity*, Ein-Boqeq, Dead sea, Israel, 27-29 October 1997. To appear in European Trans. Telecommun.

[29] S. Shamai (Shitz) and S. Verdú, "Capacity of channels with uncoded side information", *European Trans. Telecommun.*, Vol. 6, No. 5, Sept.-Oct. 1995, pp. 587-600.

[30] A.J. Viterbi, A.M. Viterbi, J. Nicolas and N.T. Sindushyana, "Perspectives on interleaved concatenated codes with iterative soft output decoding", *Proc., Int. Symp. on Turbo Codes & Related Topics*, pp. 47-54, Brest, France, Sept. 1997.

[31] N. Wiberg, "On the performance of the iterative turbo decoding algorithm", *Proc., Int. Symp. Turbo Codes & Related Topics*, pp. 223-226, Brest, France, Sept. 1997.

## Figure Captions

Fig. 1: System block diagram.

Fig. 2: Encoders and state diagrams: (a) Inner differential encoding (DE). (b) Inner D3 encoding. (c) Double differential encoder (DDE). (d) D31 encoder equivalent to DDE. (e) Outer convolutional encoder, $g = 23, 35$ octal (encoder only).

Fig. 3: The weight distribution of the interleaved serially concatenated codes compared to that of the outer convolutional code alone.

Fig. 4: Upper bounds on the ML decoded block error probabilities of the schemes in figure 1. The bounds are compared to simulated iterative decoding.

1. A serially concatenated code with an outer non-recursive convolutional code (23, 35) in octal form and the D3 inner code. A random interleaver of length $N = 1000$ bits between the two component codes is assumed. The points marked by '+' are simulation results of iterative decoding (10 iterations).

2. A serially concatenated code with the same outer code and interleaver and a differential encoder (DE) inner code. The points marked by '*' are simulation results of iterative decoding (10 iterations).

3. The non-cursive convolutional code of rate $\frac{1}{2}$ and generators $g = (23, 35)$ in octal form. The points marked by 'o' are simulations based on MAP decoding (Bahl algorithm).

Fig. 5: Performance of simulated interleaved systems with an outer $rate = 0.5$ non-recursive convolutional code with generators $g_1 = 23$, $g_2 = 35$ (octal) over a coherent AWGN channel. Interleaver length is 1000 bits. Solid lines: Bit error rate; dashed lines: block error rate; o: no inner code; *: differential inner code; +: D3 inner code.

Fig. 6: Average Mutual Information $I(\mathbf{r}; c_i)$ for different inner codes for coherent BPSK operating on the AWGN channel: DE — *; D3 — +; uncoded — o.

