RZ 3579 (# 99589) 02/07/2005
Electrical Engineering/Computer Science 20 pages

Research Report

Performance of Product Codes on Channels with Memory

S. Sankaranarayanan[†] and E. Eleftheriou*

[†]University of Arizona
Tucson, AZ 85721
USA

*IBM Research GmbH
Zurich Research Laboratory
8803 Rüschlikon
Switzerland

Performance of Product Codes on Channels with Memory

Sundararajan Sankaranarayanan, and Evangelos Eleftheriou*
University of Arizona, Tucson, AZ 85721
*IBM Zurich Research Laboratory, 8803 Rüschlikon, Switzerland

March 24, 2005

Abstract

A channel with memory has the characteristics of introducing correlated error bursts in transmitted data symbols. The inter-symbol interference and dropout effects in magnetic tape channels contribute to correlated error sequences in the read-back signal. A product-code-based error correcting scheme is used in conventional tape drive systems to combat these effects and to guarantee very high data fidelity. The read-back signal from the tape is decoded using a hard-decision distance-bounded decoder of the product code. In this paper, we present an analytical method to evaluate the performance of product codes on tape channels described as hidden Markov models. The performance measures of interest are symbol and codeword error rates of product codes. In addition to presenting the performance curves of the product code used in present day tape drive systems, we propose alternate error correcting schemes that can be considered for implementation in future generations.

1 Introduction

The exponential growth of digital information in recent years have immensely increased the demand for efficient and reliable mass storage systems. The magnetic tape drive systems is one of the mass storage devices that are mainly used for archiving and supporting backup and restore procedures. Also, it is the slow and inexpensive counterpart of the fast and expensive hard disk drives. It is not surprising that data reliability and robustness against errors due to the read/write operations or the media is very critical in these systems. One of the methods used to guarantee high reliability is the error-correction coding (ECC) schemes. The ECMA-319 Standard accepted by the industries describes the physical and magnetic characteristics of tape cartridges, quality of the recorded signals, data organization on the tapes, and reading/recording formats. The Standard specifies a Reed-Solomon (RS) codes based product code for error correction. In this paper, we propose a method to evaluate the current RS-code based product code and to investigate possible new alternatives for improved performance.

In Section 2, we present a brief introduction to product codes and hidden Markov model (HMM) for channels with memory. In Section 3, we develop a mathematical model of the error-correction coding subsystem of a magnetic tape drive system. This model is developed from the data organization format and error-correction scheme specified in the Linear Tape Open Standard. A simulation model of the ECC subsystem model is developed in Section 4 in order to evaluate the performance of a product-code-based error-correcting scheme. Although such a simulation model is a convenient way for performance evaluation of complex ECC subsystems, an analytical computation technique is preferrable for computing symbol and codeword error-rates below 10^{-6} in linear time. In Section 5, we describe the analytical computation technique for an ECC subsystem model defined by two-state HMMs. Indeed this technique can be applied for models defined by N-state HMMs. Using the simulation model and the analytical technique,

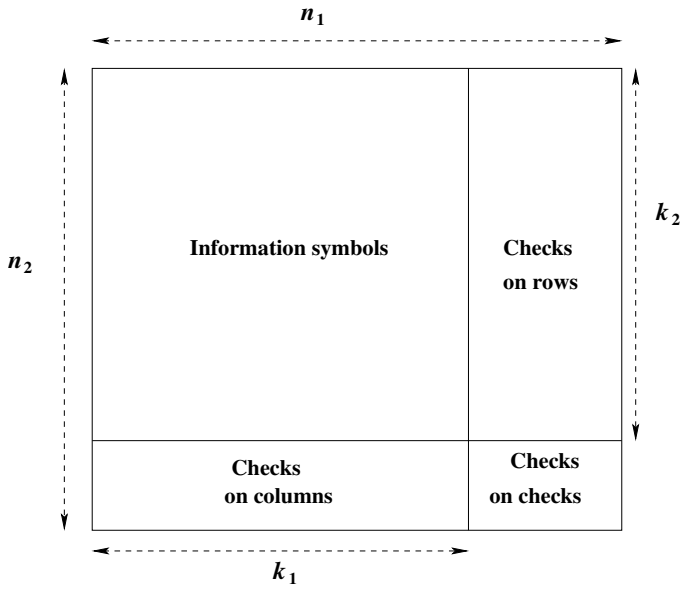


Figure 1: Structure of a product code $\mathcal{P}$.

we compute the performance of the error-correcting scheme specified in the Standard and other alternatives. These results are compiled in Section 6. In Section 7, we conclude this paper with a listing of the topics for future research work.

2 Preliminaries

2.1 Product Codes

A product code $\mathcal{P}$ is a two-dimensional code composed of two linear codes $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$. The (n_1, k_1, d_1) $\mathcal{C}^{(1)}$ is a k_1 dimensional linear code with n_1-tuple codewords that are at least d_1 Hamming-distance apart from each other. Similarly, $\mathcal{C}^{(2)}$ is defined as a (n_2, k_2, d_2) linear code. The correction capability of a maximum-likelihood decoder of $\mathcal{C}^{(1)}$ ($\mathcal{C}^{(2)}$) is $t_1 = \lfloor \frac{d_1-1}{2} \rfloor$ ($t_2 = \lfloor \frac{d_2-1}{2} \rfloor$).

The product code is a $(n_1 n_2, k_1 k_2, d_1 d_2)$ linear code whose codewords are represented as an $n_2 \times n_1$ array, shown in Fig. 1. Each $k_1 k_2$-tuple information symbol is arranged in a $k_2 \times k_1$ array. A row-column encoding scheme encodes each row of the $k_2 \times k_1$ array with $\mathcal{C}^{(1)}$ resulting in a $k_2 \times n_1$ array. Now, each column of this array is encoded using $\mathcal{C}^{(2)}$. The resultant $n_2 \times n_1$ array is a codeword in the product code, and for obvious reasons, $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$ are referred to as *row code* and *column code* respectively. A column-row encoding scheme–an alternative encoding scheme–encodes each column of the $k_2 \times k_1$ array with $\mathcal{C}^{(2)}$, and then encodes each row of the resultant $n_2 \times k_1$ array with $\mathcal{C}^{(1)}$. In A.1, we show that the row-column encoding scheme and the column-row encoding scheme result in the same codeword for a given information symbol array. Finally, it is easy to show that the minimum distance of the product code is $d_1 d_2$.

Although iterative soft-decision decoding algorithms for product codes, see [5], exist, in this paper we consider a sub-optimal hard-decision bounded-distance decoder. A *column-row decoder* decodes along the columns first, and

then decodes along the rows. If the number of errors in a column (row) codeword is less than or equal to t_2 (t_1), then it is decoded to the correct codeword, if not, the received word is left undisturbed. Unlike the encoding process, the performance of a column-row decoder need not be the same as that of a row-column decoder. The difference in their performance is attributed to the transmit and receive process of the communication system. For the application under consideration in this paper, the column-row decoder is chosen because they perform better than the row-column decoder.

2.2 Hidden Markov Model for Channel Errors

The error sequences generated by the magnetic channel of a tape storage system are correlated because of the inherent memory in the system. A hidden Markov model is a discrete-time finite state machine that is best suited for modeling correlated error sequences. The state of the HMM at discrete time t, where $t \in \mathbb{Z}^+$, is defined as the realization of the random variable $\mathbf{X}_t$. The state space of $\mathbf{X}_t$ (independent of t) is $\{s_1, s_2, \cdots, s_N\}$. The finite state machine satisfies the Markov property given in (1), and the time homogeneity given in (2).

$$\Pr(\mathbf{X}_{t+1} = x_{t+1} | \mathbf{X}_t = x_t, \mathbf{X}_{t-1} = x_{t-1}, \cdots, \mathbf{X}_0 = x_0) = \Pr(\mathbf{X}_{t+1} = x_{t+1} | \mathbf{X}_n = x_t) \quad (1)$$

$$\Pr(\mathbf{X}_{i+1} = x_{i+1} | \mathbf{X}_i = x_i) = \Pr(\mathbf{X}_{j+1} = x_{j+1} | \mathbf{X}_j = x_j) \quad (2)$$

Any state s_i of the HMM makes an error with a probability P_i. At any time t, the only observable quantity of the HMM is the presence or absence of error in the channel output. Since the state transitions are not observable, such a Markov model is referred to as the hidden Markov model. The HMM is completely defined by specifying a *state transition matrix* $\mathbf{P}^{(\mathrm{T})}$, and probability P_i,$1 \leq i \leq N$, of making errors in the state s_i. The state transition matrix $\mathbf{P}^{(\mathrm{T})}$ is a $N \times N$ whose (i,j)-th element, denoted as $\mathbf{P}^{(\mathrm{T})}_{i,j}$, is the probability of transition from s_i to s_j. The steady-state probability π_i of the HMM is the average probability of being in the state s_i, and it is computed by solving the system of equations in (3).

$$\begin{aligned} \langle \pi_1, \pi_2, \cdots, \pi_N \rangle &= \langle \pi_1, \pi_2, \cdots, \pi_N \rangle \, \mathbf{P}^{(\mathrm{T})} \\ \pi_1 + \pi_2 + \cdots + \pi_N &= 1 \end{aligned} \quad (3)$$

The average channel symbol error rate is $p_{\mathrm{s,ch}} = \sum_{i=1}^{N} \pi_i P_i$.

