# Application of Distance Enhancing Codes

Emina Soljanin, *Member, IEEE* and Adriaan J. van Wijngaarden, *Member, IEEE*

***Abstract*—During the past few years, significant progress has been made in defining high capacity constraints which prohibit specified differences between constrained sequences, thus ensuring that the minimum distance between them is larger than for the uncoded system. However, different constraints which avoid the same prescribed set of differences may have different capacities, and codes into such constraints may have different complexity of encoder/decoder architecture and different performance on more realistic channel models. These issues, which have to be considered in application of distance enhancing codes, are discussed here. We define several distance enhancing constraints which support design of high rate codes. We also define weak constraints for which the minimum distance between sequences may be the same as for the uncoded system but the number of pairs of sequences at the minimum distance is smaller. These constraints support design of even higher rate codes. We discuss the implementation issues of both types of constraints as well as their performance on the ideal channel and channels with colored noise and intertrack interference.**

***Index Terms*—Constrained coding, efficient implementation, intersymbol interference, partial response, recording.**

## I. Introduction

THE DATA recording and retrieval process is most commonly modeled as a linear, discrete-time, communications channel with inter-symbol interference (ISI), described by its transfer function and white Gaussian noise. The transfer function is given by $h(D) = (1 - D)(1 + D)^N$, where $N$ depends on and increases with the linear recording density. Several types of distance enhancing codes have been proposed for these binary-input ISI channels, such as codes with good Hamming distance on a precoded channel, codes whose spectral nulls match those of the channel, and, more recently, codes which prohibit specified differences between code sequences.

Only the last class of codes is used in commercial magnetic recording systems. There are two main reasons for this: these codes simplify the channel detectors relative to the uncoded channel and even high rate codes in this class can be realized by low complexity encoders and decoders. We here study a number of practical issues which made particular codes in this class preferable for implementation.

During the past few years, significant progress has been made in defining high capacity distance enhancing constraints for high density magnetic recording channels (see [1] and references therein). Researchers in the area quickly realized that different constraints which avoid the same prescribed set of differences may have different capacities. Thus, most of the initial research focused on finding high capacity constraints, and bounds on the capacities were recently presented in [8]. We explain the main idea behind this type of distance enhancing codes in Section II, and present several high rate constraints in Section III.

All the constraints presented in [1] eliminate some of the possible recorded sequences with a goal to increase the minimum distance between those that remain. Often, to achieve this goal, a large fraction of the possible recorded sequences has to be eliminated and this adversely affects the maximum code rate into the constraint. We here define constraints for which the minimum distance between sequences may be the same as for the uncoded system but the number of pairs of sequences at the minimum distance is smaller. (Note that this is also the case with turbo codes.) We refer to them as *weak* as opposed to *strong* constraints. Several weak constraints are also presented in Section III. The performance of some examples of both types of constraints was checked by simulation, and the results are presented in Section IV. In some cases weak constraints outperform the strong ones because of their higher rate.

In practice, to translate information bits into the sequences that obey given constraints and vice versa, one needs an encoding and decoding mapping. For high capacity constraints, the usual graph transforming techniques result in encoder/decoder architectures with formidable complexity. Fortunately, a block encoder/decoder architecture with acceptable implementation complexity for the constraints proposed here can always be designed by well known enumerative techniques [5]. A particularly efficient architecture based on the combinatorial techniques of [6] for one of the considered constraints is presented in [9] and outlined in Section V.

Coding for more realistic recording channel models that include colored noise and intertrack interference are discussed in Section VI. We point out that different constraints which avoid the same prescribed set of differences may have different performance on more realistic channels. This makes some of them more attractive for implementation.

## II. Constraints for ISI Channels

### A. Requirements

A number of papers have proposed using constrained codes to provide coding gain on channels with high ISI (see for example [1]–[4]). The main idea of this approach can be described as follows [1]. Consider a discrete-time model for the magnetic recording channel with possibly constrained input $\boldsymbol{a} = \{a_n\} \in \mathcal{C} \subseteq \{-1, 1\}^\infty$, impulse response $\{h_n\}$, and output $\boldsymbol{y} = \{y_n\}$ given by

$$y_n = \sum_m a_m h_{n-m} + \eta_n, \tag{1}$$

Manuscript received November 16, 2000.

The authors are with the Mathematical Sciences Research Center, Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: {emina; alw}@lucent.com).

Publisher Item Identifier S 0018-9464(01)02390-1.

