# Multihead Detection for Multitrack Recording Channels

Emina Soljanin, *Member, IEEE*, and Costas N. Georghiades, *Fellow, IEEE*

***Abstract*— We look at multiple-track detection for magnetic recording systems that use array heads to write and read over multiple tracks simultaneously. The recording channel is modeled as having intersymbol interference (ISI) in the axial direction, and intertrack interference (ITI) in the radial direction. Optimum multihead and single-head detectors are derived and analyzed in terms of error-probability performance for various levels of intertrack interference. Among other results, it is seen that for a range of ITI levels, codes designed to increase distance in single-head systems can provide the same coding gains for multihead systems.**

***Index Terms*—Coding, intersymbol interference, intertrack interference, multihead detection, multitrack recording,**

## I. Introduction

Multiple-head recording systems which write and read over a number of tracks simultaneously have recently received attention in the context of high-density (narrow trackwidth) systems (see [1]–[15] and references therein). There are two main motivations for using multiple-head systems, both directly related to the amount of information that can be stored on the medium. First, by writing and reading on multiple tracks simultaneously, the required redundancy for timing and gain control can be reduced, thus increasing information density [1]–[3]. Second, as shown in [4]-[9] for a special class of channels, multiple-head systems can better combat intertrack interference (ITI), which in high-density, narrow trackwidth systems, severely degrades the error-rate performance of single-head detectors. In addition, multiple-head systems are shown to be more robust to head misalignment errors [15]. On the negative side, there is, of course, the problem of practically constructing array heads, and the increased complexity of the algorithms (as we will see next) for processing data from multiple heads simultaneously. Given the potential gains in recording density and the significant improvements in timing and head-misalignment performance, multiple-head systems will almost certainly play a role in future disc recording systems.

In this correspondence we look at multiple-head detection in the presence of ITI and analyze the performance of both the optimum (maximum-likelihood) and suboptimum detectors. In the development that follows, we use a rather simplified model for ITI in which the pulse response from each track to each head is the same and only its amplitude varies with the track-to-head distance. We use a general model for intersymbol interference (ISI). Of particular interest (i.e., most common in practice) are systems where only adjacent tracks interfere. A special case of such systems with two-head detectors and a dicode model for ISI was considered by Barbosa [7], Soljanin and Georghiades [10]-[12], Siala and Kaleh [13], and Guzeoglu and Proakis [14]. We here extend their results to multiple (more than two) head receivers and for a general model for ISI.

Areal density in disc recording systems can be increased by narrowing the trackwidth. As a result of the track narrowing, however, there is a loss in signal-to-noise ratio (SNR) during readback, as well as increased ITI. These two problems can be treated simultaneously, as shown in [10], by special two-dimensional codes for the two-track, two-head, dicode channel, capable of compensating for performance degradation due to both the SNR loss and the ITI for large interference levels. Here the idea is to study systems in which the ITI problem is alleviated by employing head-arrays and the SNR is recovered by using existing one-dimensional codes.

In Section II, we introduce a simplified model for the multiple-head recording channel, and present the optimal multitrack detector and a lower bound to its error probability performance. In Section III, we present several examples of single-head and multihead detectors in order to illustrate the benefits of using the latter. In Section IV, we derive analytical expressions for performance for a general model of ISI, and discuss the possibility of using single-track codes for improving noise immunity in multihead systems. We are in particular concerned with the examples discussed in Section III, and the coding results we obtain for these examples are stronger than the results derived for the more general case. In Section V, we illustrate advantages of using detection and coding techniques presented here and in [10] for recording systems with high ITI. We give examples of potential areal density increase by track narrowing and conclude.

## II. Recording Systems with Intertrack Interference

### A. The Multihead Channel

Let

$$\overline{a}_n = (a_{n1}, a_{n2}, \cdots, a_{nN_t}) \in \{-1, 1\}^{N_t}$$

be an $N_t$-dimensional (radial) vector corresponding to $N_t$ data

Manuscript received July 14, 1995; revised February 17, 1998. The material in this correspondence was presented in part the GLOBECOM'93, Houston, TX, November 1993.

E. Soljanin is with Bell Laboratories, Innovations for Lucent Technologies, Murray Hill, NJ 07974 USA.

C. N. Georghiades is with the Electrical Engineering Department, Texas A&M University, College Station, TX 77843-3128 USA.

Publisher Item Identifier S 0018-9448(98)06744-3.

0018–9448/98$10.00 © 1998 IEEE

bits written on the medium in as many adjacent tracks at some time $n$. Further, let $\overline{a}^l = \{a_{nl}\}_{n=-\infty}^{\infty}$ be the sequence of symbols recorded on the $l$th track as a continuous-time signal

$$s_l(t) = I \sum_{n=-\infty}^{\infty} a_{nl} b_T(t - nT), \qquad 1 \le l \le N_t$$

where $I$ is some initial write-current and $b_T(t)$ is a unit-height rectangular pulse occupying the time interval $[0, T)$. The stored data are read by a rigid array of $N_h$ heads flying over these tracks. The $k$th, $1 \le k \le N_h$, head responds to the magnetization in the $l$th, $1 \le l \le N_t$, track by producing a signal as if it were positioned over that track but with an amplitude modified by a weighting parameter $\alpha_{kl}$. This is a simplifying assumption, made to make the problem mathematically tractable, since in practice the transition response from a track to a head gets wider as the distance between them increases, as discussed by Vea and Moura [16] and Lindholm [17]. Nevertheless, we believe the model does capture the essence of the channel, and we will assume it herein.

