

Non-interleaved Reed-Solomon Coding Performance on Finite State Channels

Cecilio Pimentel

Federal University of Pernambuco - UFPE
C. P. 7800 - CEP 50711-970 - Brazil
E.mail: cecilio@npd.ufpe.br

Ian F. Blake

Hewlett Packard Laboratories
1501 Page Mill Road - Palo Alto, CA 94304
E.mail: ifblake@occam.hpl.hp.com

Abstract – The analysis of a communication system operating over finite state channel (FSC) models includes the calculation of the probability of subsets of error sequences. In this paper we first present an analytical method for evaluating the performance of non-interleaved Reed-Solomon (RS) codes over channels modeled as FSC models with an arbitrary number of states. The main idea is to express the probability of the number of error symbols produced by the channel in terms of a coefficient in a formal power series. Next, the method is extended to study the effect on the performance when an interleaving with finite depth is incorporated into the communication system. The general expressions are specialized for a Gilbert-Elliott channel (GEC) with known model parameters, and numerical results are derived.

1. Introduction

One important family of discrete mathematical models that has been used to characterize the error sequence in channels with memory is the family of finite state channel (FSC) models. The achievement of reliable communication over these models will strongly depend on the design of efficient coding schemes. RS codes are non-binary codes of considerable importance in a wide variety of information transmission systems. Due to this symbol orientation, RS codes are suited to an environment where both burst and random errors occur.

The method we propose for evaluating the performance of non-interleaved RS codes on FSC models consists of finding a formal power series, the generating series, for certain sets of error sequences at the input to the decoder. We will show that these generating series can be translated directly into the codeword error probability using a linear mapping. This approach allows us to obtain a recurrence formula for the codeword error probability, which is convenient for computation.

In some applications, a useful methodology to enhance the overall burst capability is to design a RS code with some burst error ability and incorporate an interleaving into the communication system. An important parameter of an interleaving is the interleaving depth, which is defined as one less than the length of the shortest burst which can hit any codeword twice [1]. However, some practical restrictions may limit the maximum value of the interleaving depth (we refer

to [2] for an example of systems with stringent delay constraints). The interesting problem of analyzing theoretically the performance of RS codes using a particular non-ideal interleaving over FSC models will also be considered in this paper. Therefore, we compare different coding schemes under the same memory and delay constraint.

Previous analyses of error correcting codes over FSC models have been largely based on the evaluation of the codeword error probability for binary block codes through the determination of $P(m, n)$, the probability of m channel bit errors in a block of length n bits [2, 3, 4]. In [5] a recursive method was derived for the determination of the error probability of non-binary codes (e.g. Reed-Solomon codes) for the Fritchman model with one error state (renewal model). The performance of Reed-Solomon (RS) codes used with a channel modeled as a Markov chain is analyzed in references [6, 7]. Results for RS codes over FSC models are obtained from computer simulation in reference [2].

Much analytical work on RS codes for FSC models has been limited to the case of renewal models, and these results do not seem to be easily extended to more general models. Therefore, the contribution of this paper lies in presenting a more complete set of results for the performance of RS codes over general FSC models. As an application, the results are specialized for a GEC with known model parameters. The choice of coding parameters and interleaving depth to achieve a required performance is discussed for the GEC with different values of memory.

2. Communication System Description

An (n, k) primitive Reed-Solomon (RS) code defined in the Galois field $GF(2^c)$ has codewords of length $n = 2^c - 1$ symbols (where c is a positive integer), k information symbols, $n - k$ parity check symbols, and code rate $R_c = k/n$. An RS code can correct any combination of up to t error symbols within a codeword, where $t = \lfloor (n - k)/2 \rfloor$ is denoted the error correcting capability of the code, and $\lfloor x \rfloor$ is the greatest integer less than or equal to x .

We are considering a coded communication system where non-binary transmitted symbols, assuming values from $GF(2^c)$, are transmitted across a binary channel. Each symbol in a transmitted codeword is corrupted by an additive

error symbol e_k , composed of a sequence of c error bits statistically distributed according to an FSC model. Each error symbol e_k can also be regarded as an element from $\text{GF}(2^c)$, where each sequence of c bits is the vector-space representation of the corresponding field element. The k^{th} received symbol within a codeword is the sum $z_k = c_k + e_k$, $z_k \in \text{GF}(2^c)$, where the addition is over $\text{GF}(2^c)$. The transmitted symbol is received correctly (i.e., $z_k = c_k$) if the symbol e_k is the sequence of c consecutive zeros, denoted as $\mathbf{0}_c$. Otherwise, if $e_k \neq \mathbf{0}_c$ the transmitted symbol is received incorrectly.