0018–9464/01$10.00 © 2001 IEEE

where $h(D) = \sum_n h_n D^n = (1-D)(1+D)^3$ ($E^2$PR4) or $h(D) = \sum_n h_n = (1-D)(1+D)^4$ ($E^3$PR4), $\eta_n$ are independent Gaussian random variables with zero mean and variance $\sigma^2$. The quantity $1/\sigma^2$ is referred to as the signal-to-noise ratio (SNR). The minimum distance of the uncoded channel (1) is

$$d^2_{\min} = \min_{\epsilon(D)\neq 0} ||h(D)\epsilon(D)||^2,$$

where $\epsilon(D) = \sum_{i=0}^{l-1} \epsilon_i D^i$, ($\epsilon_i \in \{-1, 0, 1\}$, $\epsilon_0 = 1$, $\epsilon_{l-1} \neq 0$) is the polynomial corresponding to a normalized input error sequence $\boldsymbol{\epsilon} = \{\epsilon_i\}_{i=0}^{l-1}$ of length $l$, and the squared norm of a polynomial is defined as the sum of its squared coefficients. The minimum distance is bounded from above by $||h(D)||^2$, denoted by

$$d^2_{\mathrm{MFB}} = ||h(D)||^2. \tag{2}$$

This bound is known as the *matched-filter bound* (MFB), and is achieved when the error sequence of length $l = 1$, i.e., $\epsilon(D) = 1$, is in the set

$$\arg \min_{\epsilon(D)\neq 0} ||h(D)\epsilon(D)||^2. \tag{3}$$

For channels that fail to achieve the MFB, i.e., for which $d^2_{\min} < ||h(D)||^2$, any error sequences $\epsilon(D)$ for which

$$d^2_{\min} \leq ||h(D)\epsilon(D)||^2 < ||h(D)||^2 \tag{4}$$

are of length $l \geq 2$ and may belong to a constrained system $X_{\mathcal{L}}^{\{-1,0,1\}}$, where $\mathcal{L}$ is an appropriately chosen finite list of forbidden strings.

For code $\mathcal{C}$, we write the set of all admissible nonzero error sequences as

$$\mathcal{E}(\mathcal{C}) = \{\boldsymbol{\epsilon} \in \{-1, 0, 1\}^\infty | \boldsymbol{\epsilon} \neq \mathbf{0}, \boldsymbol{\epsilon} = (\boldsymbol{a} - \boldsymbol{b})/2, \boldsymbol{a}, \boldsymbol{b} \in \mathcal{C}\}.$$

Given the condition $\mathcal{E}(\mathcal{C}) \subseteq X_{\mathcal{L}}^{\{-1,0,1\}}$, we seek to identify the least restrictive finite collection $\mathcal{F}$ of blocks over the alphabet $\{0, 1\}$ so that

$$\mathcal{C} \subseteq X_{\mathcal{F}}^{\{0,1\}} \Longrightarrow \mathcal{E}(\mathcal{C}) \subseteq X_{\mathcal{L}}^{\{-1,0,1\}}. \tag{5}$$

### B. Definitions

A constrained code is defined by specifying $\mathcal{F}$, the list of forbidden strings for code sequences. Prior to that one needs to first characterize error sequences that satisfy (4) and then specify $\mathcal{L}$, the list of forbidden strings for error sequences. Error event characterization can be done by using any of the methods described by Karabed, Siegel, and Soljanin in [1]. Specification of $\mathcal{L}$ is usually straightforward.

A natural way to construct a collection $\mathcal{F}$ of blocks forbidden in code sequences based on the collection $\mathcal{L}$ of blocks forbidden in error sequences is the following. From the above definition of error sequences $\boldsymbol{\epsilon} = \{\epsilon_i\}$ we see that $\epsilon_i = 1$ requires $a_i = 1$ and $\epsilon_i = -1$ requires $a_i = 0$, i.e., $a_i = (1 + \epsilon_i)/2$. For each block $\boldsymbol{w}_{\mathcal{E}} \in \mathcal{L}$, construct a list $\mathcal{F}_{\boldsymbol{w}_{\mathcal{E}}}$ of blocks of the same length $l$ according to the rule:

$$\mathcal{F}_{\boldsymbol{w}_{\mathcal{E}}} = \{\boldsymbol{w}_{\mathcal{C}} \in \{-1, 1\}^l | \\ w^i_{\mathcal{C}} = (1 + w^i_{\mathcal{E}})/2 \text{ for all } i \text{ for which } w^i_{\mathcal{E}} \neq 0\}.$$

Then the collection $\mathcal{F}$ obtained as $\mathcal{F} = \bigcup_{\boldsymbol{w}_{\mathcal{E}} \in \mathcal{L}} \mathcal{F}_{\boldsymbol{w}_{\mathcal{E}}}$ satisfies requirement (5). However, the constrained system $X_{\mathcal{F}}^{\{0,1\}}$ obtained this way may not be the most efficient. (Bounds on the achievable rates of codes which avoid specified differences were found recently in [8]).

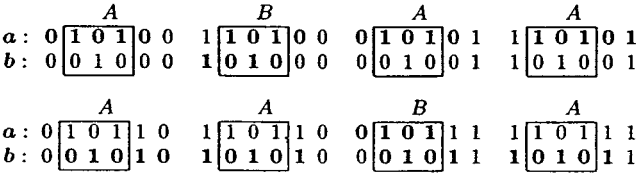



Fig. 1. Possible pairs of sequences for which error event $+-+00$ may occur.