The correlated-error sequences of a HMM are governed by $\mathbf{P}^{(\mathrm{T})}$. An *error burst* of length l is defined as an event of l consecutive error preceded and followed by error-free symbols. Similarly, an *error gap* of length g is defined as an event of g consecutive error-free symbols preceded and followed by erroneous symbols. In this paper, we have considered a two-state HMM which is commonly referred to as the Gilbert-Elliot channel (GEC) model [3], [2]. The two states of the GEC model are referred to as the *Good state* and the *Bad state*. The probability of making an error in the Good state is P_{G} (usually zero), and the probability of making an error in the Bad state is P_{B} (usually non-zero). The state transition matrix is

$$\mathbf{P}^{(\mathrm{T})} = \begin{pmatrix} b & 1-b \\ 1-a & a \end{pmatrix}. \quad (4)$$

The steady-state probabilities of the GEC are $\pi_{\mathrm{G}} = \frac{1-a}{2-a-b}$ and $\pi_{\mathrm{B}} = \frac{1-b}{2-a-b}$, and the average channel symbol error rate is $p_{\mathrm{s,ch}} = \pi_{\mathrm{G}} P_{\mathrm{G}} + \pi_{\mathrm{B}} P_{\mathrm{B}}$. The probability distribution of error gaps and error bursts in a GEC for different values of g and l are shown in Fig. 2 and Fig. 3 respectively. Each curve correspond to different values of P_{B} while $P_{\mathrm{G}} = 0$, $a = 0.90$ and $p_{\mathrm{s,ch}} = 0.01$. From Fig. 3, the average burst lengths decrease with the decrease in a.

When $P_{\mathrm{G}} = 0$ and $P_{\mathrm{B}} > 0$, the matrix $\mathbf{P}^{(\mathrm{T})}$ is in fact the *error transition matrix* that governs the transition from an *error-free state*, in this case the Good state, to an *error state*, in this case the Bad state, and vice-versa. Now,

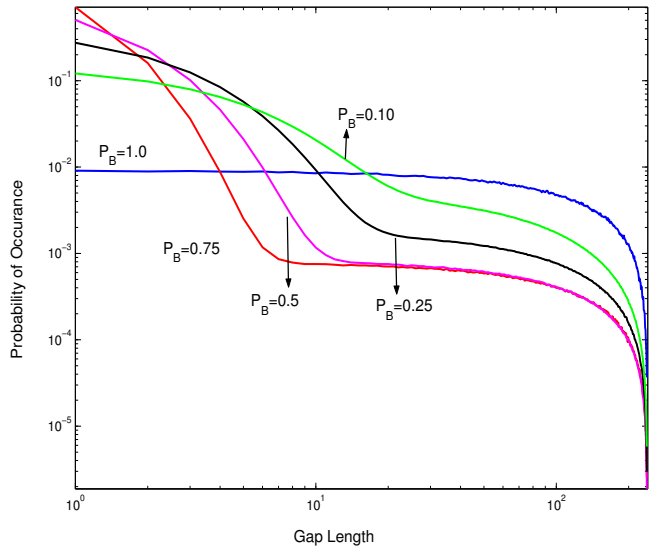


Figure 2: Gap length density of GEC channels with $P_{\mathrm{G}} = 0$, $a = 0.90$ and $p_{\mathrm{s,ch}} = 0.01$.

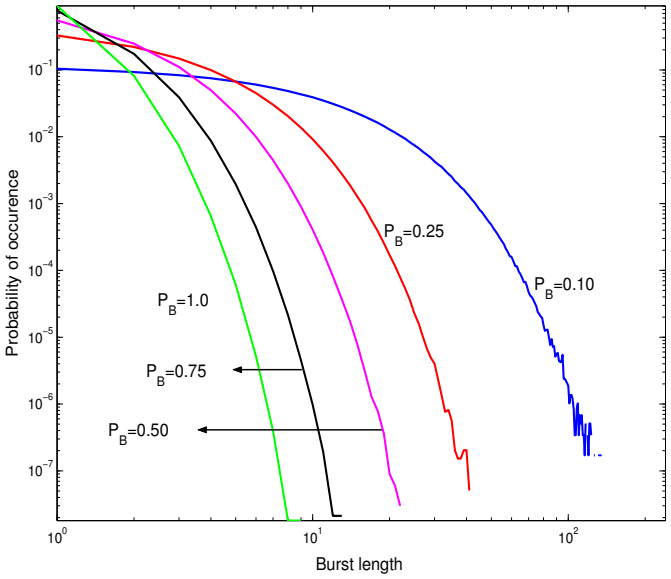


Figure 3: Burst length density of GEC channels with $P_{\mathrm{G}} = 0$, $a = 0.90$ and $p_{\mathrm{s,ch}} = 0.01$.

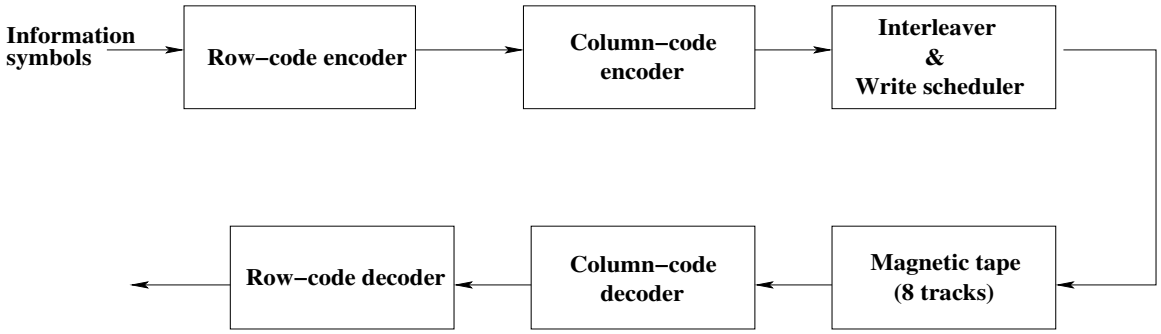


Figure 4: Illustration of the ECC subsystem of a tape drive system.

we define an error transition matrix $\mathbf{E}^{(\mathrm{T})}$ as a 2×2 matrix whose first row and column correspond to the error-free state (or 0-state), and the second row and column correspond to the error state (or 1-state). Such a matrix is a convenient representation of the HMM that will be used extensively in this paper. In A.2, we outline a method to calculate the $\mathbf{E}^{(\mathrm{T})}$ for a GEC model defined by $\mathbf{P}^{(\mathrm{T})}$, $P_{\mathrm{G}} > 0$, and $P_{\mathrm{B}} > 0$.

3 Modeling ECC Subsystem

An application where product codes are used for error correction is the tape drive system. The ECC subsystem of a tape drive system is illustrated in Fig. 4. The development of the *ECC subsystem model* for performance analysis is based on the specifications in the ECMA-319 Standard [1]. This ECMA Standard, accepted by the tape-drive manufacturing companies, specifies physical and magnetic characteristics of magnetic tape cartridges in order to facilitate physical interchange of cartridges among drives. This standard, amongst many other things, lays out methods to organize data on tapes, error correction schemes to protect this data, and recording schedules to write this data on tapes. For more details on the Standard, refer to [1].

