

OCIT  
IP.  
2337A

# SION DECODING FOR LINEAR BLOCK S AND THEIR FORMANCE ANALYSIS

chnical Report

to

**NASA**

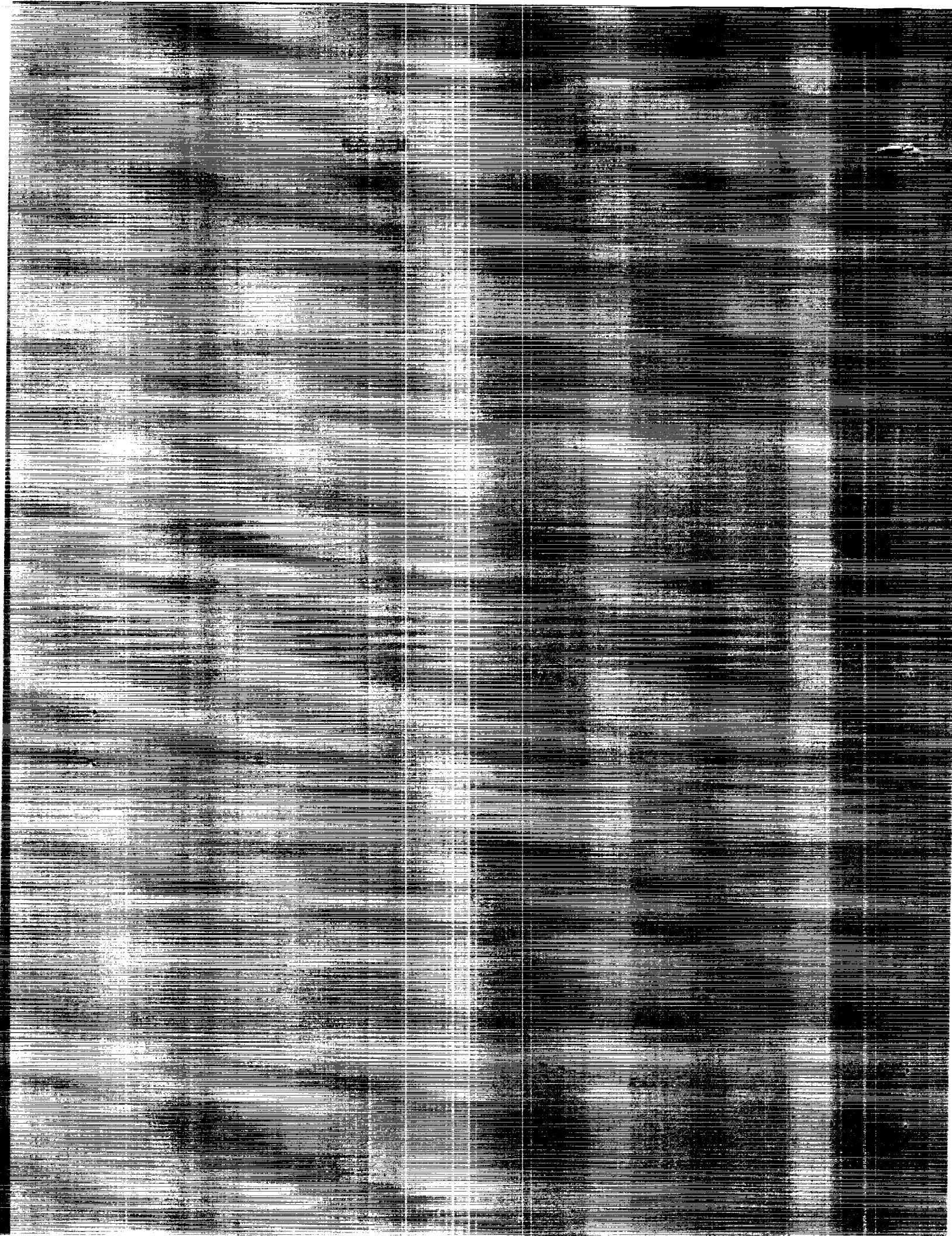
**Engineering Division  
Space Flight Center  
t, Maryland 20771**

**Number NAG 5-931  
ort No. 96-004**

**Investigator: Shu Lin**

**of Electrical Engineering  
y of Hawaii at Manoa  
lu, Hawaii 96822**

**July 10, 1996**



**PART I**

**AN ITERATIVE SOFT-DECISION  
DECODING ALGORITHM**

**Takuya Koumoto, Toyoo Takata, Tadao Kasami, Shu Lin**



# An Iterative Soft-Decision Decoding Algorithm

Takuya Koumoto<sup>†</sup> Toyoo Takata<sup>†</sup> Tadao Kasami<sup>†</sup> Shu Lin<sup>‡</sup>

<sup>†</sup>Graduate School of Information Science,  
Nara Institute of Science and Technology

<sup>‡</sup>Department of Electrical Engineering, University of Hawaii at Manoa

37-11  
41  
237092

## Abstract

This paper presents a new minimum-weight trellis-based soft-decision iterative decoding algorithm for binary linear block codes. Simulation results for the RM(64,22), EBCH(64,24), RM(64,42) and EBCH(64,45) codes show that the proposed decoding algorithm achieves practically (or near) optimal error performance with significant reduction in decoding computational complexity. The average number of search iterations is also small even for low signal-to-noise ratio.

## 1 Introduction

Recently Moorthy et. al.[2] have proposed a zero-and-minimum-weight subtrellis-based iterative decoding scheme for binary linear block codes to achieve a very good trade-off between error performance and decoding complexity. In the scheme, all the candidate codewords are generated by an algebraic decoder based on a set of test error patterns proposed by Chase[1]. The zero-and-minimum-weight trellis search around the current best candidate codeword  $c$  is performed at most once, only if (i) a sufficient condition that the optimal solution is within the minimum distance from  $c$  holds or (ii) all the test error patterns have been exhausted and no candidate codeword satisfies the sufficient condition for optimality.

For the proposed decoding algorithm in this paper, preliminarily presented in [6], the initial candidate codeword is generated by a simple decoder, the zero-th order or the first order decoding proposed in [4]. The subsequent candidate codewords (if needed) are generated by a chain of minimum-weight trellis searches. This minimum-weight trellis around a candidate codeword  $c$  consists of only the codewords in code  $C$  that are at the minimum distance from  $c$ , but does not include  $c$ . The decoding iteration stops whenever a candidate codeword is found to satisfy a sufficient condition for optimality or the latest minimum-weight trellis search results in a repetition of a previously generated candidate codeword. Let this decoding algorithm be denoted Algorithm I-

$w_1$ . The decoding process terminates faster than the Moorthy et. al. algorithm. Furthermore, the use of minimum-weight trellis search considerably reduces the possibility of being trapped into a local optimum. As a result, it achieves better error performance than the Moorthy et. al. algorithm.

A necessary condition for Algorithm I- $w_1$  to achieve good error performance is that the minimum weight codewords span the entire code. Reed-Muller(RM) codes satisfy this condition. Simulation results for the RM(64,22)(the (64,22) RM code), RM(64,42) and the EBCH(64,45)(the extended (64,45) BCH code) codes show that the proposed decoding algorithm practically achieves optimum MLD performance even in the range of relatively low SNR. The EBCH(64,24) code is an example for which the above necessary condition does not hold. For this code, the first and second minimum weight codewords span the entire code. In this case iterative decoding algorithm based on the first and second minimum-weight trellis search, denoted Algorithm I- $w_1$ - $w_2$ , is used.

We also propose another approach to overcome the problem. Let  $C_0$  be a linear subcode of  $C$  and assume that the minimum weight codewords of  $C_0$  span  $C_0$ . The decoding scheme is a combination of: (1) the iterative search using the minimum-weight trellis for  $C_0$  around the latest candidate codeword, and (2) a procedure for moving from the coset of  $C_0$  in  $C$ , containing the current candidate codeword, to another coset which is likely to contain the optimal solution. Simulation results for the EBCH(64,24) code show that this scheme achieves better error performance than Algorithm I- $w_1$ - $w_2$ .

## 2 Sufficient Conditions for Optimality

Suppose a binary  $(N, K)$  linear block code  $C$  is used for error control over the AWGN channel using BPSK signaling. Let  $z = (z_1, z_2, \dots, z_N)$  be the binary hard-decision sequence obtained from the received sequence  $r = (r_1, r_2, \dots, r_N)$ .

Let  $V_N$  denote the vector space of all binary

$N$ -tuples. For an  $N$ -tuple  $\mathbf{u} = (u_1, u_2, \dots, u_N) \in V_N$ , let  $L(\mathbf{u})$  be defined as follows:

$$L(\mathbf{u}) = \sum_{\{i: u_i \neq z_i \text{ and } 1 \leq i \leq N\}} |r_i|. \quad (1)$$

$L(\mathbf{u})$  is called the correlation discrepancy of  $\mathbf{u}$  with respect to  $\mathbf{z}$ , and the smaller  $L(\mathbf{u})$  is, the larger the correlation between  $\mathbf{u}$  and  $\mathbf{r}$  is. For  $\mathbf{u}$  and  $\mathbf{v} \in V_N$ ,  $\mathbf{u}$  is said to be better than  $\mathbf{v}$  if  $L(\mathbf{u}) \leq L(\mathbf{v})$ . For a nonempty subset  $X$  of  $V_N$  and a positive integer  $h$ , let  $h'$  denote  $\min\{|X|, h\}$  and let  $\text{best}_h X$  denote the set of the  $h'$  best  $n$ -tuples in  $X$ , that is, for any  $\mathbf{u} \in \text{best}_h X$  and  $\mathbf{v} \in X - \text{best}_h X$ ,  $L(\mathbf{u}) \leq L(\mathbf{v})$ . The best  $n$ -tuple in  $X$  will be denoted  $\text{best}X$ .

Let  $d_H(\mathbf{u}, \mathbf{v})$  denote the Hamming distance between two  $N$ -tuples,  $\mathbf{u}$  and  $\mathbf{v}$ . For  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_h \in V_N$  and positive integers  $d_1, d_2, \dots, d_h$ , let  $V_N(\mathbf{u}_1, d_1; \mathbf{u}_2, d_2; \dots; \mathbf{u}_h, d_h)$  be defined as the set:  $\{\mathbf{u} \in V_N : d_H(\mathbf{u}, \mathbf{u}_i) \geq d_i \text{ for } 1 \leq i \leq h\}$ , and let  $\underline{L}(\mathbf{u}_1, d_1; \mathbf{u}_2, d_2; \dots; \mathbf{u}_h, d_h)$  be defined as the minimum of  $L(\mathbf{u})$  over  $\mathbf{u} \in V_N(\mathbf{u}_1, d_1; \mathbf{u}_2, d_2; \dots; \mathbf{u}_h, d_h)$ .

Then we have the following early termination condition of an iterative decoding algorithm without degrading the error performance.

**Lemma:** At a stage of an iterative decoding algorithm for a block code  $B$  ( $C$  itself or a coset of a linear subcode of  $C$ ), let  $GC$  denote the set of those candidate codewords which have been generated already, and let  $\mathbf{u}_{\text{best}}$  denote the best of  $GC$ . Suppose that for  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_h \in GC$  and positive integers  $d_1, d_2, \dots, d_h$ ,  $\mathbf{u}_{\text{best}}$  is the best of  $\cup_{i=0}^h \{\mathbf{u} \in B : d_H(\mathbf{u}, \mathbf{u}_i) < d_i\}$ . If  $\mathbf{u}_{\text{best}}$  satisfies the following condition,

$$L(\mathbf{u}_{\text{best}}) \leq \underline{L}(\mathbf{u}_1, d_1; \mathbf{u}_2, d_2; \dots; \mathbf{u}_h, d_h), \quad (2)$$

then  $\mathbf{u}_{\text{best}}$  is the optimal MLD solution in  $B$ .  $\triangleleft$

Expressions for evaluating the right-hand side of (2) for  $h = 1, 2$  and  $3$  have been derived [3]. These bounds are used in our proposed iterative decoding algorithm for early termination conditions.