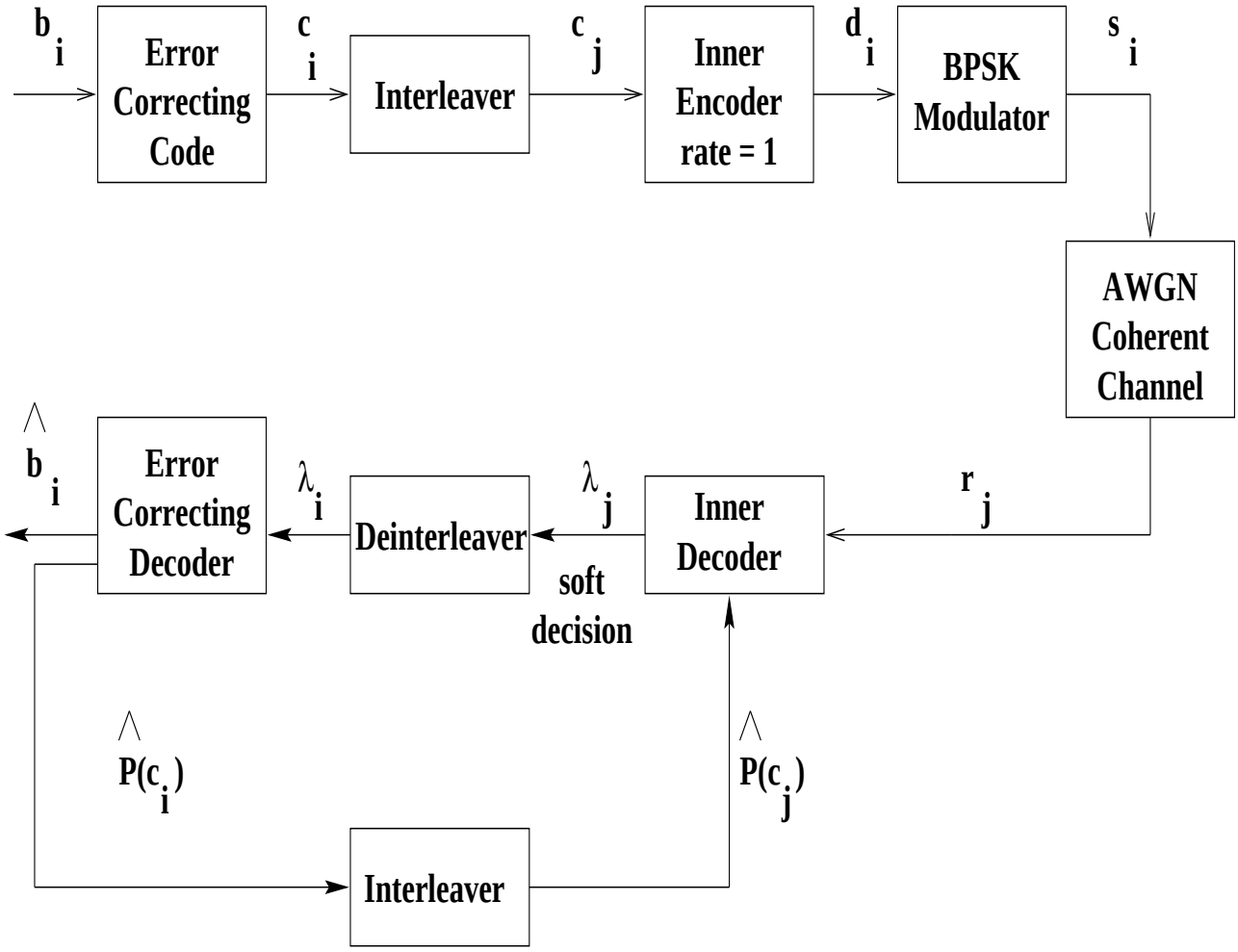
b i
Error Correcting Code
c i
Interleaver
c j
Inner Encoder rate = 1
d i
BPSK Modulator
s i
AWGN Coherent Channel
r j
Inner Decoder
λ j
soft decision
Deinterleaver
λ i
Error Correcting Decoder
b i
P(c i)
P(c j)
Interleaver