As specified in the Standard, the information symbols (along with a few CRC check bytes) are grouped to form a *dataset* that contains 16 54×468-arrays appended to form a row vector. Each one of the 54×468-arrays are referred to as *sub-datasets*, and are encoded separately. A product code $\mathcal{P}^{(\mathrm{ECMA})}$ with two RS codes $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$ defined over $\mathrm{GF}(2^8)$ is used to protect each sub-datasets. The code $\mathcal{C}^{(1)}$ is a $(240, 234, 7)$ RS code, and the code $\mathcal{C}^{(2)}$ is a $(64, 54, 11)$ RS code. The encoded sub-dataset is a 64×480-array obtained by encoding even and odd symbols of each row of the sub-dataset with $\mathcal{C}^{(1)}$ to form a *codeword pair*, and the columns with $\mathcal{C}^{(2)}$. Hence, the encoded dataset is a 1×16 vector of 64×480 encoded sub-datasets, and two codeword pairs of the encoded dataset along with header information, referred to as *quads*, are numbered row-wise. The set of 8 quads corresponding to each row of the encoded dataset is written into 8 tracks of the magnetic tape. The recording schedule of the quads of a dataset is shown in Fig. 5. The elements encircled in the recording schedule represent the quads containing codewords from the first and second sub-datasets. The tracks of the magnetic tape are modeled as identical and independent HMMs. The recording schedule of the quads guarantees that each track of the magnetic tape contains exactly 8 row codewords of $\mathcal{P}^{(\mathrm{ECMA})}$, and that they are physically separated from each other by at least $7 \times 960 = 6720$ symbols (or 53760 bits for 8-bit symbols). Since the probability of an error burst covering more than 53760 bits is very small, it is fair to assume as writing these row codewords in 8 identical and independent channels. In other words, the ECC subsystem model consists of 64 identical and independent HMMs, i.e., one for each row codeword of $\mathcal{P}^{(\mathrm{ECMA})}$. Also, this means that an error burst in a row of $\mathcal{P}^{(\mathrm{ECMA})}$ does not spill over to the next row. Hence, the ECC subsystem

Track 0	**0**	**9**	**18**	**27**	**36**	**45**	.....	**484**	**493**	**502**	**511**
Track 1	**1**	**10**	**19**	**28**	**37**	**46**	.....	**485**	**494**	**503**	**504**
Track 2	**2**	**11**	**20**	**29**	**38**	**47**	.....	**486**	**495**	**496**	**505**
Track 3	**3**	**12**	**21**	**30**	**39**	**40**	.....	**487**	**488**	**497**	**506**
Track 4	**4**	**13**	**22**	**31**	**32**	**41**	.....	**480**	**489**	**498**	**507**
Track 5	**5**	**14**	**23**	**24**	**33**	**42**	.....	**481**	**490**	**499**	**508**
Track 6	**6**	**15**	**16**	**25**	**34**	**43**	.....	**482**	**491**	**500**	**509**
Track 7	**7**	**8**	**17**	**26**	**35**	**44**	.....	**483**	**492**	**501**	**510**

Figure 5: Recording schedule of the quads of a dataset.

model is sufficiently defined by the parameters of $\mathcal{C}^{(1)}$, $\mathcal{C}^{(2)}$, and a HMM. Although, with an aim to preserve clarity, we have restricted ourselves to GEC models, the analysis techniques presented in this paper can easily be extended to N-state HMMs.

Since the quads are written in a track of the magnetic tape using a binary modulation format, it is appropriate to use a *bit-based* definition for the GEC of a ECC subsystem model. In a *bit-based* GEC, each state is equivalent to a binary symmetric channel with the crossover probability equal to P_{G} or P_{B}. The transition probability matrix of the bit-based GEC model is denoted as $\mathbf{P}^{(\mathrm{T,bit})}$. An m-bit symbol is error-free if and only if m transitions in the bit-based GEC are error-free. Hence, the error transition matrix $\mathbf{E}^{(\mathrm{T})}$ of an equivalent *symbol based* GEC can be derived by applying the method shown in A.2 on $\mathbf{P}^{(\mathrm{T})}(D) := \left\{\mathbf{P}^{(\mathrm{T,bit})}(D)\right\}^m$, where

$$\mathbf{P}^{(\mathrm{T,bit})}(D) = \begin{pmatrix} b(1-P_{\mathrm{G}}) + bP_{\mathrm{G}}D & (1-b)(1-P_{\mathrm{B}}) + (1-b)P_{\mathrm{B}}D \\ (1-a)(1-P_{\mathrm{G}}) + (1-a)P_{\mathrm{G}}D & a(1-P_{\mathrm{B}}) + aP_{\mathrm{B}}D \end{pmatrix}. \tag{5}$$

Finally, a column-row decoding is employed to correct errors introduced by the read-back process. The ECC subsystem model is illustrated in Fig. 6.

4 Simulated Performance of Product Codes

The ECC subsystem model, defined by $\mathcal{C}^{(1)}$, $\mathcal{C}^{(2)}$, and a GEC, is implemented in C to evaluate the performance of the product code $\mathcal{P}$. The performance of $\mathcal{P}$ is analyzed by plotting the row-codeword error rate (and symbol error rate) observed at the output of the row-decoder for different values of $p_{\mathrm{s,ch}}$.

During each iteration, n_2 mn_1-tuple zero vectors are transmitted over identical n_2 independent channels defined by the GEC model. An element in the output-array of size $n_2 \times n_1$ is in error if there is at least an error in the

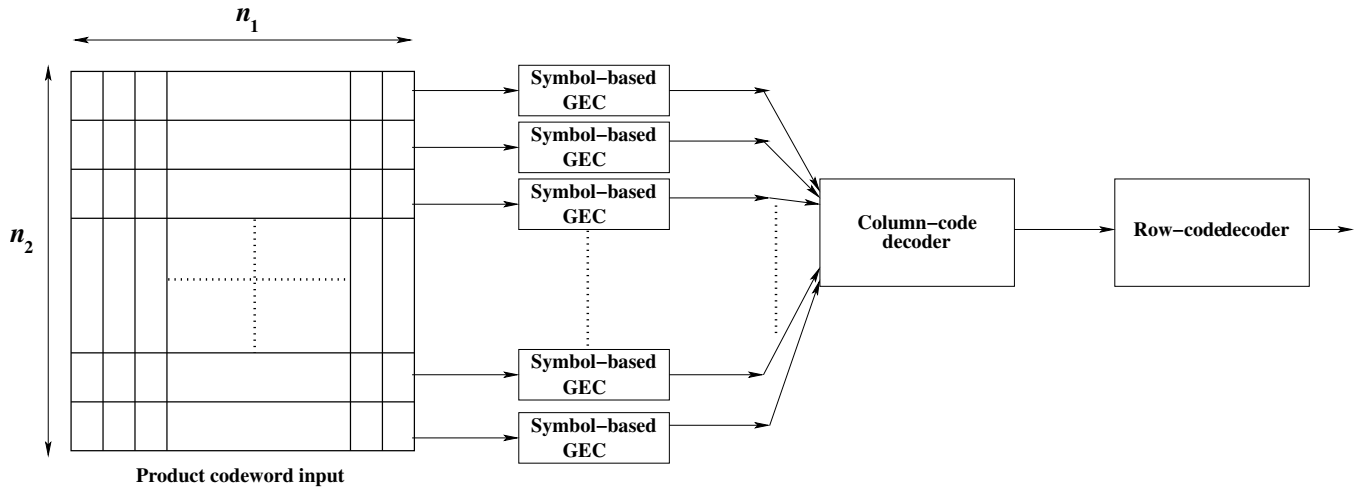


Figure 6: Illustration of the ECC subsystem model.

corresponding block of m channel-output bits. Each column of the output-array is decoded using $\mathcal{C}^{(2)}$, and then each row of the resultant array is decoded using $\mathcal{C}^{(1)}$. The number of symbol errors and row-codeword errors in the column-row decoded array of this iteration are memorized. Several iterations are performed to obtain sufficient error statistics to compute the probability of symbol error, denoted as $p_{\mathrm{s},\mathcal{C}^{(1)}}$, and row-codeword error, denoted as $p_{\mathrm{CW},\mathcal{C}^{(1)}}$. Although interleaving is not explicitly included in the simulator, we can introduce its effects by re-computing the parameters of the GEC models with respect to the given depth of intervealing.

The $p_{\mathrm{CW},\mathcal{C}^{(1)}}$ and $p_{\mathrm{s},\mathcal{C}^{(1)}}$ plots of $\mathcal{P}^{(\mathrm{ECMA})}$ shown in Fig. 7 correspond to GEC models with $P_{\mathrm{G}} = 0$, $P_{\mathrm{B}} = 1$, and different values for b. The average burst length of a GEC model increase with an increase in b.