### 3 Decoding Procedure

Let  $C$  be a binary linear  $(N, K)$  code with weight profile  $W = \{0, w_1, w_2, \dots\}$ , and let  $C_0$  be a binary linear  $(N, K_0)$  subcode of  $C$  with weight profile  $W_0 = \{0, w_{01}, w_{02}, \dots\}$ , where  $w_1$  is the minimum weight of  $C$ ,  $w_{01}$  is that of  $C_0$  and  $w_i \leq w_{0i}$ . As a special case,  $C_0$  may be  $C$  itself.

#### 3.1 Minimum-weight Subtrellis Search for a Coset

For  $B \in C/C_0$  and  $\mathbf{u} \in B$ , minimum-weight search  $(\mathbf{u}, B)$  denotes a search procedure for find-

ing a codeword in  $B$ , denoted  $\varphi_B(\mathbf{u})$ , which has the least correlation discrepancy with respect to  $\mathbf{z}$  among all codewords in  $B$  at the minimum Hamming distance  $w_{01}$  from  $\mathbf{u}$ . That is,

$$\varphi_B(\mathbf{u}) = \text{best}\{v \in B : d_H(v, \mathbf{u}) = w_{01}\}. \quad (3)$$

If  $C_0$  is spanned by the set of minimum weight codewords of  $C$ , that is,  $C_0 = C$  or  $C$  is an UEP(unequal error protection) code, then  $w_{01} = w_1$  and

$$\varphi_B(\mathbf{u}) = \varphi_C(\mathbf{u}) \triangleq \text{best}\{v \in C : d_H(v, \mathbf{u}) = w_1\}. \quad (4)$$

This search procedure is implemented by using the minimum-weight subtrellis of  $C_0$  around  $\mathbf{u}$ . This minimum-weight subtrellis is sparsely connected and much simpler than the full trellis of the code[2, 6].

Iterative minimum-weight search  $(\mathbf{u}, B)$  is to generate a sequence of candidate codewords,  $\varphi_B(\mathbf{u}), \varphi_B(\varphi_B(\mathbf{u})), \dots$ , until a certain termination condition holds. It is shown in [6] that

$$L(\varphi_B^{i+2}(\mathbf{u})) \leq L(\varphi_B^i(\mathbf{u})), \text{ for } i \geq 0, \quad (5)$$

where  $\varphi_B^0(\mathbf{u}) \triangleq \mathbf{u}$  and  $\varphi_B^{i+1}(\mathbf{u}) \triangleq \varphi(\varphi_B^i(\mathbf{u}))$  for  $i \geq 0$ , and that if  $L(\varphi_B^{i+2}(\mathbf{u})) < L(\varphi_B^i(\mathbf{u}))$  for  $0 \leq i \leq I$ , then  $\varphi_B^i(\mathbf{u})$  with  $1 \leq i \leq I$  are all different. If  $j$  is the smallest index such that

$$L(\varphi_B^{j-2}(\mathbf{u})) = L(\varphi_B^j(\mathbf{u})), \quad (6)$$

then

$$\min\{L(\varphi_B^i(\mathbf{u})) : 0 \leq i \leq j\} = \min\{L(\varphi_B^{j-2}(\mathbf{u})), L(\varphi_B^{j-1}(\mathbf{u}))\}. \quad (7)$$

The condition (6), denoted  $\text{Cond}_R$ , is used as one of termination conditions for Iterative minimum-weight search  $(\mathbf{u}, B)$  to avoid repetition.

#### 3.2 Decoding Procedure for $C$

Suppose the set of codewords of weight  $w_{01}$  in  $C_0$  spans  $C_0$  and  $K - K_0$  is not large. We propose a new decoding procedure for  $C$  which consists of iterated minimum-weight searches in a coset and coset shiftings. Two early termination conditions,  $\text{Cond}_C$  for the entire procedure and  $\text{Cond}_B$  in the subprocedure for a coset  $B \in C/C_0$ , are used besides the termination condition  $\text{Cond}_R$  in the subprocedure for a coset.

$\text{Cond}_C$  is a sufficient condition [3] that the best candidate codeword, denoted  $\mathbf{u}_{\text{best}}$ , in the set of those candidate codewords which have been generated already, denoted  $GC$ , is optimum based on  $\text{best}_h GC$  and the weight profile  $W$  of  $C$ , where  $h$

is a specified small integer. From the Lemma,  $\text{Cond}_C$  is defined as

$$L(\mathbf{u}_{\text{best}}) \leq \underline{L}(\mathbf{u}_1, d_1; \mathbf{u}_2, d_2; \dots; \mathbf{u}_l, d_l), \quad (8)$$

where  $l = \min\{h, |GC|\}$  and for  $1 \leq i \leq l$ ,  $\mathbf{u}_i \in \text{best}_h GC$  and if  $\varphi_C(\mathbf{u}_i) \in GC$ , then  $d_i = w_2$  and otherwise,  $d_i = w_1$ .

$\text{Cond}_B$  is a sufficient condition that there remain no codewords in  $B$  better than  $\mathbf{u}_{\text{best}}$ . This condition is also based on  $\text{best}_h GC_B$  where  $GC_B = GC \cap B$  and the weight profile  $W_0$  of  $C_0$  which is the same as the distance profile of any coset of  $C/C_0$ . From the Lemma,  $\text{Cond}_B$  is defined as

$$L(\mathbf{u}_{\text{best}}) \leq \underline{L}(\mathbf{u}_1, d_1; \mathbf{u}_2, d_2; \dots; \mathbf{u}_{l_B}, d_{l_B}), \quad (9)$$

where  $l_B = \min\{h, |GC_B|\}$  and for  $1 \leq i \leq l_B$ ,  $\mathbf{u}_i \in \text{best}_h GC_B$  and if  $\varphi_B(\mathbf{u}_i) \in GC_B$ , then  $d_i = w_{02}$  and otherwise,  $d_i = w_{01}$ . For instance, if the minimum or the second minimum weight of  $C_0$  is greater than that of  $C$ , then  $\text{Cond}_B$  is more effective than  $\text{Cond}_C$  only.

In the procedure, global variables  $GC_h$  and  $GC_{B,h}$  are used besides  $GC$  and  $\mathbf{u}_{\text{best}}$ .  $GC_h$  denotes the current value of  $\text{best}_h GC$  and  $GC_{B,h}$  denotes the current value of  $\text{best}_h GC_B$ . For  $B \in C/C_0$ , let  $f(r, B)$  denote the initial candidate codeword in  $B$  for a given received sequence.

### 3.2.1 Decoding Algorithm II

We assume that  $z \notin C$ .

(D1) (i) Generate  $f(r, B)$  for all  $B \in C/C_0$ , and number the  $2^{K-K_0}$  cosets of  $C/C_0$  in the increasing order of correlation discrepancy,  $L(f(r, B))$ .

(ii) Initialize  $GC \leftarrow \{f(r, B) : B \in C/C_0\}$ ,  $\mathbf{u}_{\text{best}} \leftarrow \text{best} GC$  and  $GC_h \leftarrow \text{best}_h GC$ . If  $\text{Cond}_C$  holds, then output  $\mathbf{u}_{\text{best}}$  and stop. Otherwise, initialize  $GC_B, GC_{B,h} \leftarrow \{f(r, B)\}$  for  $B \in C/C_0$ , and perform Search-in(the first coset).

(D2) Search-in( $B$ ): Execute Iterative minimum-weight search( $f(r, B), B$ ) together with updating the global variables each time a new candidate codeword is generated until either  $\text{Cond}_R, \text{Cond}_C$  or  $\text{Cond}_B$  holds. If  $\text{Cond}_C$  holds, then output  $\mathbf{u}_{\text{best}}$  and stop. Suppose that  $\text{Cond}_R$  or  $\text{Cond}_B$  holds. If all cosets have been exhausted, then output  $\mathbf{u}_{\text{best}}$  and stop; and otherwise, call Search-in(the coset next to  $B$ ).

(D3) If either  $\text{Cond}_C$  or  $\text{Cond}_B$  holds for every coset in  $C/C_0$ , the output is optimum.

For the special case where  $C = C_0$ , this decoding algorithm becomes Algorithm I- $w_1$ .

### 3.3 Choice of the Initial Candidate Codeword $f(r, B)$ for $B \in C/C_0$

For a given received sequence  $\mathbf{r} = (r_1, r_2, \dots, r_N)$ , let  $M_K$  be the location set of the most reliable

basis of the column space of a generator matrix of  $C$ , and let  $\Lambda_{N-K_0}$  be the location set of the least reliable basis for the column space of a parity-check matrix of  $C_0$ . Then it follows from Theorem 1 in [5] that

$$|M_K \cap \Lambda_{N-K_0}| = K - K_0. \quad (10)$$

For  $\mathbf{x} = (x_1, x_2, \dots, x_N) \in V_N$  and a coset  $B \in C/C_0$ , a codeword  $\mathbf{u} = (u_1, u_2, \dots, u_N) \in B$  satisfying the following condition is uniquely determined:

$$u_i = x_i, \text{ for all } i \in M_K - \Lambda_{N-K_0}. \quad (11)$$

Let  $g(\mathbf{x}, B)$  denote the above codeword  $\mathbf{u}$  in  $B$ .

Then  $g(\mathbf{z}, B)$  can be chosen as the initial candidate codeword  $f(\mathbf{r}, B)$  for  $B = C/C_0$ , where  $\mathbf{z}$  is the hard-decision binary vector obtained from the received sequence  $\mathbf{r}$ . This  $g(\mathbf{z}, B)$  for  $B = C/C_0$  is a simple generalization of the zero-th order decoding proposed by Fossorier and Lin[4] to a coset of  $C_0$  in  $C$ . Similarly, a generalization of their first order decoding can be used.

## 4 Examples

**Example 1:** Let  $C = \text{EBCH}(64, 24)$  and  $C_0 = \text{RM}(64, 22)$ . Then  $w_1 = 16$ ,  $w_2 = 18$ ,  $w_{01} = 16$  and  $w_{02} = 24$ . Decoding Algorithm II, where early termination condition  $\text{Cond}_B$  is not used, has been simulated for this code. The simulation results are shown in Figures 1 and 2. For comparison, the simulation results of Algorithm I- $w_1-w_2$  for this code are also shown. Figure 1 shows the bit error probabilities. We see that Algorithm II practically achieves optimal error performance. Figure 2 shows the average numbers of operations(addition and comparison) of Algorithm I- $w_1-w_2$  and Algorithm II. The numbers depend on the complexities of subtrellises used in the simulation. The minimum weight subtrellis of the  $\text{RM}(64, 22)$  code used in Algorithm II and the first and second weight subtrellis of the  $\text{EBCH}(64, 24)$  code used in Algorithm I- $w_1-w_2$  are obtained simply by purging the 4-section full trellis diagrams of the codes. The construction of better subtrellis is under study.

We see that the average number of operations of Algorithm II can be reduced by using  $\text{Cond}_B$ . **Example 2:** Let  $C = C_0 = \text{EBCH}(64, 45)$  with  $w_1 = 8$  and  $w_2 = 10$ . For this case, the simulation results of Algorithm I- $w_1$ , where the initial candidate codeword is provided by the first order decoding in [5], are shown in Figures 3 and 4.

Simulation results of the  $\text{RM}(64, 22)$  and  $\text{RM}(64, 42)$  codes are shown in [6].