We illustrate the above ideas on the example of the $E^2$PR4 channel. Its transfer function is $h(D) = (1-D)(1+D)^3$, and its MFB is $||(1-D)(1+D)^3 \cdot 1||^2 = 10$. The error polynomial $\epsilon(D) = 1 - D + D^2$ is the unique error polynomial for which $||(1-D)(1+D)^3\epsilon(D)||^2 = 6$, and the error polynomials $\epsilon(D) = 1 - D + D^2 + D^5 - D^6 + D^7$ and $\epsilon(D) = \sum_{i=0}^{l-1}(-1)^i D^i$ for $l \geq 4$ are the only polynomials for which $||(1-D)(1+D)^3\epsilon(D)||^2 = 8$ (see for example [1]).

It is easy to show that these error events are not in the constrained error set defined by the list of forbidden error strings $\mathcal{L} = \{+-+00, +-+-\}$, where $+$ denotes 1 and $-$ denotes $-1$. To see that, note that an error sequence that does not contain the string $+-+00$ cannot have error polynomials $\epsilon(D) = 1 - D + D^2$ or $\epsilon(D) = 1 - D + D^2 + D^5 - D^6 + D^7$, while an error sequence that does not contain string $+-+-$ cannot have an error polynomial of the form $\epsilon(D) = \sum_{i=0}^{l-1}(-1)^i D^i$ for $l \geq 4$. Therefore, by the above procedure of defining the list of forbidden code strings, we obtain the $\mathcal{F} = \{+-+\}$ NRZ constraint. Its capacity is about 0.81, and a rate 4/5 code into the constraint was first given in [7].

In [1], the following approach was used to obtain several higher rate constraints. For each of error strings in $\mathcal{L}$, we write all pairs of channel strings whose difference is the error string. To define $\mathcal{F}$, we look for the longest string(s) appearing in at least one of the strings in each channel pair. For the example above and the $+-+00$ error string, a case-by-case analysis of channel pairs is depicted in Fig. 1. We can distinguish two types (denoted by $A$ and $B$ in the figure) of pairs of code sequences involved in forming an error event. In a pair of type $A$, at least one of the sequences has a transition run of length 4. In a pair of type $B$, both sequences have transition runs of length 3, but for one of them the run starts at an even position and for the other at an odd position. This implies that an NRZI constrained system that limits the run of 1 s to 3 when it starts at an odd position, and to 2 when it starts at an even position, eliminates all possibilities shown bold-faced in Fig. 1. In addition, this constraint eliminates all error sequences containing the string $+-+-$. The capacity of the constraint is about 0.916, and rate 8/9 block codes with this constraint has been implemented in several commercial read channel chips. More about the constraint and the codes can be found in [1]–[4].

## III. High Capacity Constraints for the $E^2$PR4 Channel

We now present several high capacity constraints for the $E^2$PR4 channel derived by using the approach presented at the end of Section II-B. In each case, we state the constraint

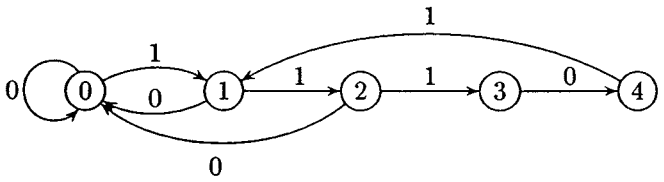
and show the strings it removes. Note that the higher capacity constraints do not remove all channel pairs involved in forming the error event $+-+00$. These are called *weak* constraints and those which remove all pairs are called *strong* constraints. For each proposed constraint, we give a graph representation, its adjacency matrix, and the Shannon capacity.

*A.* $\mathcal{F} = \{1111, 11100\}$ *NRZI Strong Constraint*

Strings eliminated by the constraint:

| | | | | |
|---|---|---|---|---|
| ***a***: | $0\boxed{1\,0\,1}0\,0$ | $1\boxed{1\,0\,1}0\,0$ | $0\boxed{1\,0\,1}0\,1$ | $1\boxed{1\,0\,1}0\,1$ |
| ***b***: | $0\boxed{0\,1\,0}0\,0$ | $1\boxed{0\,1\,0}0\,0$ | $0\boxed{0\,1\,0}0\,1$ | $1\boxed{0\,1\,0}0\,1$ |
| ***a***: | $0\boxed{1\,0\,1}1\,0$ | $1\boxed{1\,0\,1}1\,0$ | $0\boxed{1\,0\,1}1\,1$ | $1\boxed{1\,0\,1}1\,1$ |
| ***b***: | $0\boxed{0\,1\,0}1\,0$ | $1\boxed{0\,1\,0}1\,0$ | $0\boxed{0\,1\,0}1\,1$ | $1\boxed{0\,1\,0}1\,1$ |