A stationary N -state binary FSC model is defined in terms of the $N \times N$ matrices $\mathbf{P}(0)$ and $\mathbf{P}(1)$, such that the probability of a binary error sequence of length n bits generated by the channel, say $\mathbf{b}_n \triangleq (b_1 b_2 \dots b_n)$ is given by:

$$P(\mathbf{b}_n) = \Pi^T \left(\prod_{k=1}^n \mathbf{P}(b_k) \right) \mathbf{1},$$

where $\mathbf{1}$ is a column vector with all entries ones, Π is the $N \times 1$ -column matrix of stationary probabilities of a Markov chain, and the superscript $[\cdot]^T$ indicates the transpose of a matrix. For example, the GEC model is described by a two state Markov chain, as illustrated in Figure 1. The error process is generated according to the following probabilistic mechanism. When the chain is in the state 0 (the good state) the channel corrupts the transmitted bit with probability g . Otherwise, when it is in the state 1 (the burst state), the channel produces an erroneous bit with higher probability b . The GEC model is specified by the following matrices:

$$\mathbf{P}(0) = \begin{bmatrix} (1-Q)(1-g) & Q(1-b) \\ q(1-g) & (1-q)(1-b) \end{bmatrix};$$

$$\mathbf{P}(1) = \begin{bmatrix} (1-Q)g & Qb \\ qg & (1-q)b \end{bmatrix};$$

$$\Pi = [\pi_0 \pi_1]^T = \left[\frac{q}{q+Q} \quad \frac{Q}{q+Q} \right]^T.$$

The transition probability matrix of the Markov chain is given by $\mathbf{P} = \mathbf{P}(0) + \mathbf{P}(1)$.

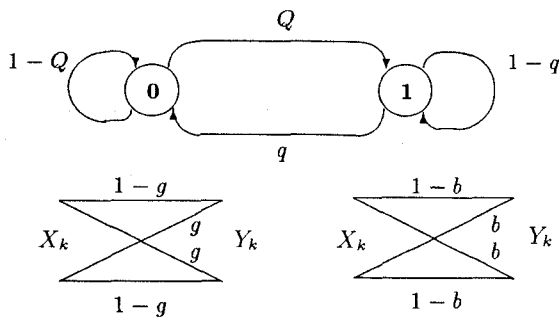


Fig. 1: Gilbert-Elliott model for burst channels.

We now define some notation. Let R be the field of real numbers. If s and z are commutative indeterminates, $[s^k z^n]P(s, z)$ denotes the coefficient of $s^k z^n$ in the formal power series $P(s, z)$. The identity matrix is denoted by \mathbf{I} .

3. Performance Analysis

An expression for the probability of m erroneous received symbols in a block of length n , denoted by $P_s(m, n)$, is developed next. Let F_c and F_e denote the generating series for sets of error symbols e_k that produce a correct and an erroneous received symbol, respectively. For RS codes, F_c enumerates a set with only one sequence, $\mathbf{0}_c$, and F_e enumerates the rest of all binary sequences of length c . Because the probability of each sequence produced by the channel depends on the non-commutative product of matrices $\mathbf{P}(0)$ and $\mathbf{P}(1)$, we define the generating series in non-commuting indeterminates. Let the indeterminates x_0 and x_1 mark an error bit (produced by the channel) equal to 0 or 1, respectively. Then

$$\begin{aligned} F_c &= x_0^c && \in R \ll x_0, x_1 \gg; \\ F_e &= (x_0 + x_1)^c - x_0^c && \in R \ll x_0, x_1 \gg, \end{aligned} \quad (1)$$

where $R \ll x_0, x_1 \gg$ is the ring of all power series in the non-commuting indeterminates x_0 and x_1 with coefficients taken from R . The set of all error symbol patterns of any length may be expressed as $(1 - F_c - F_e)^{-1}$. Notice that $P_s(m, n)$ is equal to the probability that m error symbols from the set enumerated by F_e occur in a block of n consecutive error symbols. Let the indeterminate z mark the length of an error word (an n -tuple of error symbols over $\text{GF}(2^c)$), and let s mark the number of error symbols from the set enumerated by F_e in an error word. Then, by defining a mapping Δ that replaces x_{b_i} by $\mathbf{P}(b_i)$ [8], $P_s(m, n)$ may be expressed as:

$$\begin{aligned} P_s(m, n) &= [s^m z^n] \Pi^T \Delta(1 - z(F_c + sF_e))^{-1} \mathbf{1}; \\ &= [s^m z^n] \Pi^T \Delta(1 - z\{x_0^c + s((x_0 + x_1)^c - x_0^c)\})^{-1} \mathbf{1}; \\ &= [s^m z^n] \Pi^T (\mathbf{I} - z\{\mathbf{P}(0)^c + s(\mathbf{P}^c - \mathbf{P}(0)^c)\})^{-1} \mathbf{1}. \end{aligned} \quad (2)$$

From Equation (3) it is simple to derive recurrence formulas for $P_s(m, n)$, which provides a rapid computational scheme for the problem. For a specific FSC model, $P_s(m, n)$ can be expressed as $[s^m z^n]P(s, z)$, where $P(s, z)$ is the ratio of two polynomials in s and z . The denominator polynomial is responsible for the recurrence relation, and the denominator polynomial defines the initial conditions. For example, it is easy to show that $P_s(m, n)$ for a GEC model satisfies a six-term recurrence formula.

Figure 2 shows $P_s(m, n)$, as a function of m , for GEC models with various values of memory, denoted as μ , which is defined in [9] as $\mu = (1 - q - Q)$, for $n = 127$ ($c = 7$). Throughout this section we will consider the following channel parameters, $\rho = q/Q = 20$, $b = 0.4$, $g = 0.001$. The model parameters Q and q are uniquely determined from μ and ρ . Because the GEC model has a parameter that can be interpreted as the memory of the channel, the effectiveness of coding schemes under several memory conditions can be evaluated.

The average number of erroneous symbols in a received word of length n is $\bar{n}_s = n \Pi^T (\mathbf{P}^c - \mathbf{P}(0)^c) \mathbf{1}$. Examples of values of \bar{n}_s are 12 for $\mu = 0.6$, 10 for $\mu = 0.8$, 8 for $\mu = 0.92$, 7 for $\mu = 0.96$. The curves of Figure 2 show that for $\mu < 0.8$ the probability $P_s(m, 127)$ has a maxima roughly centered around \bar{n}_s . This is a typical behavior of memoryless channels. Therefore, over the span of $127 \times 7 = 889$ bits, models with $\mu < 0.8$ make sufficient state transitions to assure

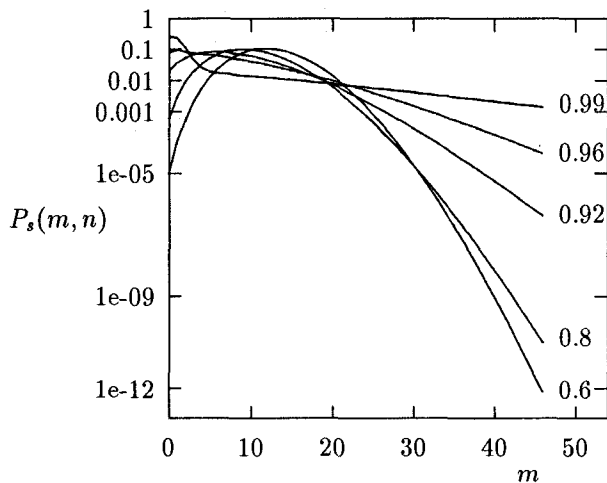


Fig. 2: $P_s(m, n)$ as a function of m , for $n = 127$, having the memory μ as a parameter, $\mu = 0.6, 0.8, 0.92, 0.96, 0.99$.