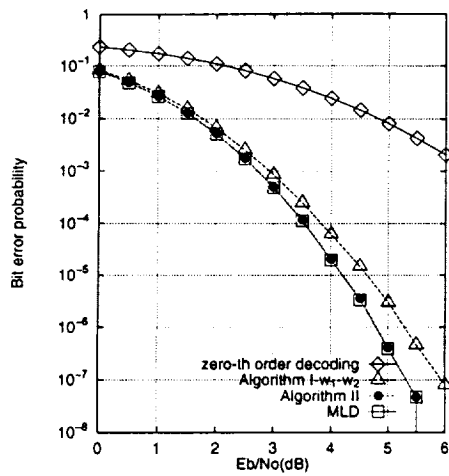


Figure 1: Bit error probabilities for the EBCH(64,24).

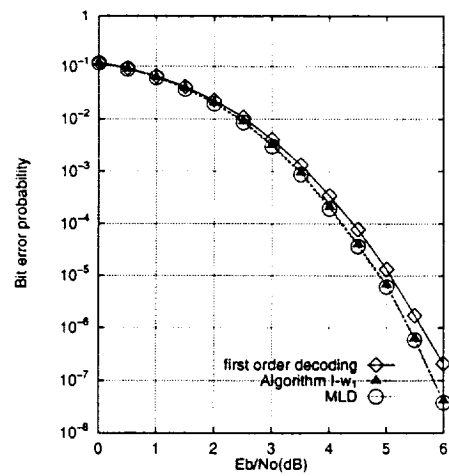


Figure 3: Bit error probabilities for the EBCH(64,45).

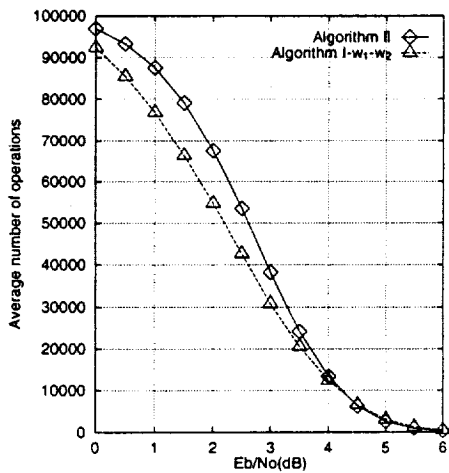


Figure 2: Average numbers of operations for EBCH(64,24).

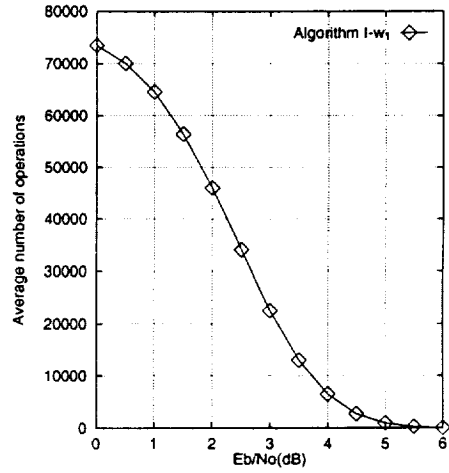


Figure 4: Average numbers of operations for EBCH(64,45).

## References

- [1] D. Chase, "A New Class for Decoding Block Codes with Channel Measurement Information," *IEEE Trans. Information Theory*, Vol. IT-18, pp.170-182, Jan. 1972.
- [2] H. T. Moorthy, S. Lin and T. Kasami, "Soft-Decision Decoding of Binary Linear Block Codes Based on an Iterative Search Algorithm," submitted to *IEEE Trans. Information Theory*, 1995.
- [3] T. Kasami, T. Takata, T. Koumoto, T. Fujiwara, H. Yamamoto and S. Lin, "The Least Stringent Sufficient Condition on Optimality of Suboptimal Decoded Codewords," *Technical Report of IE-ICE*, IT-94-82, The Inst. of Electronics, Information and Communication Engineers, Japan, Jan. 1995.
- [4] M.P.C.Fossorier and S.Lin, "Soft Decision Decoding of Linear Block Codes Based on Ordered Statistics," *IEEE Trans. Information Theory*, Vol.IT-41, pp.1217-1234, Sept. 1995.
- [5] M.P.C.Fossorier, S.Lin and J.Snyders, "On Maximum Likelihood Soft Decision Syndrome Decoding," *Proc. of IEEE International Symposium on Information Theory*, p.415, Sept. 1995.
- [6] T. Koumoto, H. Nagano, T. Takata, T. Fujiwara, T. Kasami and S. Lin, "An Iterative Soft-Decision Decoding Algorithm," *Proc. of the 18th Symp. on Information Theory and Its Applications*, pp.557-560, Oct. 1995.



**PART II**

**BOUNDS ON BLOCK ERROR  
PROBABILITY FOR  
MULTILEVEL CONCATENATED  
CODES**

**Hari T. Moorthy, Diana Stojanovic and Shu Lin**



# Bounds on Block Error Probability for Multilevel Concatenated Codes

Hari T. Moorthy, Diana Stojanovic and Shu Lin<sup>1</sup>

Department of Electrical Engineering,  
University of Hawaii,  
Honolulu, Hawaii 96822  
Email: harit@spectra.eng.hawaii.edu  
dixi@spectra.eng.hawaii.edu

48  
000094

*Abstract* — Maximum likelihood decoding of long block codes is not feasible due to large complexity. Some classes of codes are shown to be decomposable into multilevel concatenated codes (MLCC) [3]. For these codes, multistage decoding provides good trade-off between performance and complexity. In this paper, we derive an upper bound on the probability of block error for MLCC. We use this bound to evaluate difference in performance for different decompositions of some codes. Examples given show that a significant reduction in complexity can be achieved when increasing number of stages of decoding. Resulting performance degradation varies for different decompositions. A guideline is given for finding good m-level decompositions.

## I. INTRODUCTION

Multistage Decoding (MSD) is an efficient soft decision decoding method for long decomposable codes, such as multilevel concatenated codes (MLCC) [1]. Although suboptimum in performance, it greatly reduces the computational complexity as compared to optimum decoding.

Since its inception in Hemmati's paper [2], Closest Coset Decoding (CCD) of  $|u|u + v|$  codes and generalizations of CCD have been investigated by several authors. However, only a few papers on the performance analysis of this method have appeared [3-6]. Furthermore, these methods are restricted to  $|u|u + v|$  codes and some others to block modulation codes. The different code-structure of MLCC's precludes application of the analysis therein to the case of MLCCs.

In [3], an upper bound on the effective error coefficient (EEC) for 2-stage decoding of MLCC's was derived and some guidelines for choosing a good 2-level decomposition of Reed Muller codes were given.

In this paper, we first derive an upper bound on block error probability of MSD of MLCC's, when optimum decoding of each stage is performed. We first express the upper bound in terms of all the error coefficients, and then explain how these coefficients can be obtained using some combinatorial methods and weight

distributions of inner and outer codes of a MLCC. The bound enables prediction of the performance of MSD (two or more levels) without simulation. Therefore an estimate of performance degradation, when the number of decoding stages is increased, can be obtained.

## II. ANALYSIS OF CCD OF MULTILEVEL CONCATENATED CODES

For an  $M$ -level code  $C = \{B_1, B_2, \dots, B_M\} * \{A_1, A_2, \dots, A_M\}$ , consider the  $i$ -th stage of decoding for  $1 \leq i \leq M$ . At each stage, decoding is assumed to be complete (an estimate of a transmitted codeword is always given) soft decision maximum likelihood closest coset decoding [5].

Let  $P_e^{(i)}$  denote the probability that the  $i$ -th stage decoder makes an error when all previous  $(i-1)$  stages are correct. Then, conditioning on the  $i$ -th stage decoding being in error, we obtain  $P_e = \sum_{i=1}^M P_e^{(i)}$ . In general,  $P_e^{(i)}$  depends on the codeword at the output of the  $i$ -th inner encoder. However, for binary linear codes, BPSK transmission and the AWGN channel,  $P_e^{(i)}$  is the same for all transmitted codewords [7]. For simplicity, we assume that the all-zero codeword is transmitted. In this case, the output of each  $B_i$ -encoder at the transmitter is the all-zero codeword in  $B_i$ . The received  $(n_A n_B)$ -tuple  $\mathbf{r} = (r^{(1)}, r^{(2)}, \dots, r^{(n_B)})$  is sectionalized into  $n_B$  sections each of length  $n_A$ , where  $n_B$  and  $n_A$  are lengths of outer and inner code respectively.

An error at the  $i$ -th stage, when CCD is used, occurs if a nonzero codeword  $\mathbf{b} = (b_1, b_2, \dots, b_{n_B}) \in B_i$  has larger correlation than the all-zero codeword  $\mathbf{z} \in B_i$ . Let  $\mathbf{a}^j = (a_1^j, a_2^j, \dots, a_{n_A}^j)$  be the codeword with the best correlation in the coset corresponding to  $j$ -th symbol  $b_j$ , for  $j = 1, 2, \dots, n_B$ . If a standard BPSK mapping ( $0 \rightarrow 1, 1 \rightarrow -1$ ) is applied, then the difference in metrics of  $\mathbf{b}$  and  $\mathbf{z}$  can be bounded by the following expression:

$$M(\mathbf{b}) - M(\mathbf{z}) \leq \sum_{j=1, b_j \neq 0}^{n_B} \sum_{l=1, a_l^j = 1}^{n_A} (-2r_l^j), \quad (1)$$

where  $M(\cdot)$  denotes correlation metric. Thus the probability of error at the  $i$ -th stage is given by

$$P_e^{(i)} = P(M(\mathbf{b}) - M(\mathbf{z}) > 0, \mathbf{b} \in B_i, \mathbf{b} \neq \mathbf{0}). \quad (2)$$

<sup>1</sup>This work was supported by NSF grant NCR-9415374 and NASA grant NAG 5-931.

For a particular  $\mathbf{a}$  and  $\mathbf{b}$ ,

$$T(\mathbf{a}, \mathbf{b}) = \sum_{j=1, b_j \neq 0}^{n_B} \sum_{l=1, a_l' = 1}^{n_A} (-2r_l^j)$$

is a Gaussian random variable with mean  $\mu$  and variance  $\sigma^2$  given by

$$\mu = 2A \sum_{b_j \neq 0} w(\mathbf{a}), \quad \sigma^2 = 4 \sum_{b_j \neq 0} w(\mathbf{a})$$

where  $A$  is the amplitude of a transmitted BPSK signal (SNR=10 log  $(A^2/2 k/n)$ ). It then follows that

$$\begin{aligned} P^{(i)} &\leq \sum_{\mathbf{b} \in \mathbf{B}_i} \sum_{b_j \neq 0} \sum_{\mathbf{a} \in \Omega(b_j)} P(T(\mathbf{a}, \mathbf{b}) > 0) \\ &= \sum_{\mathbf{b} \in \mathbf{B}_i} \sum_{b_j \neq 0} \sum_{\mathbf{a} \in \Omega(b_j)} \frac{1}{2} \operatorname{erfc} \left( A \sqrt{\frac{1}{2} \sum_{b_j \neq 0} w(\mathbf{a})} \right), \quad (3) \end{aligned}$$

where  $\Omega(b_j)$  represents the coset corresponding to the symbol  $b_j$ , and  $w(\cdot)$  the weight of a codeword.