5 Analytical Computation of the Performance of Product Codes

Assume that the ECC subsystem model is defined by a product code $\mathcal{P}$ defined over an m-dimension field, and a bit-based GEC with parameters P_{G}, P_{B}, and $\mathbf{P}^{(\mathrm{T,bit})}$. The error transition matrix $\mathbf{E}^{(\mathrm{T})}$ of the equivalent symbol-based GEC is computed. As shown in Fig. 6, there are n_2 identically defined GECs that operate independent of each other. The characteristics of these n_2 GECs are sufficiently defined by a matrix $\mathbf{E}^{(n_2,\mathrm{ch})}$, that is the n_2-fold Kronecker product of $\mathbf{E}^{(\mathrm{T})}$, see (6). The basic properties of Kronecker product are listed in A.3.

$$\begin{aligned} \mathbf{E}^{(n_2,\mathrm{ch})} &= \left(\mathbf{E}^{(\mathrm{T})}\right)^{n_2} \\ &:= \underbrace{\mathbf{E}^{(\mathrm{T})} \otimes \mathbf{E}^{(\mathrm{T})} \otimes \cdots \otimes \mathbf{E}^{(\mathrm{T})}}_{n_2 \mathrm{times}} \end{aligned} \tag{6}$$

From the convention introduced in 2.2, the columns (and rows) of $\mathbf{E}^{(n_2,\mathrm{ch})}$ are labelled using n_2-tuple binary vectors. For example, if $n_2 = 3$, then an element in the $\langle 0,1,0\rangle$-labelled row and $\langle 1,1,0\rangle$-labelled column represents the probability of transition from the error-free state to the error state in the first GEC, from the error state to the error state in the second GEC, and from the error-free state to the error-free state in the third GEC. Unfortunately, the

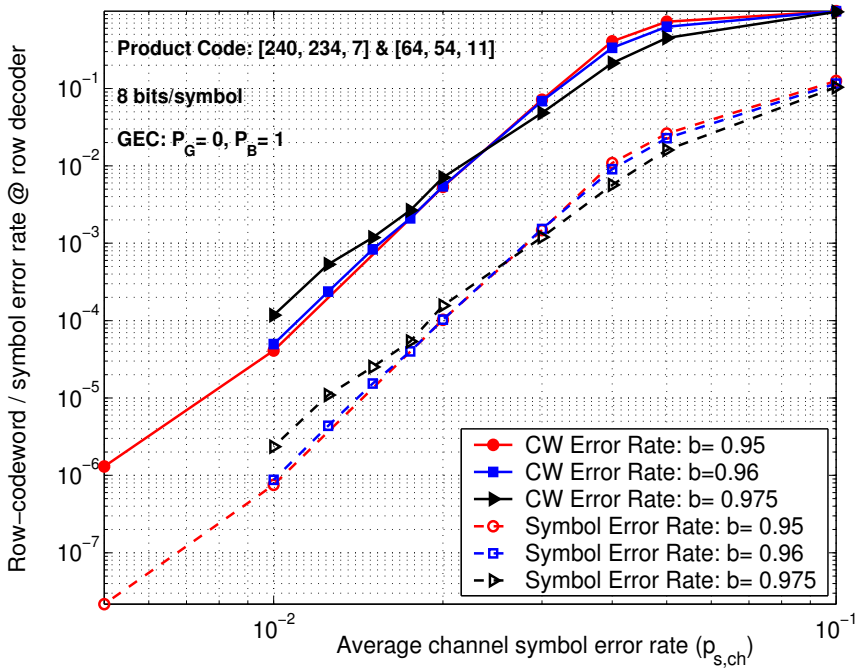


Figure 7: $p_{\mathrm{CW},\mathcal{C}^{(1)}}$ and $p_{\mathrm{s},\mathcal{C}^{(1)}}$ of $\mathcal{P}^{(\mathrm{ECMA})}$ over channels with different average burst lengths.

size of $\mathbf{E}^{(n_2,\mathrm{ch})}$ grows exponentially in n_2, and hence it is not feasible to compute this matrix for practical values of n_2.

A reduction in the computation complexity is achieved by realizing that not all information stored in $\mathbf{E}^{(n_2,\mathrm{ch})}$ is necessary for the computation of error-rates. As a first step, we have grouped the n_2 channels into two groups, namely Group-1 and Group-2. Group-1 contains the first channel, and Group-2 contains all the other channels. A symbol error in Group-1 is corrected by the column-code decoder if and only if there are at least $(t_2 - 1)$ errors in the corresponding column of Group-2. Hence, the Hamming weight of any realization vector of the channels in Group-2 is sufficient for our purposes. In other words, states with *like* Hamming weights in $\left(\mathbf{E}^{(\mathrm{T})}\right)^{n_2-1}$ can be merged to form a single state in $\tilde{\mathbf{E}}^{(n_2-1,\mathrm{ch})}$. Let $\mathcal{F}$ be a function that merges states with *like* Hamming weights such that

$$\tilde{\mathbf{E}}^{(n_2-1,\mathrm{ch})} = \mathcal{F}\left(\left(\mathbf{E}^{(\mathrm{T})}\right)^{n_2-1}\right) \tag{7}$$

The computation complexity of this matrix $\tilde{\mathbf{E}}^{(n_2-1,\mathrm{ch})}$, whose dimensions are $n_2 \times n_2$, is dramatically reduced by using the result, see A.4 for proof, given in (8).

$$\tilde{\mathbf{E}}^{(n_2-1,\mathrm{ch})} = \mathcal{F}\left(\mathbf{E}^{(\mathrm{T})} \otimes \tilde{\mathbf{E}}^{(n_2-2,\mathrm{ch})}\right),\ i > 1 \tag{8}$$

The recursive low-complexity computation technique is as follows:

$$\tilde{\mathbf{E}}^{(n_2,\mathrm{ch})} \quad := \quad \mathbf{E}^{(\mathrm{T})} \otimes \prod_{i=1}^{n_2-2} \mathcal{F}\left(\mathbf{E}^{(\mathrm{T})} \otimes \tilde{\mathbf{E}}^{(i,\mathrm{ch})}\right), \tag{9}$$

where $\tilde{\mathbf{E}}^{(1,\mathrm{ch})} := \mathbf{E}^{(\mathrm{T})}$.

The definition of the channel in Group-1 at the output of the column-code decoder is modified based on the correction capability of $\mathcal{C}^{(2)}$. The new definition $\tilde{\mathbf{E}}^{(n_2,\mathrm{col})}$ of the channel in Group-1 is developed by first observing the column and row arrangements of $\mathbf{E}^{(\mathrm{T})} \otimes \tilde{\mathbf{E}}^{(n_2-1,\mathrm{ch})}$ given in (10).

$$\tilde{\mathbf{E}}^{(n_2,\mathrm{ch})} = \begin{pmatrix} \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{1,1} & \cdots & \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{1,2^{n_2-1}+1} & \cdots & \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{1,2^{n_2-1}+\gamma} & \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{1,2^{n_2-1}+\gamma+1} & \cdots & \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{1,2^{n_2}} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{2^{n_2},1} & \cdots & \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{2^{n_2},2^{n_2-1}+1} & \cdots & \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{2^{n_2},2^{n_2-1}+\gamma} & \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{2^{n_2},2^{n_2-1}+\gamma+1} & \cdots & \tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}_{2^{n_2},2^{n_2}} \end{pmatrix}, \tag{10}$$

$$\text{where } \gamma := \sum_{i=0}^{t_2-1} \binom{n_2-1}{i}.$$

The first 2^{n_2-1} columns (and rows) of $\tilde{\mathbf{E}}^{(n_2,\mathrm{ch})}$ correspond to a 0-state in Group-1 channel. The next γ columns (and rows) correspond to a 1-state in Group-1 such that the Hamming weights of the Group-2 channels are less than t_2. The rest of the columns (and rows) correspond to a 1-state in Groupt-1 such that the Hamming weight of the Group-2 channels are at lest t_2. As mentioned above, a symbol in error at the output of the Group-1 channel is corrected by the decoder if and only if there are at least $(t_2 - 1)$ errors in the other tuples of the received word. Also, an error-free symbol at the output of this channel stays error-free at the output of the column-code decoder. Hence, the new definition of the channel in Group-1 is obtained by merging all states corresponding to the first $2^{n_2-1} + \gamma$ columns (and rows) to form an error-free state, and the rest of the states corresponding to the remaining columns (and rows) to form an error-state. Let $\mathbf{E}^{(\mathrm{T,col})}$ be the error-transition matrix of the modified GEC of the channel in Group-1.