Graph representation of the constraint:




The adjacency matrix and the capacity:

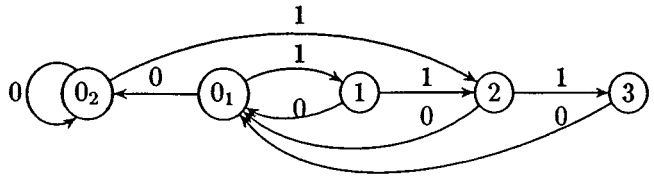
$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad C = 0.913\ldots$$

*B.* $\mathcal{F} = \{1111, 00111\}$ *NRZI Strong Constraint*

Strings eliminated by the constraint:

| | | | | |
|---|---|---|---|---|
| ***a***: | $0\,0\boxed{1\,0\,1}0$ | $0\,1\boxed{1\,0\,1}0$ | $1\,0\boxed{1\,0\,1}0$ | $1\,1\boxed{1\,0\,1}0$ |
| ***b***: | $0\,0\boxed{0\,1\,0}0$ | $0\,1\boxed{0\,1\,0}0$ | $1\,0\boxed{0\,1\,0}0$ | $1\,1\boxed{0\,1\,0}0$ |
| ***a***: | $0\,0\boxed{1\,0\,1}1$ | $0\,1\boxed{1\,0\,1}1$ | $1\,0\boxed{1\,0\,1}1$ | $1\,1\boxed{1\,0\,1}1$ |
| ***b***: | $0\,0\boxed{0\,1\,0}1$ | $0\,1\boxed{0\,1\,0}1$ | $1\,0\boxed{0\,1\,0}1$ | $1\,1\boxed{0\,1\,0}1$ |

Graph representation of the constraint:




The adjacency matrix and the capacity:

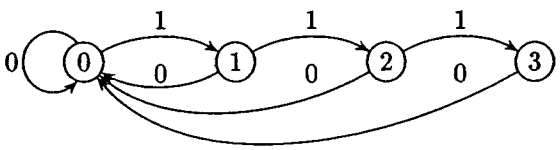
$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad C = 0.913\ldots$$

*C.* $\mathcal{F} = \{1111\}$ *NRZI Weak Constraint*

Strings eliminated by the constraint:

| | | | | |
|---|---|---|---|---|
| ***a***: | $0\boxed{1\,0\,1}0\,0$ | $1\boxed{1\,0\,1}0\,0$ | $0\boxed{1\,0\,1}0\,1$ | $1\boxed{1\,0\,1}0\,1$ |
| ***b***: | $0\boxed{0\,1\,0}0\,0$ | $1\boxed{0\,1\,0}0\,0$ | $0\boxed{0\,1\,0}0\,1$ | $1\boxed{0\,1\,0}0\,1$ |
| ***a***: | $0\boxed{1\,0\,1}1\,0$ | $1\boxed{1\,0\,1}1\,0$ | $0\boxed{1\,0\,1}1\,1$ | $1\boxed{1\,0\,1}1\,1$ |
| ***b***: | $0\boxed{0\,1\,0}1\,0$ | $1\boxed{0\,1\,0}1\,0$ | $0\boxed{0\,1\,0}1\,1$ | $1\boxed{0\,1\,0}1\,1$ |

Graph representation of the constraint:




The adjacency matrix and the capacity:

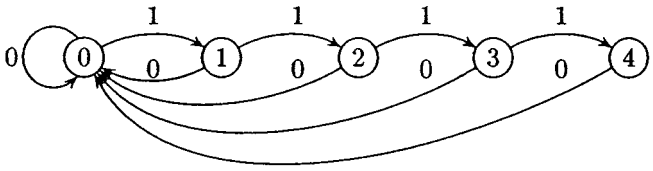
$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad C = 0.946\ldots$$

*D.* $\mathcal{F} = \{11111\}$ *NRZI Weak Constraint*

Strings eliminated by the constraint:

| | | | | |
|---|---|---|---|---|
| ***a***: | $0\boxed{1\,0\,1}0\,0$ | $1\boxed{1\,0\,1}0\,0$ | $0\boxed{1\,0\,1}0\,1$ | $1\boxed{1\,0\,1}0\,1$ |
| ***b***: | $0\boxed{0\,1\,0}0\,0$ | $1\boxed{0\,1\,0}0\,0$ | $0\boxed{0\,1\,0}0\,1$ | $1\boxed{0\,1\,0}0\,1$ |
| ***a***: | $0\boxed{1\,0\,1}1\,0$ | $1\boxed{1\,0\,1}1\,0$ | $0\boxed{1\,0\,1}1\,1$ | $1\boxed{1\,0\,1}1\,1$ |
| ***b***: | $0\boxed{0\,1\,0}1\,0$ | $1\boxed{0\,1\,0}1\,0$ | $0\boxed{0\,1\,0}1\,1$ | $1\boxed{0\,1\,0}1\,1$ |