Note that for any  $0 \neq \mathbf{b} \in \mathbf{B}_i$   $0 \neq \mathbf{a} \in \mathbf{A}_i$ ,  $\sum_{b_j \neq 0} w(\mathbf{a}) \geq d_{\min}(C)$

**Definition 1:** Define  $S_i(Y)$  as the number of terms of the form  $\frac{1}{2} \operatorname{erfc} \left( A \sqrt{Y/2} \right)$  satisfying

1.  $Y = \sum_{b_j \neq 0} w(\mathbf{a})$
2.  $\mathbf{b} \in \mathbf{B}_i, \mathbf{a} \in \Omega(b_j)$

The union bound on the probability of error at the  $i$ -th stage of decoding is then given by

$$P_e^{(i)} < \sum_Y S_i(Y) \frac{1}{2} \operatorname{erfc} \left( A \sqrt{Y/2} \right) \quad (4)$$

where the summation is taken over all  $Y$  defined above. Thus

$$P_e < \sum_i \sum_Y S_i(Y) \frac{1}{2} \operatorname{erfc} \left( A \sqrt{Y/2} \right). \quad (5)$$

$S_i(Y)$  depends on the symbol weight distribution of the  $i$ -th outer code  $\mathbf{B}_i$ , the Hamming weight distribution of the MLC code  $\mathbf{C}$  and critically on the weight distribution (of  $n_A$ -tuples) in the cosets of the partition  $\mathbf{A}_i/\mathbf{A}_{i+1}$ . For inner codes of small lengths the coset weight distribution can be evaluated easily by an exhaustive computer search program or using combinatorial methods.

### III. RESULTS FOR SOME MLCC CODES

Consider the decomposition

$$\text{RM}(64, 42, 16) = \{(8, 1, 3)(8, 4, 4)^3, (8, 7, 2)^3(8, 8, 1)\} * \{(8, 8, 1), (8, 4, 4), (8, 0)\}$$

given in [3]. The symbol weight distribution of  $B_1$  is  $N_{B_1}(0) = 1, N_{B_1}(4) = 98, N_{B_1}(6) = 1, 176, N_{B_1}(7) = 1, 344, N_{B_1}(8) = 5, 573$ . The symbol weight distribution of  $B_2$  is  $N_{B_2}(0) = 1, N_{B_2}(1) = 8, N_{B_2}(2) = 812, N_{B_2}(3) = 23, 576, N_{B_2}(4) = 443, 030, N_{B_2}(5) = 5, 315, 576, N_{B_2}(6) = 39, 867, 212, N_{B_2}(7) = 170, 859, 368, N_{B_2}(8) = 320, 361, 329$ . The coset weight distributions of  $\mathbf{A}_1/\mathbf{A}_2$  and  $\mathbf{A}_2/\{0\}$  are given in Table 1.

To compute  $S_i(Y)$  with  $i = 1, 2$ , we have to determine how many  $n_A n_B$ -tuples of weight  $Y$  can result from a  $\mathbf{b} \in \mathbf{B}_1$  of symbol weight  $w_1(\mathbf{B}_1) = X$ . In order to answer this question, we must see if a partition of the integer  $Y$  into  $X$  parts is possible with each part belonging to a valid entry in the coset weight distribution of the partition  $(8, 8, 1)/(8, 4, 4)$ . A partition of the form  $8 = 2 + 3 + 1 + 2$  is invalid because none of the non-zero cosets in the above partition have a vector of Hamming weight 1 or 3. Furthermore, the partitions must be counted in an ordered fashion (i.e., we are counting the so called distributions of an integer). For example, there are 12 distributions of 16 into parts  $\{4, 4, 2, 6\}$ . The partition table for first two nonzero weights in both first and second stage are given in Table 2.

The effective error coefficient in stage-1 is determined as  $S_1(8) = 98 \times 256 = 25,088$ . The next two error coefficients are  $S_1(10) = 98 \times 4 \times 4^3 \times 8 = 200704$  and  $S_1(12) = 98 \times (6 \times 8^2 \times 4^2 + 4^4) + 4^6 \times 1176 = 5444096$ . The higher order error coefficients can be computed in a similar manner.

In the second stage of decoding, the effective error coefficient is  $S_2(8) = 8 + 784 = 792$  and the error coefficient  $S_2(12) = 18,816$ . The first three error coefficients for this 2-level decomposition of RM  $(64, 42, 8)$  corresponding to weights 8, 10 and 12 are  $S_1(8) + S_2(8) = 25,880$ ,  $S_1(10) = 200,704$  and  $S_1(12) + S_2(12) = 5,462,912$ .

In summary, for the Reed-Muller code  $\mathbf{C} = (64, 42, 8)$ , an upper bound (with first two significant terms) on the block error probability with two-stage decoding is

$$P_{e,2\text{-stage}} < 2.5 \cdot 10^4 \frac{1}{2} \operatorname{erfc} \left( A \sqrt{8/2} \right) + 2.0 \cdot 10^5 \frac{1}{2} \operatorname{erfc} \left( A \sqrt{10/2} \right) + 5.462,912 \frac{1}{2} \operatorname{erfc} \left( A \sqrt{12/2} \right).$$

The complexity of stage-2 can be reduced by employing a 3-stage decomposition

$$\begin{aligned} \text{RM}(64, 42, 8) = \\ \{(8, 1, 8)(8, 4, 4)^3, (8, 7, 2)^2, (8, 7, 2)(8, 8, 1)\} * \\ \{(8, 8, 1), (8, 4, 4), (8, 2, 4), (8, 0)\} \quad (6) \end{aligned}$$

for which

$$P_{e,3-stage} < 2.6 \cdot 10^4 \frac{1}{2} \operatorname{erfc} \left( A\sqrt{8/2} \right) + 2.0 \cdot 10^5 \frac{1}{2} \operatorname{erfc} \left( A\sqrt{10/2} \right) + 5.4 \cdot 10^6 \frac{1}{2} \operatorname{erfc} \left( A\sqrt{12/2} \right).$$

Notice that the asymptotic loss in decoding gain, due to this simplification, predicted from the bound is practically zero, and this agrees with actual simulation results shown in Figure 1.

Table 1

Coset-#'s	Weight of vector				
	0	2	4	6	8
0	1	-	14	-	1
1-7	-	4	8	4	-

Table 2

$w_s(b)$	Partition	$N_1$	$N_2$
4	{2, 2, 4, 4}	6	$4^2 8^2 6$
4	{2, 2, 2, 6}	4	$4^3 4$
4	{2, 2, 2, 4}	4	$4^3 8$
5	Not possible	-	-
7	Not possible	-	-
6	{2, 2, 2, 2, 2, 2}	1	$4^6$

$N_1$  = Number of partitions.  $N_2$  = Number of code-words in  $B_i * [A_i/A_{i+1}]$  per partition.

For long codes, even 2-stage CCD cannot be performed in an optimal manner due to the complexity of the trellises involved. Hence it is imperative that tight bounds be derived so that the loss of coding gain at high SNRs due to  $m$ -stage ( $m \geq 3$ ) can be predicted.

The loss in coding gain of 2-stage relative to  $m$ -stage decoding can be determined based on the error coefficients derived in the paper. For the code  $C = (128, 64, 16)$  RM, and  $n = 2, 3$ , this is shown in Figure 2.

The corresponding expressions are

$$P_{e,2-stage} < 13.8 \cdot 10^6 \frac{1}{2} \operatorname{erfc} \left( A\sqrt{16/2} \right) + 5.1 \cdot 10^9 \frac{1}{2} \operatorname{erfc} \left( A\sqrt{24/2} \right)$$

$$P_{e,3-stage} < 1.6 \times 10^9 \frac{1}{2} \operatorname{erfc} \left( A\sqrt{16/2} \right) + 6.1 \times 10^{13} \frac{1}{2} \operatorname{erfc} \left( A\sqrt{24/2} \right)$$

The three stage decomposition has a trellis complexity of 512, 128, and 512 states for the first, second, and third stage, and the average number of operations per information bit is 1263. The most of the computational complexity is due to a large and fully connected trellis of the third stage, so a way to reduce it is to implement 4-stage decoding by dividing the third stage into two stages.

The difference between the bounds for 2-stage and 3-stage at block error rate  $10^{-6}$  is about 0.6 dB. Since the simulation of 2-stage decoding with MLD at each stage is not available, we compared the results of 3-stage decoding with a 2-stage decoding where the first stage is decoded using iterative search. We believe this would be very close in performance of optimum 2-stage decoding. The difference in performance for this 2-stage and optimum 3-stage decoding is 0.6 dB as well.

The third example is done for the (72, 52, 8) decomposable code. Bounds for 3 and 4-stage decoding of this code are given in Figure 3.

The state complexities are 2, 512, 8, 1 for the 4-stage decoding, and {1024, 8, 1} for the 3-stage decoding.

#### IV. TRADE-OFFS BETWEEN COMPLEXITY AND PERFORMANCE

We have shown in the previous section that the performance of 2 and 3-stage decoding of (64, 42, 8) code are almost identical. However, this is true only for 3-stage decomposition (6). Reduction in state complexity is not significant, since maximum number of states in the first stage remains 128 (the same as for 2-stage decomposition) and the maximum numbers of states at the second and third stage are 4 and 2 respectively (as compared to 8 states in 2-stage decomposition). The average number of operations per information bit, though, is reduced from 393 to 133 (almost three times).

For another 3-stage decomposition,

$$\begin{aligned} \text{RM}(64, 42, 16) = \\ \{(8, 1, 8)(8, 4, 4), (8, 4, 4)^2, (8, 7, 2)^3(8, 8, 1)\} * \\ \{(8, 8, 1), (8, 6, 4), (8, 4, 4), (8, 0)\}, \quad (7) \end{aligned}$$

maximum number of states is reduced to 16 (in the second stage), so 8 times compared to two stage decomposition. Unfortunately, this decomposition loses about 0.5 dB in performance, and the average number of operations per information bit is 227. Thus the 3-stage decomposition (6) outperforms decomposition (7), while having smaller computational complexity.

We can see that the performance of the multistage scheme depends mainly on the first decoding stage. This suggests that when increasing the number of stages, lower level stages should be divided, while the first should be kept as long as the number of states in the outer code trellis is not too big. In such manner, the performance degradation will not be large (in some cases almost negligible). Since lower level codes in multilevel decomposition are high rate codes with complex trellis structure, splitting each of them into two levels will reduce the complexity significantly (as we have seen in the example of RM(64,42,8) code). This can be used to determine a good decomposition of a code.

## V. CONCLUSION

In this paper, we have shown how to bound the performance of multistage decoding of Multilevel concatenated codes. We presented examples which show that the simulation results converge to derived bounds at moderate to high signal-to-noise ratios. Therefore, an estimate of the difference in performance for different decompositions and different numbers of decoding stages can be determined. Analysis of reduction in complexity when using large number of decoding stages has also been discussed.