this “random” behavior. However, when the memory increases, fewer transitions occur between states and long bursts are more likely. As a consequence, $P_s(m, 127)$ spreads out and decreases slowly with m . To take an example, the curves show that $P_s(0, 127) = 0.36$ for $\mu = 0.99$. This is the probability of being in the good state during all 889 bit intervals and making no error, which is equal to, $\pi_0((1-Q)(1-g))^{889} = 0.36$. The contribution of any other state sequence to $P_s(m, 127)$ is negligible. In the sequel we will discuss the effect of memory on the codeword error probability of RS codes. The decoder will decode the received word to the correct (transmitted) codeword if no more than t error symbols occur in one codeword. The probability of codeword error (PCE) is defined as the probability of occurrence of received words with more than t erroneous symbols. Thus

$$\text{PCE} = 1 - \sum_{m=0}^t P_s(m, n) = \sum_{m=t+1}^n P_s(m, n).$$

Figure 3 shows PCE for RS codes with fixed length n , versus the memory μ , for various values of k (number of information symbols). In this analysis we consider $\text{PCE} = 10^{-6}$ the required error probability for reliable communication. We can conclude from these plots that for a particular value of μ , say $\mu = 0.92$, PCE equal to 10^{-6} is achieved with the rate $R_c = 49/127 = 0.39$ for $c = 7$ ($n = 127$), and $R_c = 141/255 = 0.55$ for $c = 8$ ($n = 255$). It was also observed that for a fixed k , PCE is minimum for $\mu = 0.6$. In fact, the curves stress two distinct modes of behavior of PCE, depending upon the burst length. In the region of short bursts, say $\mu < 0.6$, as the memory increases the error bits become more concentrated within bursts and affect fewer symbols in a codeword. Therefore, short bursts help the performance of RS decoders. On the other hand, in the region of high memory, say $\mu > 0.8$, where long bursts occur, reliable communication is possible only with longer low rate codes.

One possible way to cope with long burst errors is to spread

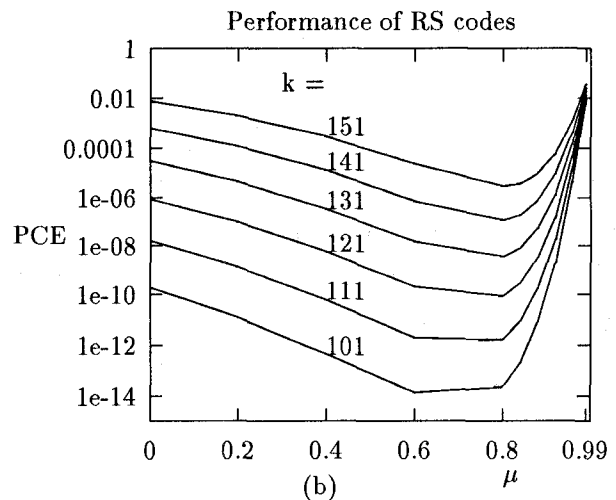
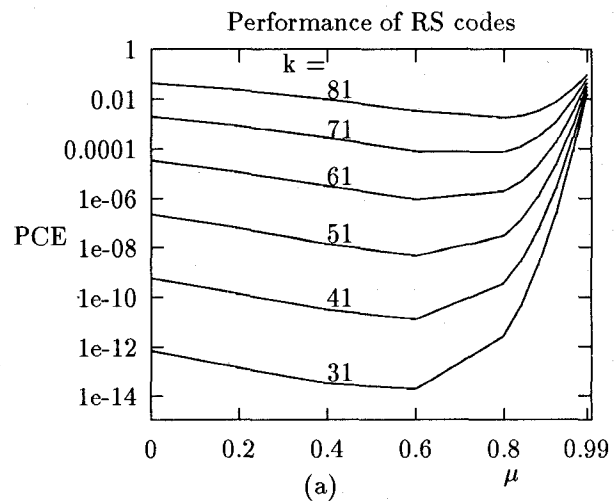


Fig. 3: PCE versus memory μ , for RS codes over GEC having the number of information symbols k as a parameter, for $c = 7$ (a); $c = 8$ (b).

the burst over many codewords so that a small fraction of the burst can hit the same codeword. This task is usually performed by an interleaver, which is a process of reordering a sequence of transmitted symbols in a one-to-one deterministic manner. At the receiver, the deinterleaver restores the original order of the transmitted symbols. One important practical consideration is how large the interleaving depth, denoted by I_d , should be in order to be considered infinite. We denote this ideal value of the interleaving depth by I_d^* . This idealized assumption may result in excessive memory requirement and decoding delay. Therefore, it is not possible to eliminate the channel memory entirely (the choice of I_d^* is not feasible), but only to reduce the burst severity. Another issue of interest is the degradation in performance resulting from finite interleaving depth. The problem of analyzing theoretically the performance of RS codes using a particular non-ideal interleaving will be considered in the next section.