Graph representation of the constraint:




The adjacency matrix and the capacity:

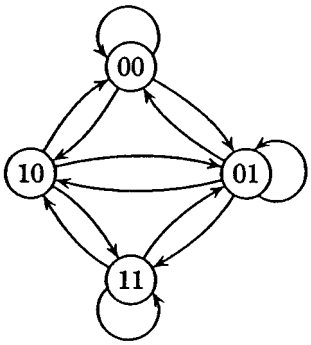
$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad C = 0.975\ldots$$

*E.* $\mathcal{F} = \{1010\}$ *NRZ Weak Constraint*

Strings eliminated by the constraint:

| | | | | |
|---|---|---|---|---|
| ***a***: | $0\boxed{1\,0\,1}0\,0$ | $1\boxed{1\,0\,1}0\,0$ | $0\boxed{1\,0\,1}0\,1$ | $1\boxed{1\,0\,1}0\,1$ |
| ***b***: | $0\boxed{0\,1\,0}0\,0$ | $1\boxed{0\,1\,0}0\,0$ | $0\boxed{0\,1\,0}0\,1$ | $1\boxed{0\,1\,0}0\,1$ |
| ***a***: | $0\boxed{1\,0\,1}1\,0$ | $1\boxed{1\,0\,1}1\,0$ | $0\boxed{1\,0\,1}1\,1$ | $1\boxed{1\,0\,1}1\,1$ |
| ***b***: | $0\boxed{0\,1\,0}1\,0$ | $1\boxed{0\,1\,0}1\,0$ | $0\boxed{0\,1\,0}1\,1$ | $1\boxed{0\,1\,0}1\,1$ |

Graph representation of the constraint (the edge labels are the same as their terminal state labels):




The adjacency matrix and the capacity:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \qquad C = 0.961\ldots$$

## IV. Simulation Results

Among the constraints presented in the previous section, those with capacities higher than 16/17 are of special practical interest. Two rate 16/17 block codes for the weak constraints $\mathcal{F} = \{1111\}$ NRZI and $\mathcal{F} = \{11\,111\}$ NRZI, with very efficient encoder/decoder architecture, were designed by using techniques presented in [9] and briefly discussed in Section V. In order to make a 16/17 block code into the constraint system $\mathcal{F} = \{1111\}$ NRZI, the codewords are allowed to have strings 11 at the beginning and the end. Therefore, the resulting code permits string 1111 at those places. We also refer to this code as a (0, 3, 2, 2) code to indicate that the minimum number of 1 s after a 0 is 0, the maximum number of consecutive "ones" within a codeword is 3 with at most two consecutive "ones" at the end and the beginning of a codeword.

We evaluated the bit-error rate performance of these codes by simulation, and compared the results with the performance of several codes into strong constraints presented in [1], and references therein. For each tested code, binary information is converted into codewords, and then into channel symbols by appropriate (NRZ or NRZI) modulation rule. Transmission over the noisy $E^2$PR4 channel was simulated by computing channel outputs and adding white Gaussian noise. At the decoder the received signal is demodulated using a Viterbi detector with memory 15. The bit error rate is measured for a range of the channel SNRs. The simulation results are shown in Fig. 2. The results show remarkably good performance of the 16/17, (0, 3, 2, 2) code. The rate of the code, its performance, and the extremely efficient encoder/decoder architecture make it the most desirable for the current linear recording densities.

## V. Constraint Code Design

Let a $(0,\, k,\, k_l,\, k_r)$ block code of length $n$ denote a set of words with at most $k$ consecutive "ones." The parameters $k_l$ and $k_r$ denote the maximum number of leading and trailing *ones* of a codeword. We briefly describe an efficient method for construction of $(0,\, k,\, k_l,\, k_r)$ codes by mapping $(n-1)$-bit

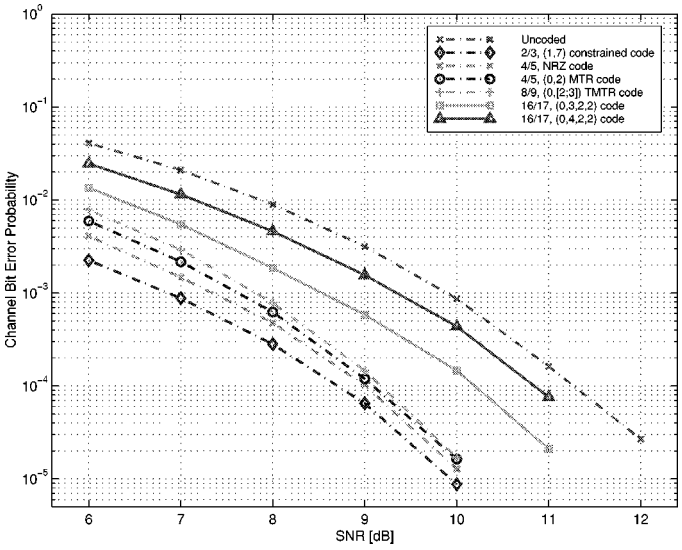



Fig. 2. Bit error rate of several constrained codes as a function of SNR for $E^2$PR4 channels.