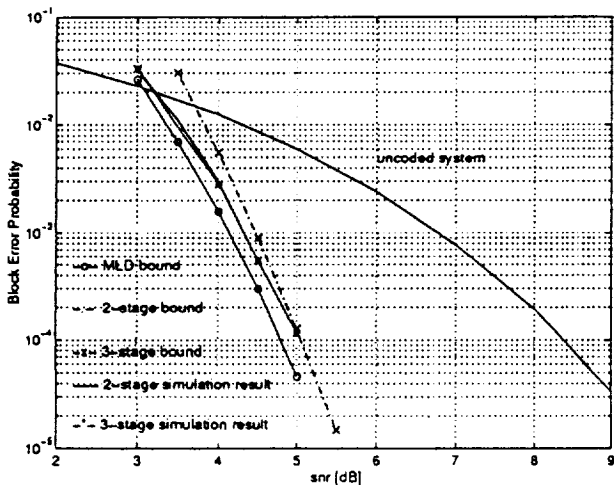


Figure 1: Simulation results and performance bounds for RM(64,42,8) code

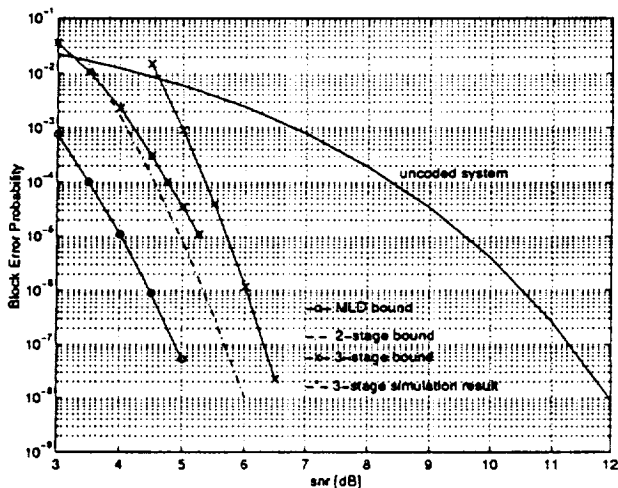


Figure 2: Simulation results and performance bounds for RM(128,64,16) code

## REFERENCES

- [1] V. A. Zinoviev, "Generalized Concatenated Codes," *Problemy Peredachi Informatsii*, Vol. 12., No. 1, pp. 5-15, 1976.
- [2] F. Hemmati, "Closest Coset Decoding of  $|u|u+v|$  codes," *IEEE J. Selected Areas Communications*, Vol. 7, pp. 982-988, Aug. 1989.
- [3] J. Wu and S. Lin, "An Upper Bound on Effective Error Coefficient of Two-Stage Decoding, and Good Two-Level Decompositions of Some Reed-Muller Codes," *IEEE Transactions on Communications*, Vol. 42, No. 3, March 1994.
- [4] M.P.C. Fossorier and S. Lin, "Weight Distribution for Closest Coset Decoding of  $|u|u+v|$  Constructed Codes," submitted to *IEEE Transactions on Information Theory*, in revision. November 1995.
- [5] T. Takata et. al., "An Improved Union Bound on Block Error Probability for Closest Coset Decoding," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E78-A, No. 9, September 1995.
- [6] T. Takata, Y. Yamashita, T. Fujiwara, T. Kasami and S. Lin, "Suboptimum Decoding of Decomposable Block Codes," *IEEE Transactions on Information Theory*, Vol. 40, No. 5, Sep. 1994.
- [7] T. Takata, S. Ujita, T. Kasami and S. Lin, "Multistage Decoding of Multilevel Block M-PSK Modulation Codes and Its Performance Analysis," *IEEE Transactions on Information Theory*, Vol. 39, No. 4, July 1993.

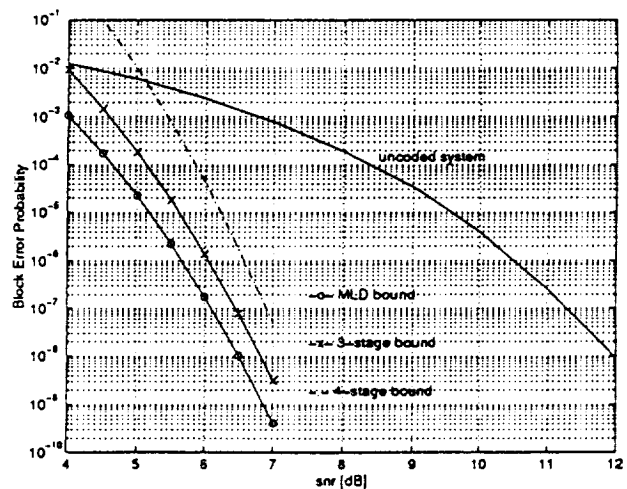


Figure 3: Performance bounds for (72,52,8) decomposable code



**PART III**

**BIT ERROR PROBABILITY FOR  
MAXIMUM LIKELIHOOD DECODING  
OF LINEAR BLOCK CODES**

**Marc P.C. Fossorier, Shu Lin and Dojun Rhee**



# Bit Error Probability for Maximum Likelihood Decoding of Linear Block Codes

Marc P.C. Fossorier<sup>1</sup>, Shu Lin<sup>1</sup> and Dojun Rhee<sup>2</sup>

<sup>1</sup> Dept. of Electrical Engineering University of Hawaii at Manoa Honolulu, HI 96822, USA.

<sup>2</sup> LSI LOGIC Corporation, 1525 Mc-Carthy Blvd, MS G-815, Milpitas, CA 95035, USA

Email: marc@wiliki.eng.hawaii.edu

*Abstract* — In this paper, the bit error probability  $P_b$  for maximum likelihood decoding of binary linear codes is investigated. The contribution of each information bit to  $P_b$  is considered. For randomly generated codes, it is shown that the conventional approximation at high SNR  $P_b \approx (d_H/N) \cdot P_s$ , where  $P_s$  represents the block error probability, holds for systematic encoding only. Also systematic encoding provides the minimum  $P_b$  when the inverse mapping corresponding to the generator matrix of the code is used to retrieve the information sequence. The bit error performances corresponding to other generator matrix forms are also evaluated. Although derived for codes with a generator matrix randomly generated, these results are shown to provide good approximations for codes used in practice. Finally, for decoding methods which require a generator matrix with a particular structure such as trellis decoding or algebraic-based soft decision decoding, equivalent schemes that reduce the bit error probability are discussed.

## I. INTRODUCTION

In this paper, we consider the minimization of the bit error probability  $P_b$  for maximum likelihood decoding (MLD) of linear block codes. Although not optimum, this minimization remains important as MLD has been widely used in practical applications. We assume that the information sequence of length  $K$  is recovered from the decoded codeword based on the inverse mapping defined from the generator matrix of the code. For block codes, the large error coefficients can justify this strategy which is explicitly or implicitly used in many decoding methods such as conventional trellis decoding, multi-stage decoding or majority-logic-decoding. Therefore, for a particular code and the same optimal block error probability, we determine the best encoding method for delivering as few erroneous information bits as possible whenever a block is in error at the decoder output. We first derive a general upper bound on  $P_b$  which applies to any generator matrix and is tight at medium to high signal to noise ratio (SNR). This bound considers the individual contribution of each information bit separately. For randomly

generated codes, we then show that the systematic generator matrix (SGM) provides the minimum bit error probability. To this end, a submatrix of the generator matrix defining an equivalent code for the bit considered is introduced. Note that a similar general result holds for the optimum bit error probability related to the BSC [2]. We finally discuss how to achieve this performance whenever the systematic encoding is not the natural choice, as for trellis decoding [3] or for MLD in conjunction with algebraic decoding [4]-[8]. For example, for trellis decoding of the (32,26,4) Reed-Muller (RM) code, at low SNR a performance degradation of more than 1 dB is recovered with the proposed method. Minimizing the bit error probability associated with MLD becomes even more important whenever the considered block code is used as the inner code of a concatenated coding system [9].

## II. BIT ERROR PROBABILITY FOR MLD

Suppose an  $(N, K, d_H)$  binary linear code  $\mathcal{C}$  with generator matrix  $G$  is used for error control over the AWGN channel. Defining

$$P_b = \frac{1}{K} \sum_{j=1}^K P_b(j), \quad (1)$$

where  $P_b(j)$  represents the error probability for the  $j^{\text{th}}$  bit in a block of  $K$  information bits delivered by the decoder, we obtain from the union bound

$$P_b(j) \leq \sum_{i=d_H}^N \tilde{w}_i(j) \tilde{Q}(\sqrt{i}), \quad (2)$$

where  $\tilde{Q}(x) = (\pi N_0)^{-1/2} \int_x^\infty e^{-n^2/N_0} dn$ . We call  $\tilde{w}_i(j)$  the effective error coefficient associated with the  $j^{\text{th}}$  information bit with respect to the generator matrix  $G$ .

We can prove the following theorem.

**Theorem 1** Let  $w_i$  represent the number of codewords of weight  $i$  in the code  $\mathcal{C}$  generated by  $G$  and let  $w_i(j)$  represent the number of codewords of weight  $i$  in the subcode generated by the matrix  $G(j)$  obtained after deleting row- $j$  in  $G$ ; then

$$\tilde{w}_i(j) = w_i - w_i(j). \quad (3)$$

Theorem 1 depends on the mapping defined by  $G$  as it implicitly assumes that the inverse mapping corresponding to  $G$  is used to retrieve the information bits

<sup>2</sup>Supported by the NSF Grant NCR-94-15374 and the NASA Grant NAG-5-931

from the decoded code sequence. Since for a linear code, this mapping is one-to-one and thus invertible, Theorem 1 is valid for any representation of  $G$ , systematic as well as non-systematic. Combining (1) and (2), the average bit error probability is expressed as

$$P_b \leq \sum_{i=d_H}^N \left( \frac{1}{K} \sum_{j=1}^K \tilde{w}_i(j) \right) \tilde{Q}(\sqrt{i}), \quad (4)$$

For a code defined by a matrix  $G$  randomly generated, we associate with each information bit  $j \in [1, K]$  a matrix

$$D_\alpha(j) = \begin{bmatrix} \bar{0} & 1 \\ I_{\alpha-1} & \bar{1}^T \end{bmatrix}, \quad (5)$$

where  $\bar{1}$  and  $I_{\alpha-1}$  represent the all-1 vector and the identity matrix of dimension  $\alpha - 1$  respectively. The matrix  $D_\alpha(j)$  is defined as the dependency matrix associated with dimension  $j$  of the generator matrix  $G$ . This matrix allows to derive the following theorem.

**Theorem 2** *Let consider an  $(N, K)$  linear block code  $\mathcal{C}$  with a generator matrix generated randomly. Then the value  $\tilde{w}_i(j)$  corresponding to the dimension  $j$  with dependency matrix  $D_\alpha(j)$  is well approximated by*

$$\tilde{w}_i(j) \approx 2^{-(N-K)} \sum_{l=0}^{\lfloor \frac{\alpha-1}{2} \rfloor} \binom{\alpha}{2l+1} \binom{N-\alpha}{i-(2l+1)}. \quad (6)$$

Theorem 2 indicates that the larger  $\alpha$ , the larger the corresponding  $P_b(j)$ . Consequently,  $\alpha = 1$  gives the smallest bit error probability. For this case,  $D_1 = [1]$  which corresponds to a systematic encoding. Therefore, the optimum bit error probability for MLD at medium to high SNR is achieved by a systematic encoding if the inverse-mapping defined by  $G$  is used to retrieve the information bits. This strategy is intuitively correct since whenever a code sequence estimated by the decoder is in error, the best strategy to recover the information bits is simply to determine them independently. Otherwise, errors propagate. For  $\alpha = 1$ , (6) becomes [1]

$$\tilde{w}_i(j) \approx 2^{-(N-K)} \binom{N-1}{i-1} \approx (i/N) w_i. \quad (7)$$

In that case only, at high SNR, the bit error probability for MLD follows

$$P_b \approx \left( \frac{d_H}{N} \right) w_{d_H} \tilde{Q}(\sqrt{d_H}). \quad (8)$$

For Reed-Muller (RM) codes of length  $N \leq 64$ , we computed the ratios  $\tilde{w}_{d_H}(j)/w_{d_H}$  corresponding to (6) for various forms of generator matrices. In all cases, the value computed from (6) is the exact ratio, although the weight distribution of RM codes is far from a binomial distribution.

### III. APPLICATIONS

#### A. ML trellis decoding

ML trellis decoding is based on the trellis oriented generator matrix (TOGM) of the code considered [3]. If this matrix is used for encoding, trellis decoding becomes suboptimum with respect to the bit error probability of MLD. We present a simple method to overcome this problem.

Let  $G_t$  denote the TOGM of the code  $\mathcal{C}_t$ . Then, by row additions only, it is possible to obtain the generator matrix  $G$  of an equivalent code  $\mathcal{C}$  which contains the  $K$  columns of the identity matrix. This matrix is known as the reduced echelon form (REF). These operations modify the mapping between information bits and codewords, but since no column permutation has been realized, each codeword of  $\mathcal{C}$  is still uniquely represented by a path in the trellis of  $\mathcal{C}_t$ . Therefore ML trellis decoding of the received sequence is still possible if we use  $G$  for encoding. The trellis decoder estimates the code sequence which is closest to the received sequence. Then the information bits are easily retrieved due to the systematic nature of  $G$ . Since no restrictions on  $G_t$  apply, the matrix  $G$  can be obtained for any possible trellis decomposition.