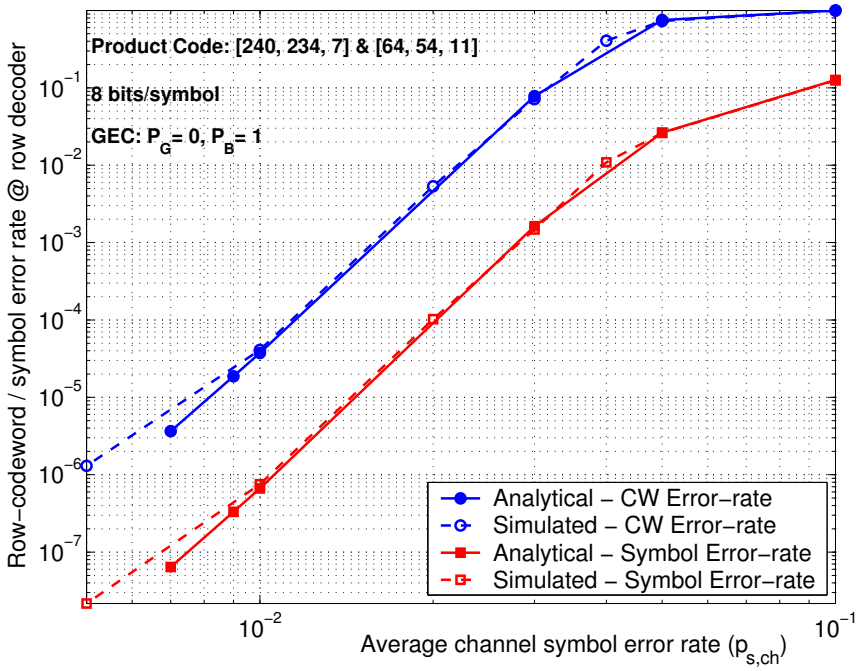


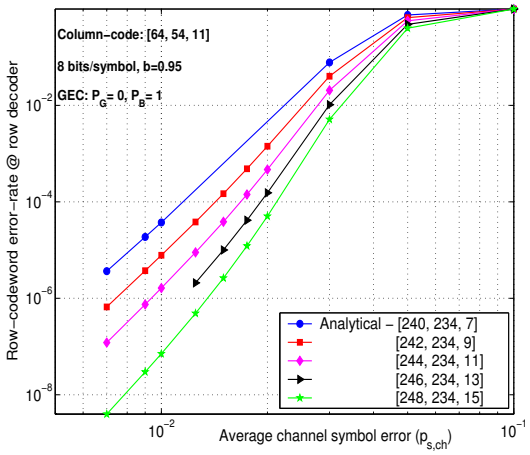
Figure 8: Comparison of $p_{\mathrm{CW},\mathcal{C}^{(1)}}$ of $\mathcal{P}^{(\mathrm{ECMA})}$ computed with analytical and simulation methods.

In [6], Wolf introduced an elegant method to compute the symbol and codeword error-rate performance of ECC over a channel modelled as a HMM. It is straightforward to apply this method on $\mathbf{E}^{(\mathrm{T,col})}$ in order to compute $p_{\mathrm{CW},\mathcal{C}^{(1)}}$ and $p_{\mathrm{s},\mathcal{C}^{(1)}}$ of $\mathcal{C}^{(1)}$. The performance of $\mathcal{P}^{(\mathrm{ECMA})}$, for $b = 0.95$ and other conditions as specified in Sect. 4, is computed using the analytical method. In Fig. 8, we can observe that the performance computed analytically agree with that computed by running simulations.

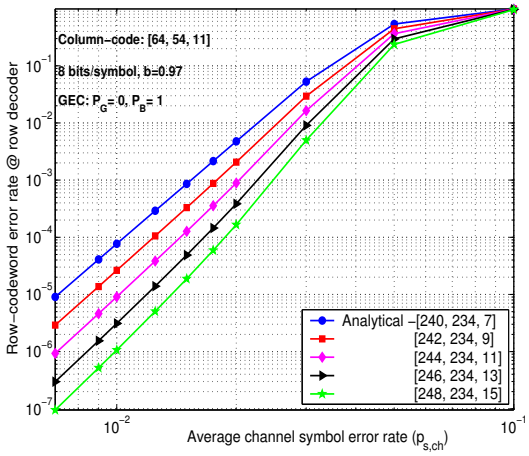
6 Alternate Error-Correcting Schemes

The analytical computation technique allows to investigate the performance of alternate RS-based product codes for tape drive systems. To design suitable product codes, it is essential to study the effectiveness of increasing the minimum distance of $\mathcal{C}^{(1)}$ to combat very long bursts. To this effect, we considered five different RS codes over GF(2^8) for $\mathcal{C}^{(1)}$, namely $(240, 234, 7)$, $(242, 234, 9)$, $(244, 234, 11)$, $(246, 234, 13)$, and $(248, 234, 15)$, but fixed $\mathcal{C}^{(2)}$ to the one defined in ECMA Standard. The plots in Fig. 9 are analytical result for transmission over a GEC with $P_{\mathrm{G}} = 0$, $P_{\mathrm{B}} = 1$, and values of $b \in \{0.95, 0.97, 0.98, 0.99\}$. The plots are not adjusted for code rates. These plots indicate that when the average burst length increases, the difference in performance between codes with the largest and smallest minimum distances decreases.

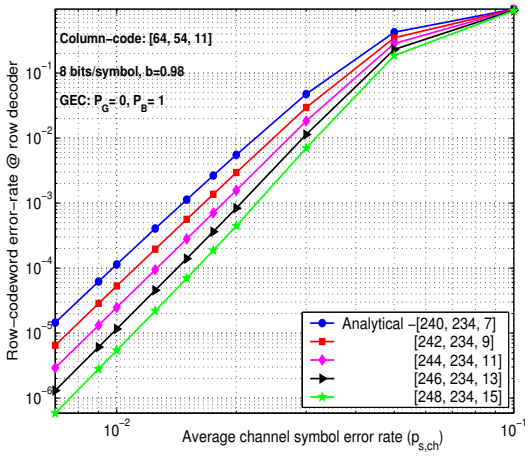
Another way to improve the performance of the product codes is to use 10-bits-per-symbol RS codes as component codes of the product code. The reason behind this approach is the fact that 10-bits-per-symbol RS codes are more resistant to bursts than their 8-bits-per-symbol counterparts. The first ECC scheme uses a $(386, 374, 13)$ RS code over GF(2^{10}) as $\mathcal{C}^{(1)}$, and a $(51, 43, 9)$ RS code over GF(2^{10}) as $\mathcal{C}^{(2)}$. The second ECC scheme $(480, 468, 13)$ RS code



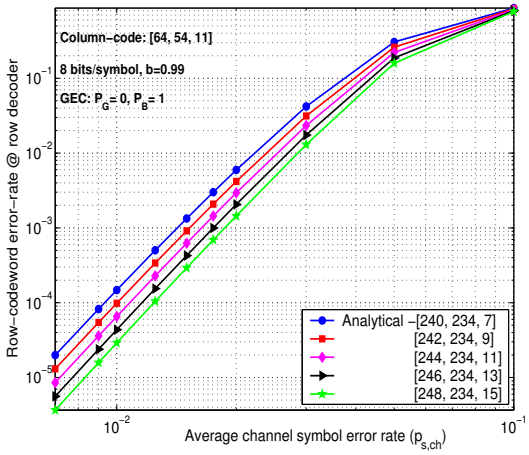
(a) $b = 0.95$



(b) $b = 0.97$



(c) $b = 0.98$



(d) $b = 0.99$

Figure 9: $p_{\mathrm{CW},\mathcal{C}^{(1)}}$ of product codes with different row codes for various burst conditions.

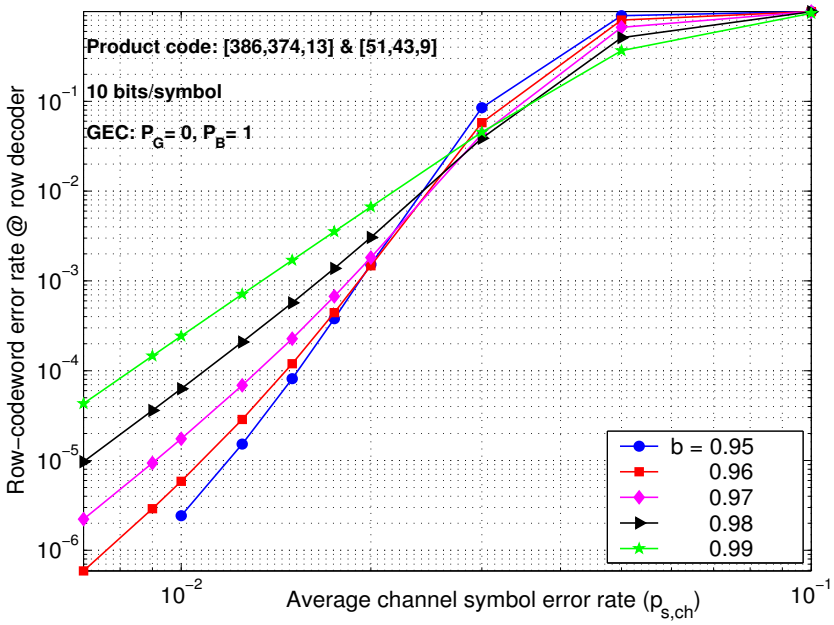


Figure 10: $p_{\mathrm{CW},\mathcal{C}^{(1)}}$ of the first ECC scheme for various burst conditions.

over GF(2^{10}) as $\mathcal{C}^{(1)}$, and a $(64, 54, 11)$ RS code over GF(2^{10}) as $\mathcal{C}^{(2)}$. The third ECC scheme $(480, 468, 13)$ RS code over GF(2^{10}) as $\mathcal{C}^{(1)}$, and a $(128, 108, 21)$ RS code over GF(2^{10}) as $\mathcal{C}^{(2)}$.