data words onto $n$-bit codewords of a $(0,\, k,\, k_l,\, k_r)$ code presented in [9]. If the number of sequences of length $n$ that satisfy the $(0,\, k,\, k_l,\, k_r)$ constraints exceeds $2^{n-1}$, it is always possible to perform this mapping, although the complexity may be high. For several codeword lengths it is possible to realize this mapping procedure using a small digital combinatorial circuit. The codes that have been developed are characterized by a set of data-dependent mapping rules. The encoder first checks whether the data word satisfies the constraint. If this is the case, the encoder copies the data word and inserts one extra bit that indicates that the other bits of the codeword are a copy of the data word. Otherwise, the extra bit indicates that a mapping rule had to be applied to transform the data word into an $(n-1)$-bit coded sequence that satisfies the imposed $(0,\, k,\, k_l,\, k_r)$ constraints. The coded sequence, which uniquely represents the data word, consists of bits which identify the mapping used and how the data bits were rearranged. The mapping structure is often symmetric to simplify the design process and the structure of the digital combinatorial circuitry.

The construction technique has the following characteristics:

- A mapping is defined for a class of data words rather than for one word. Based on the specific constraints, the set of data words and the set of codewords are partitioned into disjoint subsets. For every set a mapping is specified. The words belonging to the same subset are transformed in the same fashion.
- Symmetry is used whenever possible, i.e., if $Y = \psi(X)$ represents the mapping of $X$ onto $Y$, then $X'$ and $Y'$, representing the words $X$ and $Y$ in reversed order, satisfy the relation $Y' = \psi(X')$ whenever possible. This significantly simplifies the design.
- The relative position of the information bits is changed only if necessary. Whenever changes are necessary to fulfill the constraints, the data symbols are permuted instead of transformed.
- The translation of the $(n-1)$ information bits into $n$-bit codewords and vice versa is done in parallel.

- The combinatorial circuitry that results from the technique classifies the incoming data word, and rearranges the data bits based on the classification in order to satisfy the constraints.

It is particularly interesting to use this technique for constructing tightly constrained codes with a short and medium word length (4 to 32 bits, and possibly up to 64 bits with slightly looser constraints). The constructed codes can also be used as a basic code to obtain longer and higher rate codes with less tight constraints using interleaving with uncoded data symbols. The proposed techniques are discussed in more detail in [9], and illustrated by constructing a rate 16/17, (0, 3, 2, 2) MTR code. The code also fulfills (0, 15, 9, 9) RLL constraints, and can be used as a basic code for the construction of higher rate constrained codes by interleaving the codewords with uncoded bits.

## VI. Channels with Colored Noise and Intertrack Interference

Magnetic recording systems always operate in the presence of colored noise intertrack interference, and data dependent noise. Codes for these more realistic channel models are studied in [11]. Below, we briefly outline the problem.

Data recording and retrieval process is usually modeled as a linear, continuous-time, communications channel described by its Lorentzian step response and additive white Gaussian noise. The most common discrete-time channel model is given by (1). Magnetic recording systems employ channel equalization to the most closely matching transfer function $h(D) = \sum_n h_n D^n$ of the form $h(D) = (1 - D)(1 + D)^N$. This equalization alters the spectral density of the noise, and a better channel model assumes that the $\eta_n$ in (1) are identically distributed, Gaussian random variables with zero mean, variance $\sigma^2$, and normalized cross-correlation $E\{\eta_n \eta_k\}/\sigma^2 = \rho_{n-k}$.

In practice, there is always intertrack interference (ITI), i.e., the read head picks up magnetization from an adjacent track. Therefore, the channel output is given by

$$y_n = \sum_m a_m h_{n-m} + \sum_m x_m g_{n-m} + \eta_n, \tag{6}$$

where $\{g_n\}$ is the discrete-time impulse response of the head to the adjacent track, and $\boldsymbol{x} = \{x_n\} \in \mathcal{C}$ is the sequence recorded on that track. We assume that the noise is white.