3.1. The Effect of Interleaving

The aim of this section is to study the effect of symbol interleaving in a block coded system over FSC models. We assume an (I_d, n) block interleaver/deinterleaver. Such an interleaver

consists of an array of n (codeword length) columns and I_d (interleaving depth) rows, where each entry of the array stores one RS symbol (or c bits). The symbols are written into the array by rows and read out by columns. The deinterleaver performs the inverse operation, that is, the received symbols are written into by columns and read out to the decoder by rows. Notice that two consecutive received symbols within a received word are corrupted by two error symbols separated exactly by I_d . The probability of m erroneous received symbols in a received word of length n for the interleaved channel, denoted by $P_s^I(m, n)$, can be obtained from Equation (2) where F_c and F_e are replaced by F_c^I and F_e^I by applying the mapping:

$$\begin{aligned} F_{cI} &\mapsto F_c F_{\Omega^{I_d-1}}; \\ F_{eI} &\mapsto F_e F_{\Omega^{I_d-1}}, \end{aligned}$$

where $F_{\Omega^{I_d-1}} = (x_0 + x_1)^{(I_d-1)c}$ is the generating series for Ω^{I_d-1} , the set of all sequences of $I_d - 1$ symbols or $(I_d - 1)c$ bits, and F_c and F_e are given by Equations (1). So

$$\begin{aligned} P_s^I(m, n) &= \\ &= [s^m z^n] \Pi^T \Delta (1 - z(F_{cI} + s F_{eI}))^{-1} \mathbf{1}; \\ &= [s^m z^n] \Pi^T \Delta (1 - z(F_c + s F_e) (x_0 + x_1)^{(I_d-1)c})^{-1} \mathbf{1}; \\ &= [s^m z^n] \Pi^T (\mathbf{I} - z\{\mathbf{P}(0)^c + s(\mathbf{P}^c - \mathbf{P}(0)^c)\} \mathbf{P}^{(I_d-1)c})^{-1} \mathbf{1}. \end{aligned}$$

The effect of I_d on PCE when the memory varies is shown in Figure 4. This figure depicts PCE for the interleaved (127,71) RS code, as a function of the memory, for several values of I_d . We conclude from the plots that the optimum interleaving depth I_d^* varies with the memory in the following way: $I_d^* = 1$ for $0 \leq \mu < 0.4$, $I_d^* = 2$ for $0.4 \leq \mu < 0.6$, $I_d^* = 4$ for $0.6 \leq \mu < 0.84$, $I_d^* = 8$ for $0.84 \leq \mu < 0.92$, $I_d^* = 16$ for $0.92 \leq \mu < 0.96$, $I_d^* = 64$ for $0.96 \leq \mu < 0.99$.

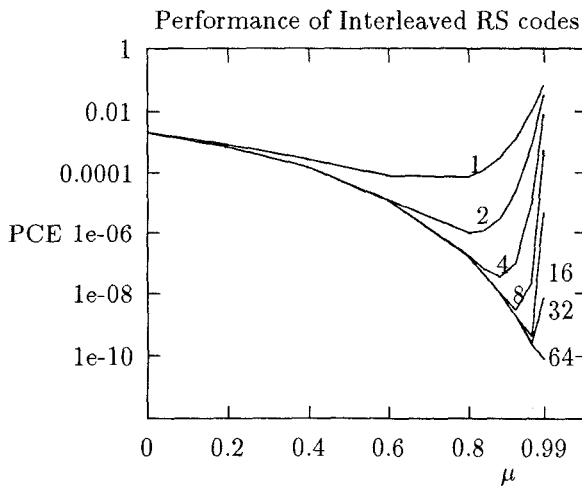


Fig. 4: PCE for the interleaved (127,71) RS code, as a function of memory, having I_d as a parameter. $I_d = 1, 2, 4, 8, 16, 32, 64$.

Several combinations of rate and interleaving degree can be used to achieve $\text{PCE} = 10^{-6}$, as we can see from Figure 5. This figure shows PCE as a function of t (error correcting capability) for various values of I_d . For the GEC model considered, $\mu = 0.92$, we can see that the following combination of RS codes and interleaving can be used: (127,59) RS and

$I_d = 2$; (127,75) RS and $I_d = 4$; (127,79) RS and $I_d = 8$. Therefore, the trade-off between code rate and interleaving depth is clear from these plots.