Note that the code rates in the second and third schemes are equal to those used specified in ECMA Standard. One of the major criteria in designing an ECC scheme for a tape drive system is that an erased track should be completely recoverable. This criterion is satisfied by the second and third schemes. In order to evaluate the performance of all three schemes, we assume an ECC subsystem model similar to the one described in Sect. 3. Also, the parameters of the GEC models are $P_{\mathrm{G}} = 0$, $P_{\mathrm{B}} = 1$, and values of $b \in \{0.95, 0.96, 0.97, 0.98, 0.99\}$. The performance plots of these schemes are shown in Fig. 10, Fig. 11, and Fig. 12. Once again, these plots are not adjusted for code rates. From the plots, it is obvious that these schemes perform better than the one specified in ECMA Standard. But, to make a definitive statement one has to investigate the processing overhead imposed by these new schemes, and also devise a procedure to adjust code rates where necessary to make fair comparisons.

7 Conclusion

In this paper, we developed a model of the ECC subsystem of a magnetic tape drive system in order to investigate its performance. Two computation techniques were proposed to compute the codeword and symbol error-rates of product codes over channels with memory. The first technique involved describing the ECC subsystem model in C, and running simulations to compute the error-rate performances. The time-complexity of this technique at high signal-to-noise ratio was prohibitive, and hence we developed an analytical computation technique. This technique was used to evaluate the performance of various 8-bits-per-symbol and 10-bits-per-symbol error-correcting schemes.

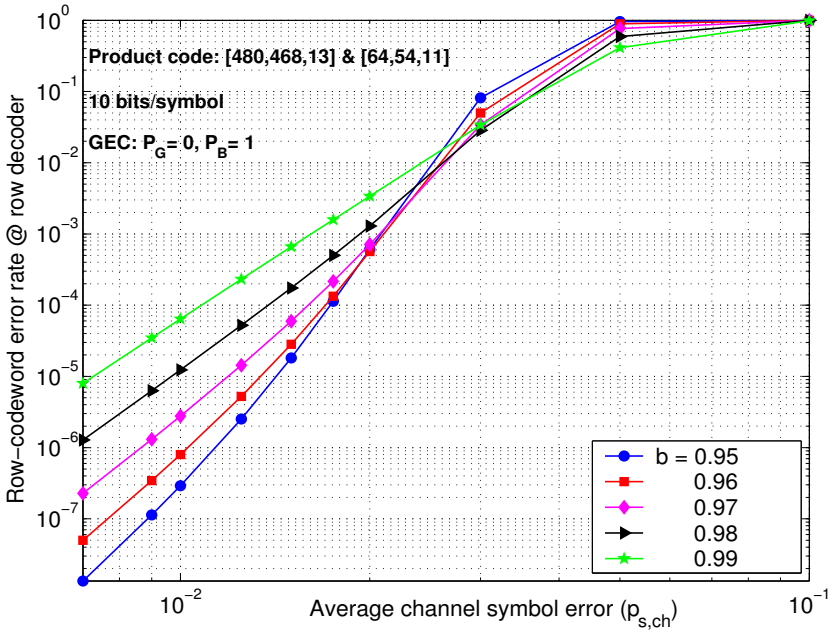


Figure 11: $p_{\mathrm{CW},\mathcal{C}^{(1)}}$ of the second ECC scheme for various burst conditions.

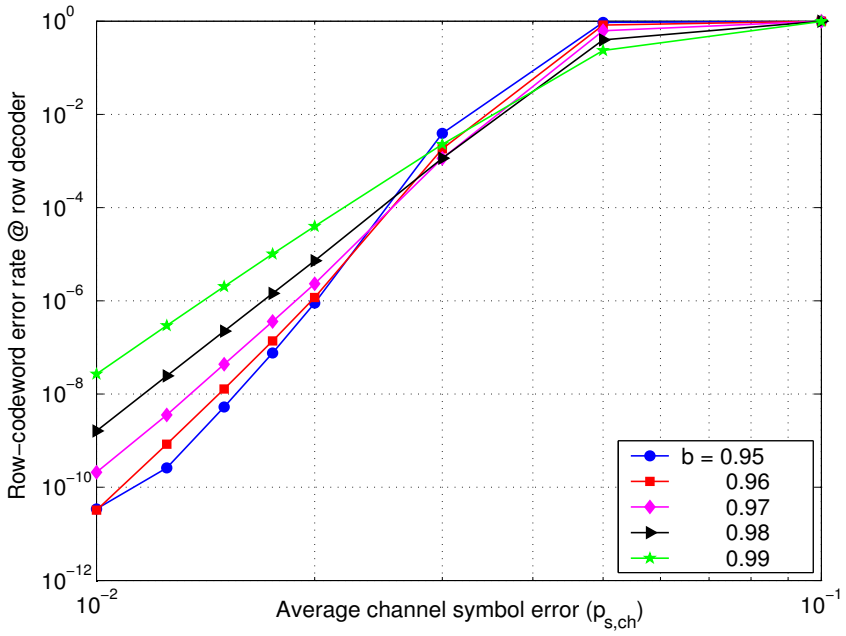


Figure 12: $p_{\mathrm{CW},\mathcal{C}^{(1)}}$ of the third ECC scheme for various burst conditions.

The next step in this line of research should be to investigate the performance of iterative hard-decision decoding schemes for product codes. Intuitively, such an approach should improve on the performance of the column-row decoder considered in this paper. Also, in this paper we have not considered the possibility of a decoder error which might result in a burst of errors. Finally, it is well known that the accuracy of the error-rates is critically dependent on the accuracy of the ECC subsystem model. Hence, it is necessary to run experiments to design a HMM that serves as a better approximation of the magnetic tape channel.

A Appendix

A.1 Order of Row and Column Encoding is Irrelevant

Using the row-column encoding scheme, the (p,q)-th element, where $k_2 < p \leq n_2$ and $k_1 < q \leq n_1$, of $\mathbf{c}$ for a given $k_2 \times k_1$ array of information symbols can be written as

$$\mathbf{c}_{pq} = \langle \mathbf{m}_1^{(1)}\mathbf{G}^{(1)}\delta_q^{(1)}, \mathbf{m}_1^{(2)}\mathbf{G}^{(1)}\delta_q^{(1)}, \cdots, \mathbf{m}_1^{(k_2)}\mathbf{G}^{(1)}\delta_q^{(1)} \rangle \;\; \mathbf{G}^{(2)}\delta_p^{(2)}, \tag{11}$$

where $\mathbf{m}_i^{(1)}$ is the i-th information symbol vector that is encoded with $\mathbf{G}^{(1)}$, $\delta_q^{(1)}$ is a vector of length n_1 such that the q-th element is 1 and every other element is 0, and $\delta_p^{(2)}$ is a vector of length n_2 such that the p-th element is 1 and every other element is 0.

Let $\mathbf{m}_j^{(2)}$ be the j-th information symbol vector that is encoded with $\mathbf{G}^{(2)}$, and $m_{j,r}^{(2)}$ be the r-th element of this vector. Using this convention, $\mathbf{m}_i^{(1)}$ can be written in terms of column vectors as shown in Eq. (12).

$$\mathbf{m}_i^{(1)} = \langle m_{1,i}^{(2)}, m_{2,i}^{(2)}, \cdots, m_{k_1,i}^{(2)} \rangle \tag{12}$$

By substituting this expression in Eq. (11), we obtain the expression shown in Eq. (13) where the (i,j)-th element of a matrix is denoted by $\mathbf{G}_{i,j}$.