In the ideal case (1), the probability of detecting $\boldsymbol{b}$ given that $\boldsymbol{a}$ was recorded is equal to $Q(d(\boldsymbol{\epsilon})/\sigma)$, where $d(\boldsymbol{\epsilon})$ is the distance between $\boldsymbol{a}$ and $\boldsymbol{b}$ given by

$$d^2(\boldsymbol{\epsilon}) = \sum_n \left( \sum_m \epsilon_m h_{n-m} \right)^2. \tag{7}$$

Therefore, a lower bound, and a close approximation for small $\sigma$, to the minimum probability of an error-event in the system is given by $Q(d_{\min, \mathcal{C}}/\sigma)$, where

$$d_{\min, \mathcal{C}} = \min_{\boldsymbol{\epsilon} \in \mathcal{E}_{\mathcal{C}}} d(\boldsymbol{\epsilon})$$

is the channel minimum distance of code $\mathcal{C}$. We refer to

$$d_{\min} = \min_{\boldsymbol{\epsilon} \in \{-1, 0, 1\}^\infty} d(\boldsymbol{\epsilon}) \tag{8}$$

as the minimum distance of the uncoded channel, and to the ratio $d_{\min, \mathcal{C}}/d_{\min}$ as the gain in distance of code $\mathcal{C}$ over the uncoded channel.

In the case of colored noise, the probability of detecting $\boldsymbol{b}$ given that $\boldsymbol{a}$ was recorded equals to $Q(\Delta(\boldsymbol{\epsilon})/\sigma)$, where $\Delta(\boldsymbol{\epsilon})$ is the distance between $\boldsymbol{a}$ and $\boldsymbol{b}$ given by

$$\Delta^2(\boldsymbol{\epsilon}) = \frac{\left[ \sum_n \left( \sum_m \epsilon_m h_{n-m} \right)^2 \right]^2}{\sum_n \sum_k \left( \sum_m \epsilon_m h_{n-m} \right) \rho_{n-k} \left( \sum_m \epsilon_m h_{k-m} \right)}.$$

Therefore a lower bound to the minimum probability of an error-event in the system is given by $Q(\Delta_{\min, \mathcal{C}}/\sigma)$, where

$$\Delta_{\min, \mathcal{C}} = \min_{\boldsymbol{\epsilon} \in \mathcal{E}_{\mathcal{C}}} \Delta(\boldsymbol{\epsilon}).$$

In the case of ITI (6), we are interested in the probability of detecting sequence $\boldsymbol{b}$ given that sequence $\boldsymbol{a}$ was recorded on the track being read and sequence $\boldsymbol{x}$ was recorded on an adjacent track. This probability is

$$Q(\delta(\boldsymbol{\epsilon}, \boldsymbol{x})/\sigma),$$

where $\delta(\boldsymbol{\epsilon}, \boldsymbol{x})$ is the distance between $\boldsymbol{a}$ and $\boldsymbol{b}$ in the *presence* of $\boldsymbol{x}$ given by [10]

$$\delta^2(\boldsymbol{\epsilon}, \boldsymbol{x}) = \frac{1}{\left[ \sum_n \left( \sum_m \epsilon_m h_{n-m} \right)^2 \right]} \cdot \left[ \sum_n \left( \sum_m \epsilon_m h_{n-m} \right)^2 + \sum_n \left( \sum_m x_m g_{n-m} \right) \left( \sum_m \epsilon_m h_{n-m} \right) \right]^2.$$

Therefore a lower bound to the minimum probability of an error-event in the system is proportional to $Q(\delta_{\min, \mathcal{C}}/\sigma)$, where

$$\delta_{\min, \mathcal{C}} = \min_{\boldsymbol{\epsilon} \neq 0, \boldsymbol{x} \in \mathcal{C}} \delta(\boldsymbol{\epsilon}, \boldsymbol{x}).$$

Distance $\delta_{\min, \mathcal{C}}$ can be bounded as follows [10]:

$$\delta_{\min, \mathcal{C}} \geq (1 - M)\, d_{\min, \mathcal{C}}, \tag{9}$$

where $M = \max_{n, \boldsymbol{x} \in \mathcal{C}} \sum_m x_m g_{n-m}$, i.e., $M$ is the maximum absolute value of the interference. Note that $M = \sum_n |g_n|$. We will assume that $M < 1$. The bound is achieved if and only if there exists an $\boldsymbol{\epsilon}$, $d(\boldsymbol{\epsilon}) = d_{\min, \mathcal{C}}$, for which $\sum_m \epsilon_m h_{n-m} \in \{-1, 0, 1\}$ for all $n$, and there exists an $\boldsymbol{x} \in \mathcal{C}$ such that $\sum_m x_m g_{n-m} = \mp M$ whenever $\sum_m \epsilon_m h_{n-m} = \pm 1$.