Figure 1

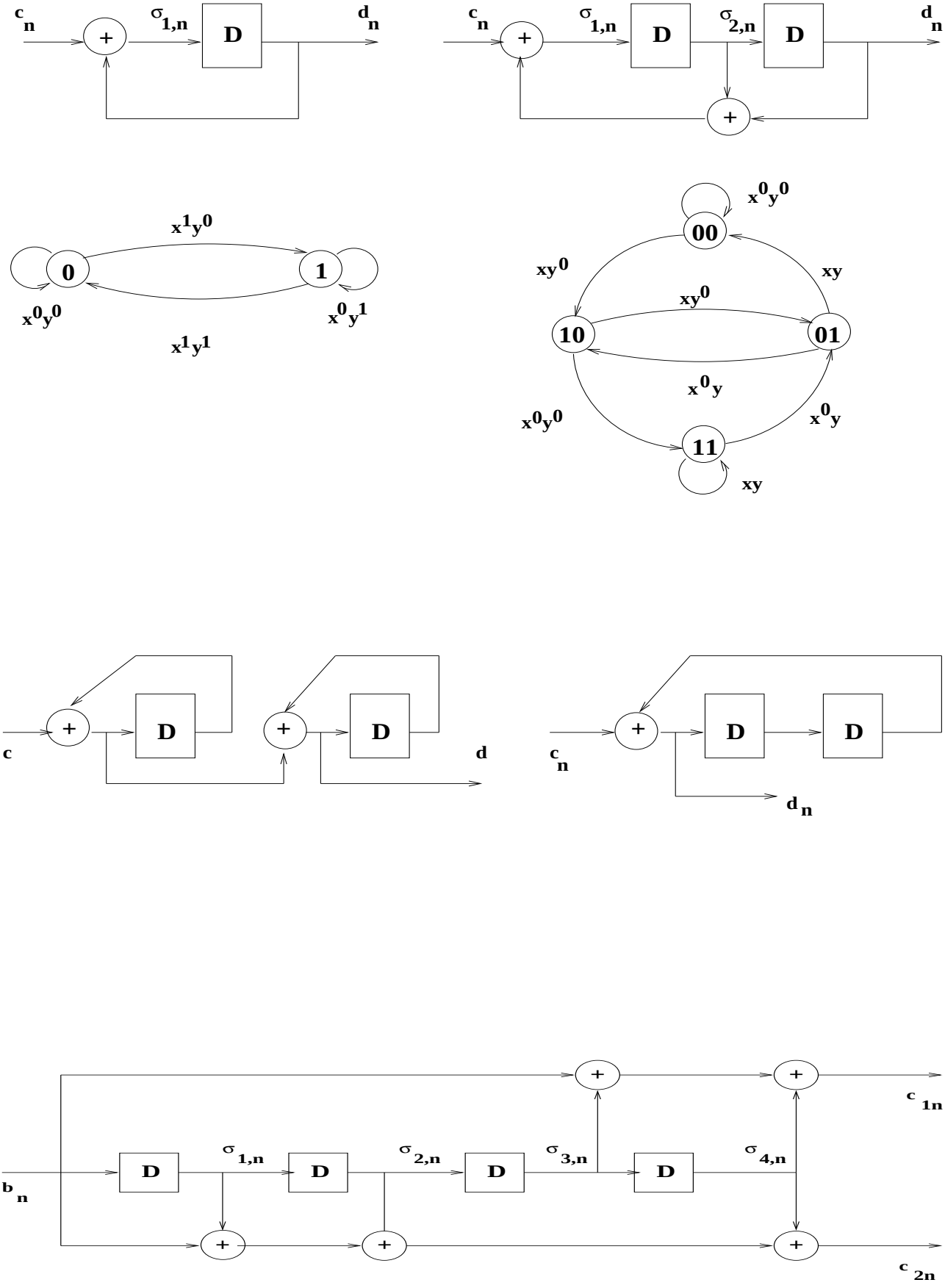
c n
σ 1,n
D
d n
σ 2,n
x 1 y 0
0
1
x 0 y 0
x 1 y 1
x 0 y 1
x 0 y 0
00
xy 0
xy
10
01
x 0 y
11
c
d
b n
σ 3,n
σ 4,n
c 1n
c 2n

Figure 2

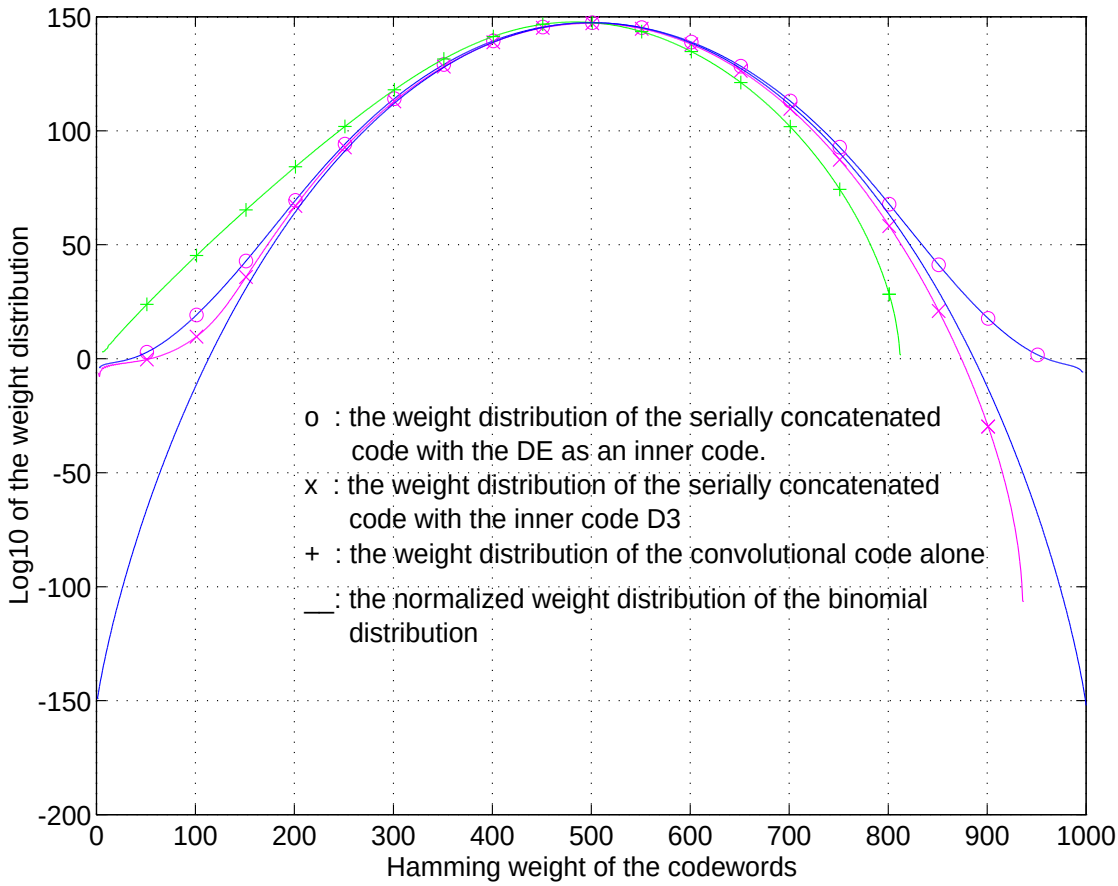
150
100
50
0
-50
-100
-150
-200
0
100
200
300
400
500
600
700
800
900
1000
Log10 of the weight distribution
Hamming weight of the codewords
o : the weight distribution of the serially concatenated code with the DE as an inner code.
x : the weight distribution of the serially concatenated code with the inner code D3
+ : the weight distribution of the convolutional code alone
__: the normalized weight distribution of the binomial distribution