In [10], a specific ML trellis decoding algorithm for the (63,57,3) Hamming code is proposed. The decoding is realized based on a generator matrix in cyclic form. It is also shown that an equivalent systematic representation outperforms the cyclic form by 0.4 dB at the BER  $10^{-5}$ . However, the decoding of the systematic code requires an additional step. By processing the generator matrix in cyclic form as described in this section, this additional step can be removed as the encoding matrix becomes  $G = [I_{57}P_6]$ . On the other hand, the cyclic structure no longer exists, but the encoder remains very simple.

Figures 1 and 2 depict the simulation results for the (32,16,8) and (32,26,4) RM-codes respectively. For both codes, we simulated ML decoding based on the REF and the conventional TOGM described in [3], and plotted the first term of the union bound derived from (4). As expected from the results of Section II, we observe a larger gap in error performance for the (32,26,4) RM-code. At the bit error rate (BER)  $10^{-6}$ , the gap in performance for this code is about 0.2 dB, which is of the same order as the difference between closest coset decoding (CCD) and ML trellis decoding [11]. Also, we observe a much significant gap at high BER of 0.4 dB for the (32,16,8) code and 1.1 dB for the (32,26,4) code. This behavior becomes important if a concatenated coding scheme is used.

The extension of this method to multi-stage trellis decoding does not follow in a straightforward way. In general, multi-stage decoding methods exploit the decomposable structure of the code considered, so that row additions on the associated generator matrix can

destroy this structure. For example, CCD of  $|u|u+v|$ -constructed codes exploits the repetition of the  $u$ -component code [11]. As a result, row additions in each component code generator matrix are allowed, but not from one matrix to the other. In addition, the propagation of decoding errors between decoding stages also has to be considered when searching for the optimum encoding matrix associated with multi-stage decoding.

### B. MLD in conjunction with algebraic decoding

Several soft decision decoding algorithms in conjunction with an algebraic decoder have been proposed [4]-[8]. In general, algebraic decoding is associated with a particular generator matrix form  $G_a$ . Therefore, if this form is used for encoding, the corresponding algorithm becomes suboptimum with respect to the bit error probability of MLD. Algebraic decoding algorithms can be divided into two classes, depending on whether the decoder delivers an estimate of the transmitted codeword of length  $N$  or of the information sequence of length  $K$ . In the first case, the method of Section A extends in a straightforward fashion. Hence decoding of cyclic codes can be realized this way. However, a similar method is also possible for the second class of algebraic decoders. Again, this method is transparent with respect to algebraic decoding, so that the conventional algebraic decoder corresponding to the code considered can still be used. This method simply consists of recording the row operations processed to obtain  $G$  in REF form  $G_a$  and applying the inverse operations to the information sequence delivered by the algebraic decoder.

Figure 3 depicts the improvement achieved by this method for Chase algorithm-2 with majority-logic-decoding for the (64,42,8) RM-code. The proposed method outperforms Chase algorithm-2 with conventional majority-logic-decoding by 0.15 dB at the BER  $10^{-5}$ .

### C. Concatenated coding

We consider the concatenated scheme presented in [12] where the inner code is a (64,40) subcode of the (64,42) RM code and the outer code is the NASA standard (255,223) RS code over  $GF(2^8)$ . The outer code is interleaved to a depth of 5. For this scheme, Figure 4 represents the simulated bit error performance for encoding with the TOGM and the REF. We observe that the systematic encoding outperforms the TOGM by about 0.2 dB at the BER  $10^{-5}$ . More importantly, we also notice that while the error performance curves corresponding to the inner codes differ by a constant value due to different error coefficients, the difference in bit error probability between the error performance curves corresponding to the concatenated system increases as the SNR increases.

## IV. CONCLUSION

In this paper, we have showed that for many good codes, the SGM provides the best bit error probability for MLD when the inverse mapping of the generator matrix  $G$  is used to retrieve the information sequence. Based on the presented results, we can conclude that a careful choice of the generator matrix becomes important when comparing different optimum, near-optimum or suboptimum soft decision decoding schemes. Generally, tenths of dB's separate the bit error performance of such schemes, so that a poor choice of the generator matrix of one of the scheme may result in an important relative degradation.

By exploiting the fact that modifying the mapping between information bits and codewords is transparent to the decoder, we modified conventional trellis decoding and MLD in conjunction with an algebraic decoder so that these schemes achieve the same bit error performance as for systematic encoding. Hence the decoding becomes independent of the encoding and can simply be viewed as a process providing the most likely codeword of the codebook. As a result, the decoder structure remains the same as the conventional one but in some cases the decoded sequence requires an additional simple reprocessing.

## REFERENCES

- [1] F. J. Mac Williams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, Amsterdam: North-Holland Mathematical Library, 1977.
- [2] A. B. Kiely, J. T. Coffey and M. R. Bell, "Optimal Information Bit Decoding of Linear Block Codes," *IEEE Transactions on Information Theory*, vol. IT-41, pp. 130-140, January 1995.
- [3] G. D. Forney Jr, "Coset Codes II: Binary Lattices and Related Codes," *IEEE Transactions on Information Theory*, vol. IT-34, pp. 1152-1187, September 1988.
- [4] G. D. Forney Jr, "Generalized Minimum Distance Decoding," *IEEE Transactions on Information Theory*, vol. IT-12, pp. 125-131, April 1966.
- [5] D. Chase, "A Class of Algorithms for Decoding Block Codes with Channel Measurement Information," *IEEE Transactions on Information Theory*, vol. IT-18, pp. 170-182, January 1972.
- [6] T. Kaneko, T. Nishijima, H. Inazumi and S. Hirasawa, "An Efficient Maximum Likelihood Decoding of Linear Block Codes with Algebraic Decoder," *IEEE Transactions on Information Theory*, vol. IT-40, pp. 320-327, March 1994.
- [7] H. Thirumoorthy, S. Lin and T. Kasami, "Soft Decision Decoding of Binary Linear Block Codes

Based on an Iterative Search Algorithm," *submitted to IEEE Transactions on Information Theory*, (revised November 1995).

- [8] M. P. C. Fossorier and S. Lin, "Complementary Reliability-Based Decodings of Binary Linear Block Codes," *submitted IEEE Transactions on Information Theory*, December 1995.
- [9] G. D. Forney Jr. "Concatenated Codes". Cambridge: M.I.T. Press, 1966.
- [10] A. M. Michelson and D. F. Freeman, "Viterbi Decoding of the (63,57) Hamming Codes - Implementation and Performance Results," *IEEE Transactions on Communications*, vol. COM-43, pp. 2653-2656, November 1995.
- [11] F. Hemmati, "Closest Coset Decoding of  $|u|u+v|$  Codes," *IEEE Journal on Selected Areas in Communications*, vol. JSAC-7, pp. 982-988, August 1989.
- [12] T. Kasami, T. Takata, K. Yamashita, T. Fujiwara and S. Lin. "On the Bit Error Probability of a Concatenated Coding Scheme," *The Proceedings International Symposium on Information Theory and its Applications*, Sydney, Australia, p. 1-6, November 1994, submitted *IEEE Transactions on Communications*, 1995.

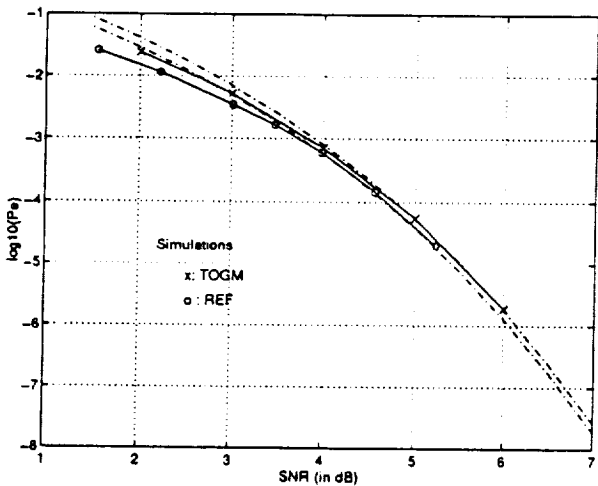


Figure 1: Simulated and theoretical bit error probabilities for the (32,16,8) RM code with TOGM and REF.

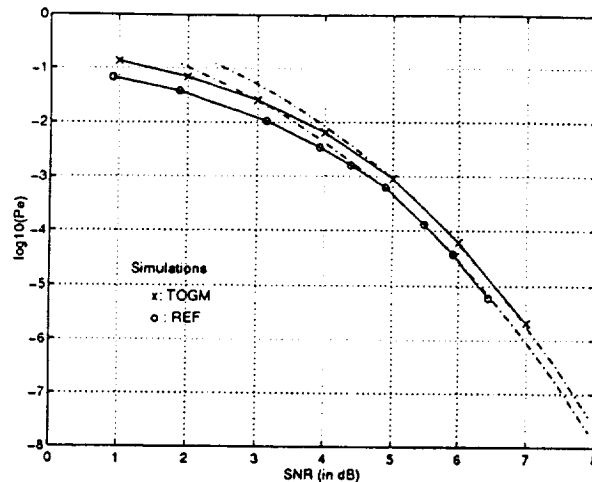


Figure 2: Simulated and theoretical bit error probabilities for the (32,26,4) RM code with TOGM and REF.

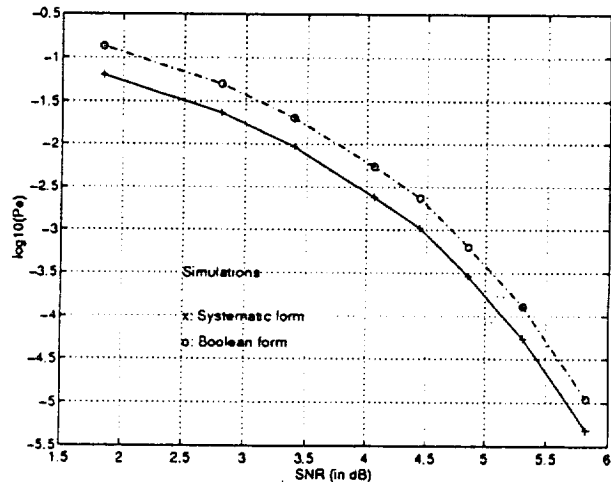


Figure 3: Simulated bit error probabilities for Chase algorithm-2 of the (64,42,8) RM code with majority-logic-decoding in Boolean and systematic forms.

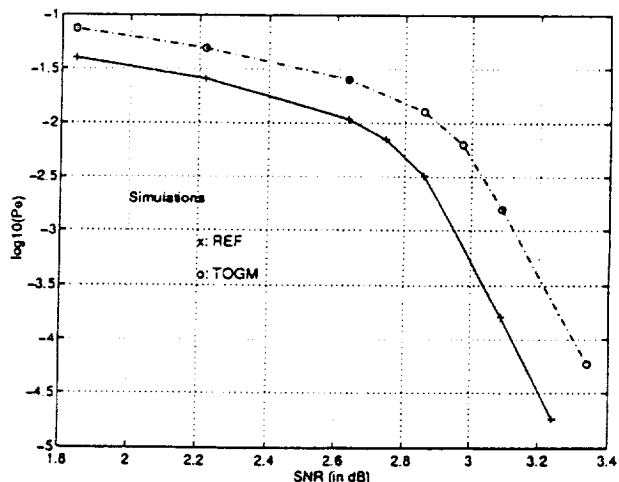


Figure 4: Simulated bit error probabilities for (255,223) RS outer code and (64,40,8) inner code, and encoding with REF and TOGM.