*An Example:* There are codes that provide gain in minimum distance on channels with ITI and colored noise, but not on the AWGN channel with the same transfer function. This is best illustrated using the example of the partial response channel with the transfer function $h(D) = (1-D)(1+D)^2$ known as EPR4. It is well known that for the EPR4 channel $d^2_{\min} = 4$. Moreover, the following result was shown in [10]:

*Proposition 1:* Error events $\epsilon(D)$ such that

$$d^2(\epsilon) = d^2_{\min} = 4$$

take one of the following two forms:

$$\epsilon(D) = \sum_{j=0}^{k-1} D^{2j}, \qquad k \geq 1$$

or

$$\epsilon(D) = \sum_{i=0}^{l-1} (-1)^i D^i, \qquad l \geq 3.$$

In the case of ITI (6), we assume that the impulse response to the reading head from an adjacent track is described by $g(D) = \alpha H(D)$, where the parameter $\alpha$ depends on the track to head distance. Under this assumption, the bound (9) gives $\delta^2_{\min} \geq d^2_{\min}(1-4\alpha)^2$. The following result was shown in [10]:

*Proposition 2:* Error events $\epsilon(D)$ such that

$$\min_{\boldsymbol{x} \in \mathcal{C}} \delta^2(\boldsymbol{\epsilon}, \boldsymbol{x}) = \delta^2_{\min} = d^2_{\min}(1-4\alpha)^2 = 4(1-4\alpha)^2$$

take the following form:

$$\epsilon(D) = \sum_{i=0}^{l-1} (-1)^i D^i, \qquad l \geq 5.$$

For all other error sequences for which $d^2(\boldsymbol{\epsilon}) = 4$, we have $\min_{\boldsymbol{x} \in \mathcal{C}} \delta^2(\boldsymbol{\epsilon}, \boldsymbol{x}) = 4(1-3\alpha)^2$.

Therefore, an improvement in error-probability performance of this channel can be accomplished by limiting the length of strings of alternating symbols in code sequences to four. For the NRZI type of recording, this can be achieved by a code that limits the runs of successive ones to three.

Channel equalization to the EPR4 target introduces cross-correlation among noise samples for a range of current linear recording densities (see [11] and references therein). The following result was obtained in [11]:

*Proposition 3:* Error events $\epsilon(D)$ such that

$$\Delta^2(\boldsymbol{\epsilon}) = \Delta^2_{\min}$$

take the following form:

$$\epsilon(D) = \sum_{i=0}^{l-1} (-1)^i D^i, \qquad l \geq 3, \quad l \text{ odd}.$$

Therefore, an improvement of error-probability performance can be accomplished by codes which eliminate the error sequences $\boldsymbol{\epsilon}$ containing the strings $-1+1-1$ and $+1-1+1$. Such codes were extensively studied in [1].

## VII. Conclusions

In order to apply distance enhancing codes, issues like code rate, encoder/decoder complexity and performance in real circumstances have to be considered. We presented several strong and weak constraints for the $E^2PR4$ channel, tested their performance by simulation on the ideal channel, and discussed implementation issues and performance in the presence of ITI and colored noise. This work illustrates that very often the code with the minimum distance on the ideal channel may not be the best choice in practice.

## References


[1] R. Karabed, P. H. Siegel, and E. Soljanin, "Constrained coding for binary channels with high intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1777–1797, Sept. 1999.

[2] W. G. Bliss, "An 8/9 rate time-varying trellis code for high density magnetic recording," *IEEE Trans. Magn.*, vol. 33, no. 5, pp. 2746–2748, Sept. 1997.

[3] K. K. Fitzpatrick and C. S. Modlin, "Time-varying MTR codes for high density magnetic recording," in *Proc. 1997 IEEE Global Telecommun. Conf. (GLOBECOM '97)*, Phoenix, AZ, Nov. 1997, pp. 1250–1253.

[4] J. Moon and B. Brickner, "Maximum transition run codes for data storage systems," *IEEE Trans. Magn.*, vol. 32, pp. 3992–3994, Sept. 1996.

[5] T. Cover, "Enumerative source encoding," *IEEE Trans. Inform. Theory*, pp. 73–77, Jan. 1973.

[6] A. J. van Wijngaarden and K. A. S. Immink, "Combinatorial construction of high rate runlength-limited codes," in *Proc. 1996 IEEE Global Telecommun. Conf. (GLOBECOM '96)*, London, U.K., Nov. 1996, pp. 343–347.

[7] R. Karabed and P. H. Siegel, "Coding for higher order partial response channels," in *Proc. 1995 SPIE Int. Symp. on Voice, Video, and Data Communications*, vol. 2605, Philadelphia, PA, Oct. 1995, pp. 115–126.

[8] B. E. Moision, A. Orlitsky, and P. H. Siegel, "On codes that avoid specified differences," *IEEE Trans. Inform. Theory*, to be published.

[9] A. J. van Wijngaarden and E. Soljanin, "A combinatorial technique for constructing high rate MTR–RLL codes," *IEEE J. Selec. Areas Commun.*, to be published.

[10] E. Soljanin, "On-track and off-track distance properties of Class 4 partial response channels," in *Proc. 1995 SPIE Int. Symp. on Voice, Video, and Data Communications*, vol. 2605, Philadelphia, PA, Oct. 1995, pp. 92–102.

[11] B. E. Moision, P. H. Siegel, and E. Soljanin, "Distance enhancing codes for high-density magnetic recording channel,", preprint.