Figure 3.

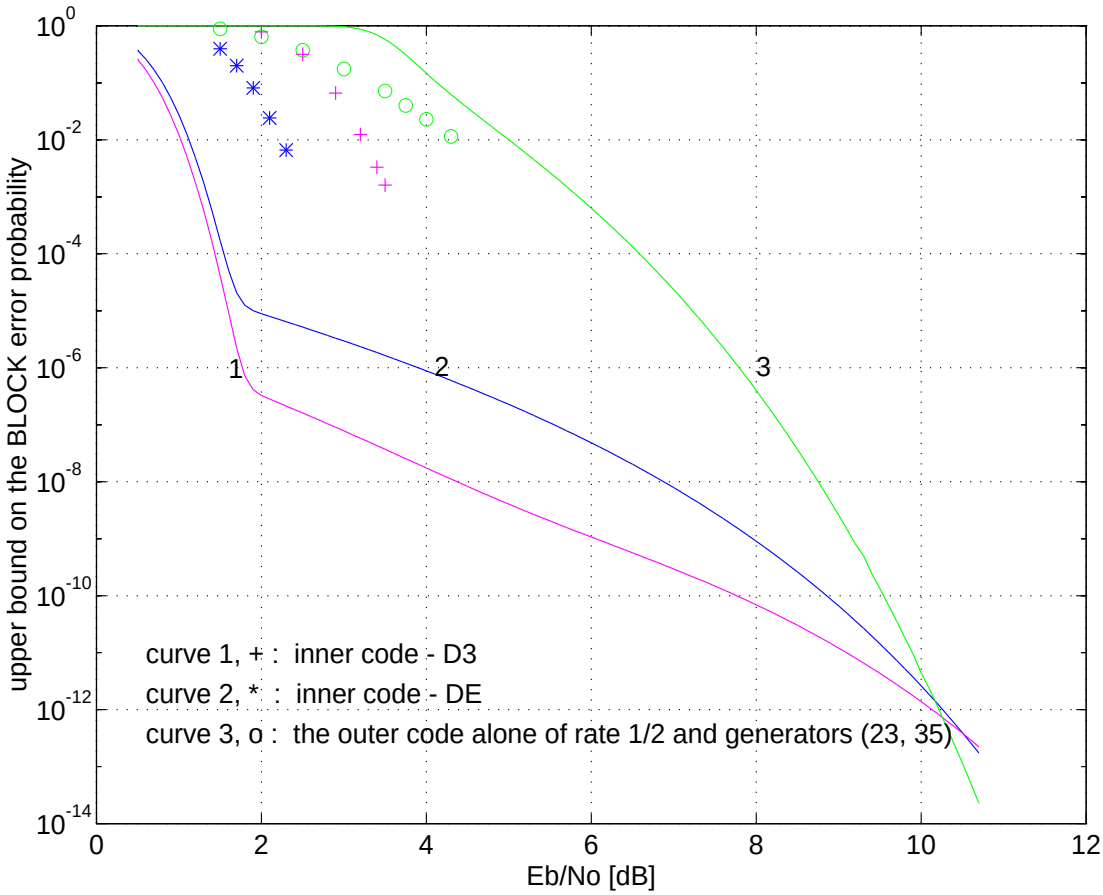
upper bound on the BLOCK error probability
Eb/No [dB]
curve 1, + : inner code - D3
curve 2, * : inner code - DE
curve 3, o : the outer code alone of rate 1/2 and generators (23, 35)
1
2
3