**PART IV**

**MULTILEVEL CONCATENATED  
BLOCK MODULATION CODES  
FOR THE FREQUENCY  
NON-SELECTIVE RAYLEIGH  
FADING CHANNEL**

**Dojun Rhee and Shu Lin**



# MULTILEVEL CONCATENATED BLOCK MODULATION CODES FOR THE FREQUENCY NON-SELECTIVE RAYLEIGH FADING CHANNEL

Dojun Rhee and Shu Lin<sup>1</sup>

LSI LOGIC CORPORATION  
MS G-815  
1525 McCarthy Blvd.  
Milpitas, California 95035  
E-mail : drhee@lsil.com

Dept. of Electrical Engineering  
2540 Dole Street, Holmes Hall #483  
University of Hawaii at Manoa  
Honolulu, Hawaii 96822  
slin@spectra.eng.hawaii.edu

*Abstract* — This paper is concerned with construction of multilevel concatenated block modulation codes using a multi-level concatenation scheme [1] for the frequency non-selective Rayleigh fading channel. In the construction of multilevel concatenated modulation code, block modulation codes are used as the inner codes. Various types of codes ( block or convolutional, binary or nonbinary ) are being considered as the outer codes. In particular, we focus on the special case for which Reed-Solomon (RS) codes are used as the outer codes. For this special case, a systematic algebraic technique for constructing  $q$ -level concatenated block modulation codes is proposed. Codes have been constructed for certain specific values of  $q$  and compared with the single-level concatenated block modulation codes using the same inner codes. A multilevel closest coset decoding scheme for these codes is proposed.

## I. INTRODUCTION

Single-level concatenated trellis coded modulation (TCM) for AWGN channel was first introduced by Deng and Costello in 1989 [2, 3]. Almost at the same time, Kasami et. al. presented a single-level concatenated block coded modulation (BCM) scheme for reliable data transmission over the AWGN channel [4]. Error performance of the single-level concatenated TCM and BCM schemes for the Rayleigh fading channel was first investigated by Vucetic and Lin in 1991 [5] and then by Vucetic in 1993 [6].

In the single-level concatenated BCM scheme [4], every information bits of inner code are protected by the same degree. In the AWGN channel, it is possible to design the inner code such that the bit error rate of each information bit is almost same. However, in the Rayleigh fading channel, it is not easy to design the inner code such that the bit error rate of each information bit is almost same. Therefore, in the single-level

concatenated BCM system, bit-error-rate of some information bits dominates the overall BER of the coded system and results in poor bit error performance. However, in a multilevel concatenated BCM scheme, it is possible to provide different degree of protection to information bits with different bit error rates. Outer codes in each level are designed to stop the error propagation to the next level of decoding.

Multilevel concatenated BCM codes constructed in this paper are designed to achieve better bit error performance than the single-level concatenated BCM code by stopping the error propagation to the next level decoding. Simulation results show that these codes achieve very impressive real coding gains over the uncoded reference system and single-level concatenated BCM codes using the same inner codes.

## II. MULTILEVEL CONCATENATED BLOCK CODED MODULATION SCHEME

The proposed multi-level concatenated coded modulation schemes are constructed using a multi-level concatenation approach [1].

In a  $q$ -level concatenated coded modulation system,  $q$  pairs of outer and inner codes are used as shown in Figure 1. In block coding, Reed-Solomon (RS) codes are used as the outer codes, and coset codes constructed from a block modulation code and its subcodes are used as the inner codes. The encoding and decoding are accomplished in  $q$  levels respectively.

### Outer Code Construction

For  $1 \leq i \leq q$  and  $1 \leq j \leq m_i$ , let  $B_{i,j}$  be an  $(N, K_i)$  RS (or shortened RS) code over  $GF(2^p)$  with minimum Hamming distance  $D_i = N - K_i + 1$ . In the  $i$ -th level outer code encoder, a  $(N, K_i, D_i)$  RS code is interleaved with depth  $m_i$ . For  $1 \leq j \leq m_i$ , let  $B_{i,j}$  represent the  $j$ -th code among  $m_i$  interleaved RS codes. After  $i$ -th level outer code encoding, symbols from  $GF(2^p)$  in each RS code are converted into  $p$ -bits binary representation. After conversion,  $m_i N p$ -bits are stored in the  $m_i$  by  $N p$  array such that every column has  $m_i$  bits and each bit in this column is selected from each RS code  $B_{i,j}$  for  $1 \leq j \leq m_i$ . Since  $B_{i,j} = B_{i,k}$

<sup>1</sup>This research was supported by NSF Grant NCR-911540, BCS-9020435 and NASA Grant NAG 5-931

for  $1 \leq j, k \leq m_i$ , let  $\mathbf{B}_i = \{B_{i,1}, B_{i,2}, \dots, B_{i,m_i}\}$  represents an  $i$ -th level outer code for  $1 \leq i \leq q$ . The  $i$ -th level outer code encoder is shown in Figure 2. Later, these  $q$ -sets of RS codes will be used as the outer codes in  $q$  levels of concatenation.

### Inner Coset Code Construction

Let  $\Lambda_0$  be a block modulation code over a certain elementary signal set  $S$  with length  $n$ , dimension  $k_0$  and minimum squared Euclidean distance  $\Delta_0$ . We require that

$$k_0 = m_1 + m_2 + \dots + m_q \quad (1)$$

From  $\Lambda_0$ , we form a sequence of subcodes,  $\Lambda_0, \Lambda_1, \Lambda_2, \dots, \Lambda_q$ , where  $\Lambda_q$  consists of the all-zero codeword, i.e.  $\Lambda_q = \{0\}$ . The dimension of these subcodes satisfy the following conditions: For  $1 \leq i \leq q$ ,  $\Lambda_i$  is a linear subcode of  $\Lambda_{i-1}$  with dimension

$$k_i = k_{i-1} - m_i \quad (2)$$

and minimum squared Euclidean distance  $\Delta_i$ . From 1 and 2, we have

$$\begin{aligned} k_1 &= m_2 + m_3 + \dots + m_{q-1} + m_q \\ k_2 &= m_3 + m_4 + \dots + m_q \\ &\vdots \\ k_{q-1} &= m_q \\ k_q &= 0 \end{aligned} \quad (3)$$

We also note that  $\Delta_0 \leq \Delta_1 \leq \dots \leq \Delta_q$  and that  $\Lambda_q$  consists of only the all-zero codeword with  $\Delta_q = \infty$ .

Now we are going to construct  $q$  coset codes from  $\Lambda_0, \Lambda_1, \dots, \Lambda_q$ . These  $q$  coset codes will be used as the inner codes in the proposed  $q$ -level concatenated coded modulation scheme. First we partition  $\Lambda_0$  into  $2^{m_1}$  cosets modulo  $\Lambda_1$ . Let  $\Lambda_0/\Lambda_1$  denote the set of cosets of  $\Lambda_0$  modulo  $\Lambda_1$ . The minimum squared Euclidean distance of each coset in  $\Lambda_0/\Lambda_1$  is  $\Delta_1$ . The minimum squared distance between two cosets in  $\Lambda_0/\Lambda_1$  is  $\Delta_0$ .  $\Lambda_0/\Lambda_1$  is called the **coset code of  $\Lambda_0$  modulo  $\Lambda_1$** . Next we partition each coset in  $\Lambda_0/\Lambda_1$  into  $2^{m_2}$  cosets modulo  $\Lambda_2$ . Let  $\Lambda_0/\Lambda_1/\Lambda_2$  denote the set of cosets of a coset in  $\Lambda_0/\Lambda_1$  modulo  $\Lambda_2$ . It is clear that the minimum squared Euclidean distance of a coset in  $\Lambda_0/\Lambda_1/\Lambda_2$  is  $\Delta_2$ , and the minimum squared distance among the cosets of a coset in  $\Lambda_0/\Lambda_1$  modulo  $\Lambda_2$  is  $\Delta_1$ . We call  $\Lambda_0/\Lambda_1/\Lambda_2$  the **coset code of  $\Lambda_0/\Lambda_1$  modulo  $\Lambda_2$** . We continue the above partition process to form coset codes. For  $1 \leq i \leq q$ , let  $\Lambda_0/\Lambda_1/\dots/\Lambda_{i-1}$  be the coset code of  $\Lambda_0/\Lambda_1/\dots/\Lambda_{i-2}$  modulo  $\Lambda_{i-1}$ . We partition each coset in  $\Lambda_0/\Lambda_1/\dots/\Lambda_{i-1}$  into  $2^{m_i}$  cosets modulo  $\Lambda_i$ . Then  $\Lambda_0/\Lambda_1/\dots/\Lambda_i$  is the coset code of  $\Lambda_0/\Lambda_1/\dots/\Lambda_{i-1}$  modulo  $\Lambda_i$ . The minimum squared Euclidean distance of a coset in  $\Lambda_0/\Lambda_1/\dots/\Lambda_i$  is  $\Delta_i$ , and the minimum squared distance among the cosets of a coset in  $\Lambda_0/\Lambda_1/\dots/\Lambda_{i-1}$  modulo  $\Lambda_i$  is

$\Delta_{i-1}$ . Note that each coset in  $\Lambda_0/\Lambda_1/\dots/\Lambda_{q-1}$  consists of  $2^{m_q}$  codewords in  $\Lambda_0$ . Since  $\Delta_q = \{0\}$ , each coset in  $\Lambda_0/\Lambda_1/\dots/\Lambda_q$  consists of only one codeword in  $\Lambda_0$ . Hence the minimum squared Euclidean distance of each coset is  $\Delta_q = \infty$ . The minimum squared distance among the cosets of a coset in  $\Lambda_0/\Lambda_1/\dots/\Lambda_{q-1}$  modulo  $\Lambda_q$  is  $\Delta_{q-1}$ . The above partition process results in a sequence of  $q$  coset codes,

$$A_1 = \Lambda_0/\Lambda_1 \quad (4)$$

$$A_2 = \Lambda_0/\Lambda_1/\Lambda_2$$

$$\vdots$$

$$A_q = \Lambda_0/\Lambda_1/\dots/\Lambda_q \quad (5)$$

These  $q$  coset codes are used as inner codes in the proposed  $q$ -level concatenated coded modulation scheme. This  $q$ -level concatenated modulation code  $\mathbf{C}$  is denoted as follows:

$$\mathbf{C} \triangleq \{\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_q\} * \{A_1, A_2, \dots, A_q\} \quad (6)$$

If  $\Lambda_0, \Lambda_1, \dots, \Lambda_{q-1}$  have simple trellis diagrams, the coset inner codes,  $A_1, A_2, \dots, A_q$ , also have simple trellis diagrams. If the coset inner codes have simple trellis structure, then we can use Viterbi decoding to decode coset inner codes. This will decrease decoding complexity of inner codes drastically.

### III. MULTILEVEL CLOSEST COSSET DECODING

A multilevel closest coset decoding for the proposed scheme is presented in this section. Each level decoding consists of the inner closest coset decoding and the outer code decoding. At  $i$ -th level decoding, the inner coset code  $A_i$  is decoded by using decoded estimates from first level decoder to  $i-1$ th level decoder. In the multilevel concatenated scheme, the  $i-1$ th level outer code is designed to reduce the error propagation from  $i-1$ th level decoder to  $i$ -th level decoder. Since decoded information at each level is passed to next level, decoding at each level depends on decoded information from the preceding level. Therefore, error propagation may occur. To reduce the probability of error propagation, outer codes must be selected by considering the specific channel characteristic. In following sections, multilevel concatenated codes are constructed by following rule. In the Rayleigh fading channel, strong outer codes must be used for levels where inner codes have small minimum symbol and product distances.