$$\begin{aligned}
\mathbf{c}_{pq} &= \left[\left(m_{1,1}^{(2)}\mathbf{G}_{1,p}^{(2)} + m_{1,2}^{(2)}\mathbf{G}_{2,p}^{(2)} + \cdots + m_{1,k_2}^{(2)}\mathbf{G}_{k_2,p}^{(2)} \right) \mathbf{G}_{1,q}^{(1)} \right] + \\
&\qquad \left[\left(m_{2,1}^{(2)}\mathbf{G}_{1,p}^{(2)} + m_{2,2}^{(2)}\mathbf{G}_{2,p}^{(2)} + \cdots + m_{2,k_2}^{(2)}\mathbf{G}_{k_2,p}^{(2)} \right) \mathbf{G}_{2,q}^{(1)} \right] + \\
&\qquad \cdots + \left[\left(m_{k_1,1}^{(2)}\mathbf{G}_{1,p}^{(2)} + m_{k_1,2}^{(2)}\mathbf{G}_{2,p}^{(2)} + \cdots + m_{k_1,k_2}^{(2)}\mathbf{G}_{k_2,p}^{(2)} \right) \mathbf{G}_{k_1,q}^{(k_1)} \right] \\
&= \left[\left(\mathbf{m}_1^{(2)}\mathbf{G}^{(2)}\delta_p^{(2)} \right) \mathbf{G}_{1,q}^{(1)} \right] + \left[\left(\mathbf{m}_2^{(2)}\mathbf{G}^{(2)}\delta_p^{(2)} \right) \mathbf{G}_{2,q}^{(1)} \right] + \cdots + \left[\left(\mathbf{m}_{k_1}^{(2)}\mathbf{G}^{(2)}\delta_p^{(2)} \right) \mathbf{G}_{k_1,q}^{(1)} \right] \\
&= \langle \mathbf{m}_1^{(2)}\mathbf{G}^{(2)}\delta_p^{(2)}, \mathbf{m}_2^{(2)}\mathbf{G}^{(2)}\delta_p^{(2)}, \cdots, \mathbf{m}_{k_1}^{(2)}\mathbf{G}^{(2)}\delta_p^{(2)} \rangle \;\; \mathbf{G}^{(1)}\delta_q^{(1)}
\end{aligned} \tag{13}$$

The expression in Eq. (13) can be interpreted as encoding the column vectors with $\mathbf{G}^{(2)}$ and then encoding the resultant n_2 rows with $\mathbf{G}^{(1)}$. Hence, the column-row encoding scheme yields the same codeword as the row-column encoding process.

A.2 Computation of Error Transition Matrix

The GEC model with $\mathbf{P}^{(\mathrm{T})}$, $P_{\mathrm{G}} > 0$, and $P_{\mathrm{B}} > 0$ is illustrated in Fig. 13. A variable D is introduced in $\mathbf{P}^{(\mathrm{T})}$ to track the error occurences in each state.

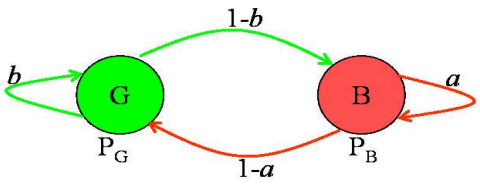


Figure 13: Illustration of the GEC model.

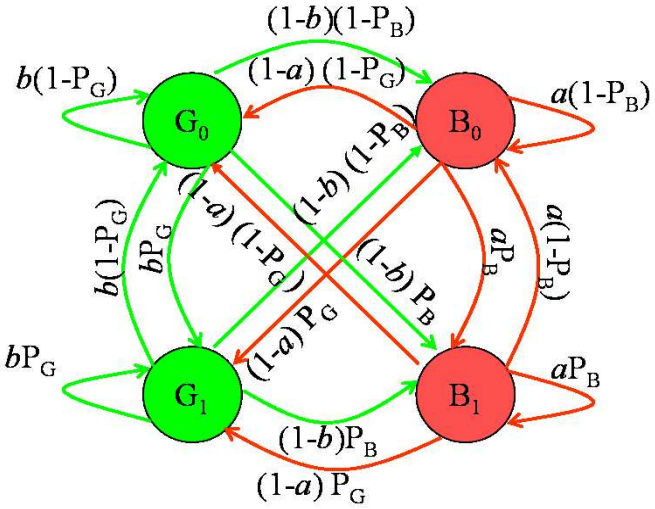


Figure 14: Illustration of the split GEC model.

$$\begin{aligned}
\mathbf{P}^{(\mathrm{T})}(D) &= \begin{pmatrix} b(1-P_{\mathrm{G}}) + bP_{\mathrm{G}}D & (1-b)(1-P_{\mathrm{B}}) + (1-b)P_{\mathrm{B}}D \\ (1-a)(1-P_{\mathrm{G}}) + (1-a)P_{\mathrm{G}}D & a(1-P_{\mathrm{B}}) + aP_{\mathrm{B}}D \end{pmatrix} \\
&= \begin{pmatrix} b(1-P_{\mathrm{G}}) & (1-b)(1-P_{\mathrm{B}}) \\ (1-a)(1-P_{\mathrm{G}}) & a(1-P_{\mathrm{B}}) \end{pmatrix} + D \begin{pmatrix} bP_{\mathrm{G}} & (1-b)P_{\mathrm{B}} \\ (1-a)P_{\mathrm{G}} & aP_{\mathrm{B}} \end{pmatrix} \qquad (14)
\end{aligned}$$

Since there is a non-zero probability of making errors in each state, we obtain a new GEC model by splitting each state into two new states with subscripts 0 and 1, see Fig. 14. A state with subscript 0 produces no errors, while a state with subscript 1 produces errors with unit probability. The state transition matrix $\tilde{\mathbf{P}}^{(\mathrm{T})}$ of the new GEC is written as:

$$\begin{aligned}
\tilde{\mathbf{P}}^{(\mathrm{T})} &= \begin{pmatrix} \mathbf{P}^{(\mathrm{T})}(0) & \mathbf{P}^{(\mathrm{T})}(1) - \mathbf{P}^{(\mathrm{T})}(0) \\ \mathbf{P}^{(\mathrm{T})}(0) & \mathbf{P}^{(\mathrm{T})}(1) - \mathbf{P}^{(\mathrm{T})}(0) \end{pmatrix} \\
&= \begin{pmatrix} b(1-P_{\mathrm{G}}) & (1-b)(1-P_{\mathrm{B}}) & bP_{\mathrm{G}} & (1-b)P_{\mathrm{B}} \\ (1-a)(1-P_{\mathrm{G}}) & a(1-P_{\mathrm{B}}) & (1-a)P_{\mathrm{G}} & aP_{\mathrm{B}} \\ b(1-P_{\mathrm{G}}) & (1-b)(1-P_{\mathrm{B}}) & bP_{\mathrm{G}} & (1-b)P_{\mathrm{B}} \\ (1-a)(1-P_{\mathrm{G}}) & a(1-P_{\mathrm{B}}) & (1-a)P_{\mathrm{G}} & aP_{\mathrm{B}} \end{pmatrix} \qquad (15)
\end{aligned}$$

The matrix $\tilde{\mathbf{P}}^{(\mathrm{T})}$ of the new GEC model has two states that produce no errors, and two other states that produce errors with unit probability. If π_{G_0}, π_{B_0}, π_{G_1}, and π_{B_1} are the steady state probabilities, then the required error transition matrix $\mathbf{E}^{(\mathrm{T})}$ is obtained by merging *like*-states as shown in Eq. (16).

$$\mathbf{E}^{(\mathrm{T})} = \begin{pmatrix} \frac{\pi_{\mathrm{G}_0}}{\pi_{\mathrm{G}_0}+\pi_{\mathrm{B}_0}} & \frac{\pi_{\mathrm{B}_0}}{\pi_{\mathrm{G}_0}+\pi_{\mathrm{B}_0}} & 0 & 0 \\ 0 & 0 & \frac{\pi_{\mathrm{G}_1}}{\pi_{\mathrm{G}_1}+\pi_{\mathrm{B}_1}} & \frac{\pi_{\mathrm{B}_1}}{\pi_{\mathrm{G}_1}+\pi_{\mathrm{B}_1}} \end{pmatrix} \tilde{\mathbf{P}}^{(\mathrm{T})} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \tag{16}$$

This method can be extended to compute error transition matrix for any N-state HMM.

A.3 Basic Properties of Kronecker Product

The *Kronecker product* of two matrices, $\mathbf{A}$ of dimensions $(\rho_{\mathbf{A}}, \gamma_{\mathbf{A}})$ and $\mathbf{B}$ of dimensions $(\rho_{\mathbf{B}}, \gamma_{\mathbf{B}})$, is a matrix $\mathbf{C}$ of dimensions $(\rho_{\mathbf{A}}\rho_{\mathbf{A}}, \gamma_{\mathbf{A}}\gamma_{\mathbf{B}})$ defined as $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$, where $\mathbf{C}_{i,j} = \mathbf{A}_{i,j}\mathbf{B}, \forall 1 \leq i \leq \rho_{\mathbf{A}}$ and $1 \leq j \leq \rho_{\mathbf{B}}$. The basic properties of Kronecker product are as follows:

1. Associativity: $\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C}$.
2. Distributivity over addition: $(\mathbf{A} + \mathbf{B}) \otimes (\mathbf{C} + \mathbf{D}) = (\mathbf{A} \otimes \mathbf{C}) + (\mathbf{B} \otimes \mathbf{C}) + (\mathbf{A} \otimes \mathbf{D}) + (\mathbf{B} \otimes \mathbf{D})$.
3. Compatibility with multiplication: $(\mathbf{A} \times \mathbf{B}) \otimes (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \otimes \mathbf{C}) \times (\mathbf{B} \otimes \mathbf{D})$.
4. Compatibility with inversion: $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$.
5. Commutativity up to permutation: $\mathbf{A} \otimes \mathbf{B} = \mathbf{\Phi}\,(\mathbf{B} \otimes \mathbf{A})\,\mathbf{\Psi}$, where $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are permutation matrices.