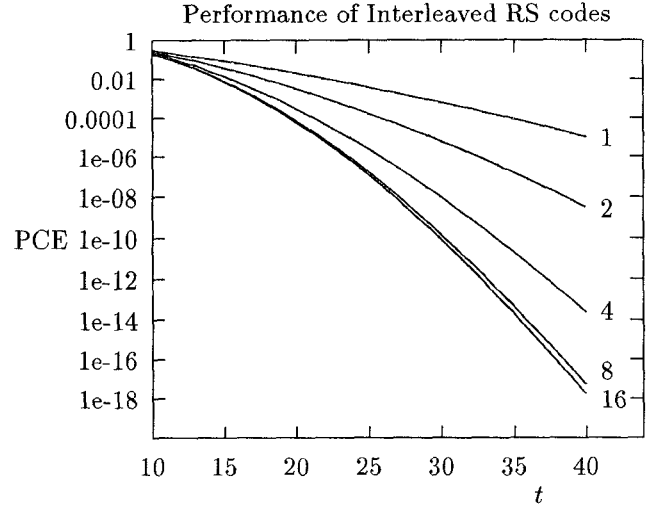


Fig. 5: PCE as a function of t , for (127, k) RS codes, $\mu = 0.92$, having I_d as a parameter. $I_d = 1, 2, 4, 8, 16$.

4. Conclusions

We have studied the evaluation of the codeword error probability of error correcting codes on FSC models. The main idea to find an expression for this measure for a specific coding scheme is to express the probability of the correctable error pattern as a coefficient in a formal power series. All numerical plots presented in this chapter were generated by first deriving recurrence formulas from the matrix expressions. Therefore, the performance of longer codes can be easily evaluated.

In the first part of this paper we have derived an expression for the codeword error probability of non-interleaved RS codes. Next, we have considered the cascade of coding and interleaving. The results presented here have been applied to investigate the tradeoffs among the code rate, error-correcting capability, and interleaving depth to achieve a desired performance. We have also compared the performance of interleaved binary codes with interleaved non-binary codes under the same memory and delay requirements.

The majority of previous analytical works of FEC schemes on channel with memory (e.g. Rician fading channel) have assumed ideal interleaving. An important application of modeling the real channel with FSC models is that we can use powerful combinatorial techniques to derive various statistics of the burst process necessary for the design and analysis of coding schemes and interleaving with finite interleaving depth.

References

- [1] "Interleaved Coding for Burst Channels", Cyclotomics' Final Report on Phase I, Small Business Innovation Research, NSF Grant No. ECS-8260180, April, 1983.
- [2] H. Bischl and E. Lutz, "Packet Error Rate in the Non-interleaved Rayleigh Channel", IEEE Transactions on Communications, Vol.43, pp.1375-1383, April, 1995.

- [3] J. R. Yee and E. J. Weldon Jr., "Evaluation of the Performance of Error-Correcting Codes on a Gilbert Channel", IEEE Transactions on Communications, Vol.43, pp.2316-2323, August, 1995.
- [4] L. N. Kanal and A.R.K. Sastry, "Models for Channels with Memory and Their Applications to Error Control", Proceedings of the IEEE, Vol. 66, pp. 724-744, July 1987.
- [5] A. I. Drukarev and K. P. Yiu, "Performance of Error-Correcting Codes on Channels with Memory", IEEE Transactions on Communications, Vol.34, pp.513-521, June,1986.
- [6] T. Berman and J. Freedman, "Non-Interleaved Reed-Solomon Coding Over a Bursty Channel", IEEE Military Communications Conference - MILCOM, San Diego - CA, pp.25.3.1-25.3.4, October 1992.
- [7] D. Kim and Lloyd R. Welch and R. E. Peile, "Correlation Effect to Reed-Solomon Code Performance and Experiment of Error/Erasur Decoder Co-operating with Interleaver and Hidden Markov Modeler on Satellite Channel", IEEE Military Communications Conference - MILCOM, San Diego - CA, pp. 25.4.1-25.4.5, October 1992.
- [8] C. Pimentel and I.F. Blake, "Derivation of Burst Error Statistics for Finite State Channel Models Using a Combinatorial Approach", Proceedings of the 30th Annual Conference on Information Sciences and Systems, Princeton, USA, March 1996.
- [9] M. Mushkin and I. Bar-David, "Capacity and Coding for the Gilbert-Elliott Channels", IEEE Transactions on Information Theory, Vol. 35, pp. 1277-1289, November, 1989.