### IV. EXAMPLE

Consider a two-level concatenated coded 8-PSK modulation system for the frequency non-selective Rayleigh fading channel. In this system, two 3-level 8-PSK modulation codes of length 8 are used as the inner codes and two RS codes over the Galois field  $GF(2^9)$  are used as the outer codes as shown in Figure 3. The two



inner codes are constructed from two 3-level 8-PSK modulation codes,  $\Lambda_0 = \lambda[(8, 4, 4) * (8, 7, 2) * (8, 7, 2)]$  and  $\Lambda_1 = \lambda[(8, 1, 8) * (8, 4, 4) * (8, 4, 4)]$ .  $\Lambda_0$  has dimension 18, minimum symbol distance 2, minimum product distance 4, and minimum squared Euclidean distance 2.344. And  $\Lambda_1$  has dimension 9, minimum symbol distance 4, minimum product distance 16, and minimum squared Euclidean distance 4.688.  $\Lambda_1$  has a very simple trellis structure and can be decoded in either 3 stages, 2 stages, or one stage. Either way, the decoding complexity is very simple.

Since (8,4,4) code is a subcode of (8,7,2) and (8,1,8) is subcode of (8,4,4),  $\Lambda_1$  is a subspace of  $\Lambda_0$ . Now, partition  $\Lambda_0$  with respect to  $\Lambda_1$ . Let  $\Lambda_0/\Lambda_1$  denote this partition. Then  $\Lambda_0/\Lambda_1$  consists of  $2^9$  cosets, denoted  $\Omega_i$  with  $1 \leq i \leq 2^9$ , modulo  $\Lambda_1$ . Each coset in  $\Lambda_0/\Lambda_1$  has a trellis structure identical to that of  $\Lambda_1$ . Let  $[\Lambda_0/\Lambda_1]$  denote the set of coset representatives of the cosets in  $\Lambda_0/\Lambda_1$ .  $[\Lambda_0/\Lambda_1]$  is called a coset code. In the proposed two-level concatenated coded 8-PSK modulation system,  $[\Lambda_0/\Lambda_1]$  is used as the first-level inner code and  $\Lambda_1$  is used as the second-level inner code. Let  $C$  denote this two-level concatenated coded modulation system.

The two outer codes used in the proposed system are the (511,411,101) and (511,499,13) RS codes with symbols from the Galois field  $GF(2^9)$ . The (511,411,101) RS code is used as the first-level outer code and the (511,499,13) RS code is used as the second-level outer code. Each code symbol can be represented by a binary 9-bits.

The spectral efficiency of the above two-level concatenated coded 8-PSK system is 2.0034 bits/symbol. Its bit error performance over the frequency non-selective Rayleigh fading channel using two-level closest coset decoding is simulated and shown in Figure 4.

For comparison purpose, we construct a single-level concatenated modulation code with  $\Lambda_0$  as inner code. Since  $\Lambda_0$  has 18 information bits, (511,455,57) RS code is used as outer code with interleaving depth 2. Each symbols over  $GF(2^9)$  in outer code is converted into an 9-bits. Then form an 511 by 18 array where each column consists of 18 bits. Each column is then encoded into a codeword in  $\Lambda_0$ . Let  $C(S)$  denote the above single-level concatenated BCM code. The spectral efficiency of the code  $C(S)$  is 2.0034 bits/symbol. The error performance of the coherently detected 8-PSK single-level concatenated modulation code  $C(S)$  over the Rayleigh fading channel is also shown in Figure 4. The two-level concatenated modulation code  $C$  achieves a 1.4757 dB real coding gain over the single-level concatenated modulation code  $C(S)$  at a bit error rate (BER)  $10^{-5}$  with the same spectral efficiency. Also, two-level concatenated code achieves a 32.259 dB real coding gain at a BER  $10^{-5}$  over the uncoded QPSK modulation without bandwidth expansion.

Other two-, three- and six-level concatenated BCM schemes have been devised for the Rayleigh fading channel and they all achieve very impressive real coding gains over the uncoded reference system and single-level concatenated BCM schemes using same inner codes.

#### REFERENCES

- [1] V. A. Zinoviev, "Generalized Cascade Codes," *Problemy Peredach i Informatsii*, Vol. 12. 1, pp. 2-9, trans. from Russian by Plenum Publishing Co., 1976.
- [2] R. H. Deng and D. J. Costello, Jr., "High Rate Concatenated Coding Systems Using Bandwidth Efficient Trellis Inner Codes," *IEEE Transactions on Communications*, Vol. COM-37, pp. 420-427, May 1989.
- [3] R. H. Deng and D. J. Costello, Jr., "High Rate Concatenated Coding Systems Using Multi-dimensional Bandwidth Efficient Trellis Inner Codes," *IEEE Transactions on Communications*, Vol. COM-37, pp. 1091-1096, October 1989.
- [4] T. Kasami, T. Takata, T. Fujiwara, and S. Lin, "A Concatenated Coded Modulation Scheme for Error Control," *IEEE Trans. on Communications*, Vol. COM-38, No. 6, pp. 752-763, June 1990.
- [5] B. Vucetic, "Concatenated Multilevel Block Coded Modulation," *IEEE Trans. on Communications*, Vol. 41, No. 1, January 1993.
- [6] B. Vucetic and S. Lin "Block Coded Modulation and Concatenated Schemes for Error Control On Fading Channels," *Journal of Discrete Applied Mathematics*, No. 33, pp. 257-269, 1991.

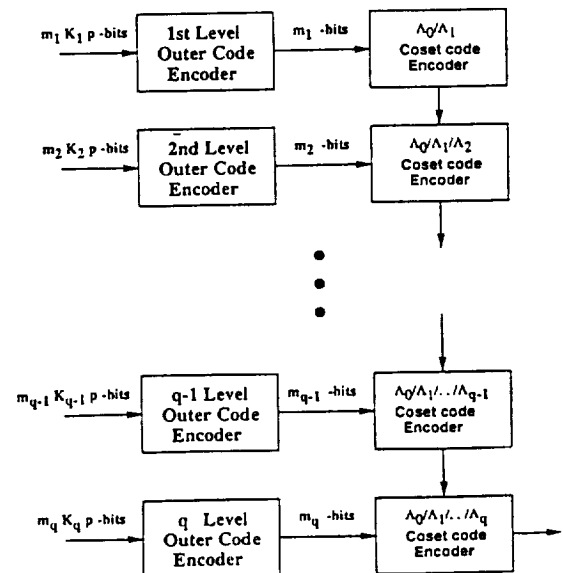


Figure 1 q-level concatenated block modulation code encoder

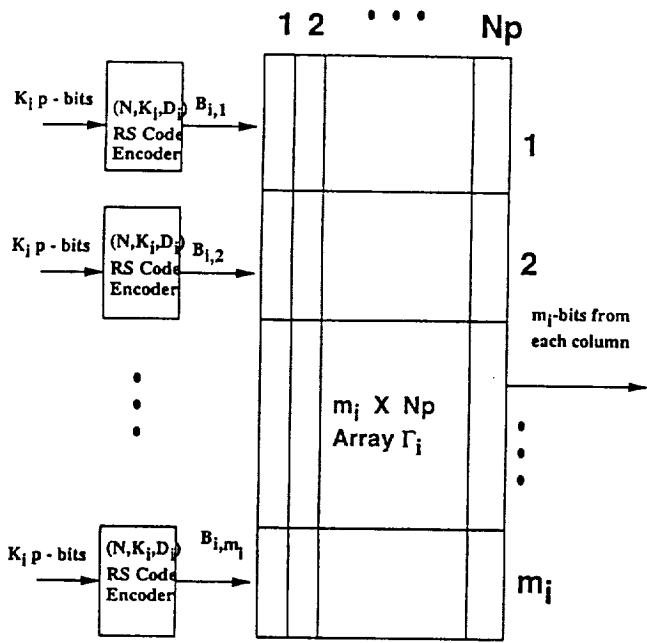


Figure 2 The  $i$ -th level outer code encoder

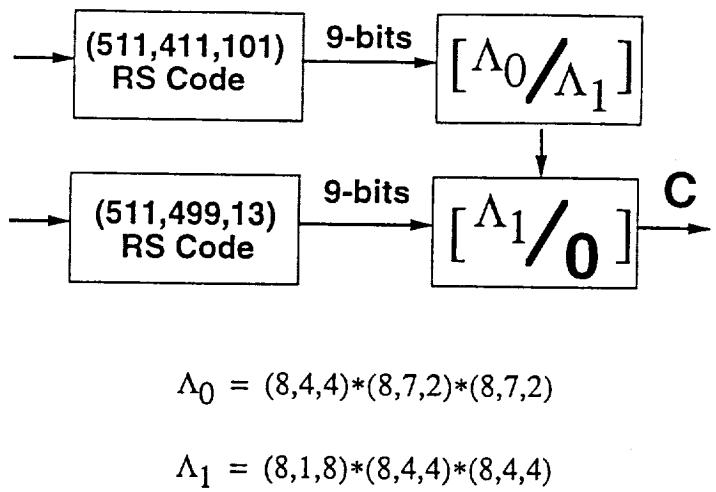


Figure 3 Two-level concatenated block modulation code

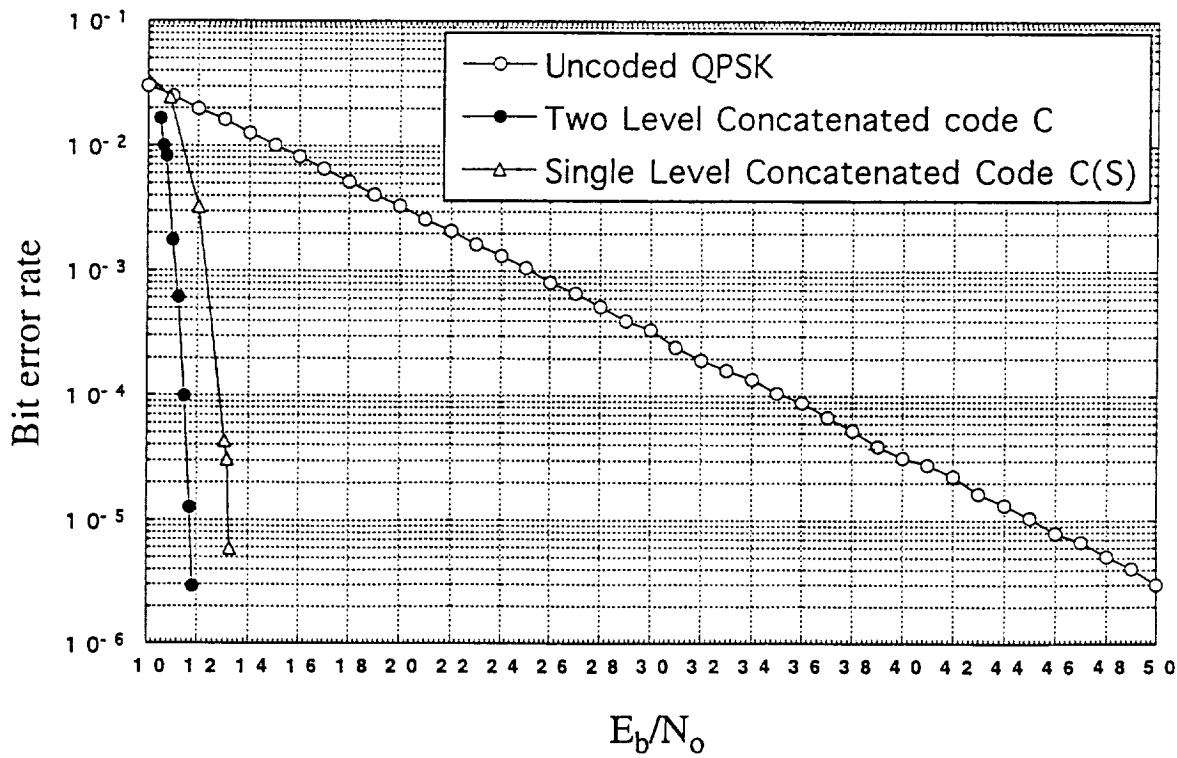


Figure 4. Bit error performance of 8-PSK two-level concatenated block modulation code over the frequency non-selective Rayleigh fading channel.