For proof of the properties and other details, refer to [4].

A.4 Proof of the Local State-Merging Technique

Theorem 1. *For any* $N > 1$, $\tilde{\mathbf{E}}^{(N,ch)} = \mathcal{F}\left(\mathbf{E}^{(T)} \otimes \tilde{\mathbf{E}}^{(N-1,ch)}\right)$.

Proof. The columns of $\mathbf{E}^{(N,\mathrm{ch})}$ are grouped based on the Hamming weights of the corresponding states, and in turn each group is divided into two subgroups based on the most significant bit (MSB) of the states. Such a grouping of columns is illustrated in Eq. (17), where HW(states) refers to the Hamming weight of states in that group. A similar arrangement is enforced on the rows of $\mathbf{E}^{(N,\mathrm{ch})}$.

$$\mathbf{E}^{(N,\mathrm{ch})} = \begin{pmatrix} & \overbrace{\qquad\qquad\qquad}^{\text{HW(states)=1 \& MSB is 0}} & & \overbrace{\qquad\qquad\qquad}^{\text{HW(states)=1 \& MSB is 1}} & & \\ \mathbf{E}^{(N,\mathrm{ch})}_{1,1} & \mathbf{E}^{(N,\mathrm{ch})}_{1,2} \cdots \mathbf{E}^{(N,\mathrm{ch})}_{1,\binom{N}{1}-1} & & \mathbf{E}^{(N,\mathrm{ch})}_{1,\binom{N}{1}} \cdots \mathbf{E}^{(N,\mathrm{ch})}_{1,\binom{N}{1}+1} & \cdots & \mathbf{E}^{(N,\mathrm{ch})}_{1,2^N} \\ \vdots & \ddots & & \vdots & \ddots & \vdots \\ \mathbf{E}^{(N,\mathrm{ch})}_{2^N,1} & \cdots & & \cdots & \cdots & \mathbf{E}^{(N,\mathrm{ch})}_{2^N,2^N} \end{pmatrix} \tag{17}$$

The states with *like* Hamming weights in $\mathbf{E}^{(N,\mathrm{ch})}$ are merged together to form a single state in order to obtain $\tilde{\mathbf{E}}^{(N,\mathrm{ch})}$. This operation can be written as:

$$
\begin{aligned}
\tilde{\mathbf{E}}^{(N,\mathrm{ch})} &= \mathcal{F}\left(\mathbf{E}^{(N,\mathrm{ch})}\right) \\
&= \begin{pmatrix}
\frac{1}{\binom{N}{0}} & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & \overbrace{\frac{1}{\binom{N}{1}} \quad \cdots \quad \frac{1}{\binom{N}{1}}}^{\binom{N}{1}} & & & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & \overbrace{\frac{1}{\binom{N}{2}} \quad \cdots \quad \frac{1}{\binom{N}{2}}}^{\binom{N}{2}} & & & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \frac{1}{\binom{N}{N}}
\end{pmatrix} \mathbf{E}^{(N,\mathrm{ch})} \\
&\quad \begin{pmatrix}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & \overbrace{1 \quad \cdots \quad 1}^{\binom{N}{1}} & & & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & \overbrace{1 \quad \cdots \quad 1}^{\binom{N}{2}} & & & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1
\end{pmatrix}^{\mathrm{T}} \\
&= \mathbf{\Pi}_{\mathrm{R}}\ \mathbf{E}^{(N,\mathrm{ch})}\ \mathbf{\Pi}_{\mathrm{C}},
\end{aligned}
\tag{18}
$$

where $(.)^{\mathrm{T}}$ is the transpose operator. For example, in Eq. (19), there are $\binom{N}{1}$ states with a unit Hamming weight, and the columns corresponding to these states in $\mathbf{E}^{(N,\mathrm{ch})}$ are summed up by the second column of $\mathbf{\Pi}_{\mathrm{C}}$.

Both $\mathbf{\Pi}_{\mathrm{R}}$ and $\mathbf{\Pi}_{\mathrm{C}}$ can be written as a product of two matrices as shown in Eqs. (19) and (20). The scale-factors in rows of $\mathbf{\Pi}_{\mathrm{R}}$ correspond to the steady-state probabilities of states with *like* Hamming weights. In addition, computing these scale factors are greatly simplified by the fact that all channels are identical and independent.

$$\mathbf{\Pi}_{\mathrm{R}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{\binom{N-1}{1}}{\binom{N}{1}} & \frac{\binom{N-1}{0}}{\binom{N}{1}} & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \frac{\binom{N-1}{2}}{\binom{N}{2}} & \frac{\binom{N-1}{1}}{\binom{N}{2}} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{\binom{N}{0}} & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & \overbrace{\frac{1}{\binom{N-1}{1}} \cdots \frac{1}{\binom{N-1}{1}}}^{\binom{N-1}{1}} & & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & \overbrace{\frac{1}{\binom{N-1}{0}} \cdots \frac{1}{\binom{N-1}{0}}}^{\binom{N-1}{0}} & & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \overbrace{\frac{1}{\binom{N-1}{2}} \cdots \frac{1}{\binom{N-1}{2}}}^{\binom{N-1}{2}} & & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \overbrace{\frac{1}{\binom{N-1}{1}} \cdots \frac{1}{\binom{N-1}{1}}}^{\binom{N-1}{1}} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \frac{1}{\binom{N}{N}} \end{pmatrix} \tag{19}$$

$$\mathbf{\Pi}_{\mathrm{C}} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & \overbrace{1 \cdots 1}^{\binom{N-1}{1}} & & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & \overbrace{1 \cdots 1}^{\binom{N-1}{0}} & & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \overbrace{1 \cdots 1}^{\binom{N-1}{2}} & & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \overbrace{1 \cdots 1}^{\binom{N-1}{1}} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}^{\mathrm{T}} \tag{20}$$

By substituting these expressions for $\mathbf{\Pi}_{\mathrm{R}}$ and $\mathbf{\Pi}_{\mathrm{C}}$ in Eq. (19), we prove the theorem statement as shown in Eq.

(21).

$$
\begin{aligned}
\tilde{\mathbf{E}}^{(N,\mathrm{ch})} &= \boldsymbol{\Pi}_{\mathrm{R}}\, \mathbf{E}^{(N,\mathrm{ch})}\, \boldsymbol{\Pi}_{\mathrm{C}} \\
&= \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & \frac{\binom{N-1}{1}}{\binom{N}{1}} & \frac{\binom{N-1}{0}}{\binom{N}{1}} & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \frac{\binom{N-1}{2}}{\binom{N}{2}} & \frac{\binom{N-1}{1}}{\binom{N}{2}} & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1
\end{pmatrix} \tilde{\mathbf{E}}^{(N-1,\mathrm{ch})} \\
&\quad \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1
\end{pmatrix}^{\mathrm{T}} \\
&= \mathcal{F}\left(\mathbf{E}^{(\mathrm{T})} \otimes \tilde{\mathbf{E}}^{(N-1,\mathrm{ch})}\right). \qquad (21)
\end{aligned}
$$

□

References

[1] "Data interchange on 12.7 mm 384-track magnetic tape cartridges," *ECMA*-319 Standard, 2001.

[2] E. O. Elliot, "Estimates of error-rates for codes on burst-noise channels," *Bell Syst. Tech J.*, vol. 42, pp.1977-1997, 1963.

[3] E. N. Gilbert, "Capacity of burst-noise channel," *Bell Syst. Tech J.*, vol. 39, pp.1253-1265, 1960.

[4] A. Graham, *Kronecker products and matrix calculus with applications*. New York: John Wiley and Sons, 1981.

[5] R. Pyndiah, "Near-optimum decoding of product codes: Block turbo codes," *IEEE Tran. Commun.*, vol. 46, pp. 1003-1010, Aug. 1998.

[6] J. K. Wolf, "ECC performance of interleaved RS codes with burst errors," *IEEE Trans. Magn.*, vol. 34, pp. 75-79, Jan. 1998.