Figure 4.

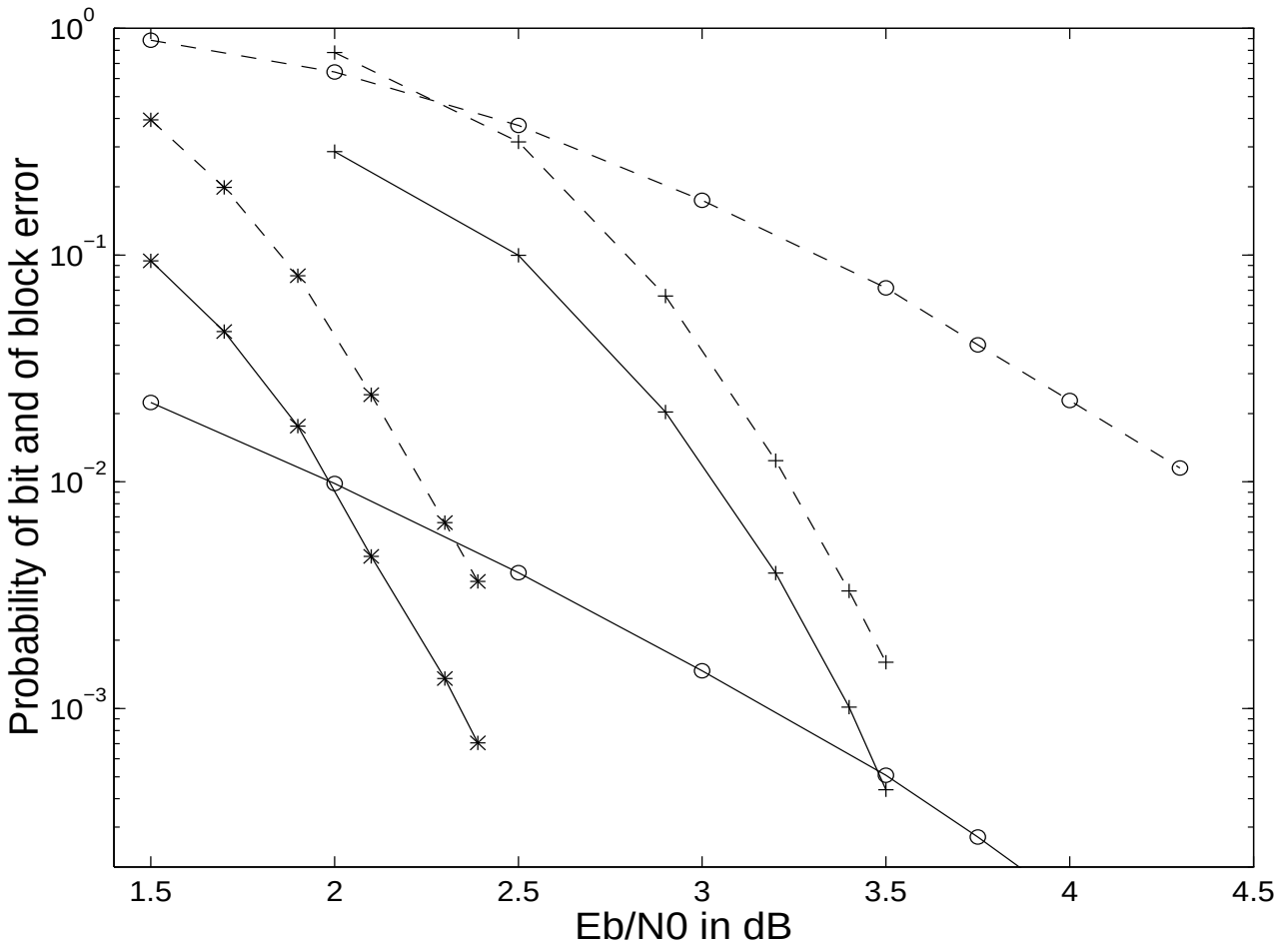
Probability of bit and of block error
$10^0$
$10^{-1}$
$10^{-2}$
$10^{-3}$
1.5
2
2.5
3
3.5
4
4.5
Eb/N0 in dB

Figure 5.

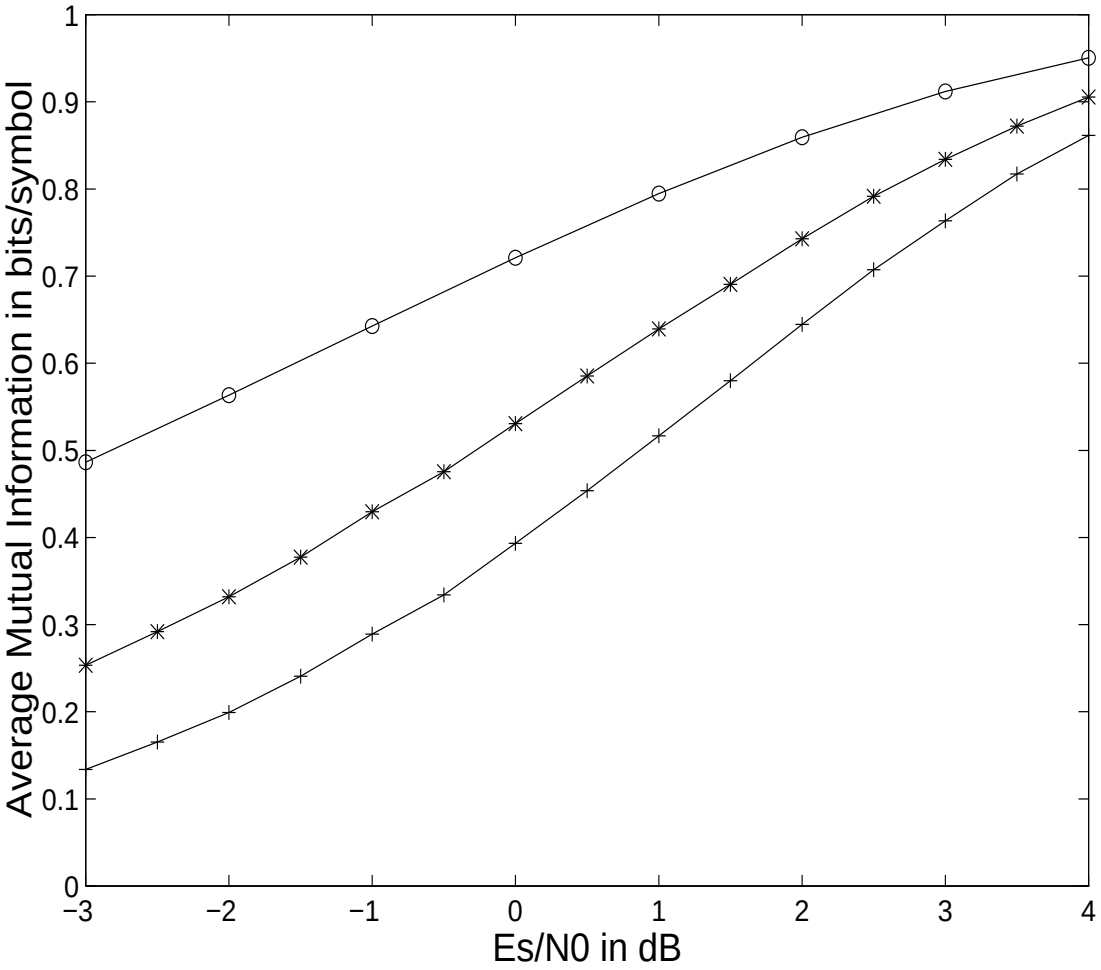
Average Mutual Information in bits/symbol
Es/N0 in dB
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
−3
−2
−1
0
1
2
3
4

Figure 6.