We will refer to the $\{\alpha_{kl}\}$ as the intertrack-interference (ITI) parameters, and define

$$\mathcal{A} = \{\overline{\alpha}_k = (\alpha_{k1}, \alpha_{k2}, \cdots, \alpha_{kN_t})\}_{k=1}^{N_h}$$

to be the $N_h \times N_t$ matrix whose entries are the ITI parameters. In saturation recording, the total read-head response is the superposition of responses to individual flux reversals, which makes the reading process essentially linear. Thus the signal read by the $k$th head is given by

$$r_k(t) = \sqrt{E} \sum_{l=1}^{N_t} \alpha_{kl} \left( \sum_{n=-\infty}^{\infty} a_{nl} h(t - nT) \right) + n_k(t) \quad (1)$$

where $h(t)$ is a unit-energy pulse, $n_k(t)$ are independent, zero-mean, white Gaussian random processes of variance $\sigma^2$, and $E$ is a constant related to the output voltage amplitude. We will refer to $E/\sigma^2$ as the signal-to-noise ratio (SNR) per track.

The above set of equations can be compactly written as [18]

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{a} + \boldsymbol{n}$$

where

$$\boldsymbol{a} = \{\overline{a}_n\}_{n=-\infty}^{\infty}$$
$$\boldsymbol{r} = \{r_k(t)\}_{k=1}^{N_h}$$
$$\boldsymbol{n} = \{n_k(t)\}_{k=1}^{N_h}$$

and $\boldsymbol{H}\colon \mathcal{C} \to L_2^{N_h}(-\infty, \infty)$ denotes a mapping defined by

$$\boldsymbol{H}\boldsymbol{a} = \left\{ (Ha)_k(t) = \sqrt{E} \sum_{l=1}^{N_t} \alpha_{kl} \sum_{n=-\infty}^{\infty} a_{nl} h(t - nT) \right\}_{k=1}^{N_h} \quad (2)$$

from the set of all possible recorded sequences

$$\mathcal{C} \subseteq \{\boldsymbol{a} = \{\overline{a}_n \in \{-1, 1\}^{N_t}\}_{n=-\infty}^{\infty}\} \quad (3)$$

to the Hilbert space $L_2^{N_h}(-\infty, \infty)$. The multitrack detector observes the $N_h$ signals and uses them to produce a maximum-likelihood estimate of the recorded data.

### B. Optimal Multitrack Detection

Given the observed data $\boldsymbol{r}$, the optimum detector performs maximum-likelihood sequence estimation (MLSE), i.e., it chooses an $\hat{\boldsymbol{a}}$ satisfying

$$\hat{\boldsymbol{a}} = \arg \max_{\boldsymbol{a} \in \mathcal{C}} p(\boldsymbol{r}|\boldsymbol{a}).$$

Given the Gaussian and independent nature of the data, the optimum detector implements

$$\hat{\boldsymbol{a}} = \arg \min_{\boldsymbol{a} \in \mathcal{C}} \|\boldsymbol{r} - \boldsymbol{H}\boldsymbol{a}\|^2 = \arg \max_{\boldsymbol{a} \in \mathcal{C}} \Omega(\boldsymbol{a})$$

where

$$\Omega(\boldsymbol{a}) \triangleq \frac{1}{\sqrt{E}} \left( \langle \boldsymbol{r}, \boldsymbol{H}\boldsymbol{a} \rangle - \frac{1}{2} \langle \boldsymbol{H}\boldsymbol{a}, \boldsymbol{H}\boldsymbol{a} \rangle \right). \quad (4)$$

It is straightforward to show that the optimum multihead detector chooses an $\hat{\boldsymbol{a}}$ that minimizes the following log-likelihood function [8]:

$$\Omega(\boldsymbol{a}) = \sum_{n=-\infty}^{\infty} \overline{a}_n \mathcal{A}' \overline{y}'_n - \frac{\sqrt{E}}{2} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \overline{a}_n \mathcal{A}' \mathcal{A} \overline{a}'_m x_{n-m} \quad (5)$$

where

$$\overline{y}_n = (y_{n1}, y_{n2}, \cdots, y_{nN_h})$$
$$y_{nk} = \int_{-\infty}^{\infty} h(t - nT) r_k(t)\, dt, \qquad 1 \le k \le N_h \quad (6)$$

and

$$x_n = \int_{-\infty}^{\infty} h(t) h(t - nT)\, dt.$$

For the majority of magnetic recording systems, the channel response is equalized so that

$$\{x_n\} \overset{\text{DFT}}{\longleftrightarrow} X(D) = A^2 F(D) F^*(D^{-1})$$
$$F(D) = (1 - D)(1 + D)^N$$

where $A^2$ is a normalizing constant and $N$ is a nonnegative integer.

For the purpose of error-probability analysis and coding, we derive an equivalent discrete-time multihead, multitrack channel model [8]. This channel model describes the relation between the discrete input $\overline{a}_n$ and the above defined discrete output $\overline{y}_n$ as follows:

$$\overline{y}_n = \sqrt{E} \sum_{m=-\infty}^{\infty} \overline{a}_m \mathcal{A}' x_{n-m} + \overline{\nu}_n \quad (7)$$

where $\overline{\nu}_n = (\nu_{n1}, \nu_{n2}, \cdots, \nu_{nN_h})$ are vectors containing the noise samples for the $N_h$ heads at time $n$. The crosscorrelation between $\nu_{nk}$ and $\nu_{mj}$, $-\infty \le n, m \le \infty$, $1 \le k, j \le N_h$, is given by $\sigma^2 x_{n-m} \delta_{kj}$.

### C. Error Probability Analysis for the Optimal Multitrack Detector

Let $\boldsymbol{a}=\{\overline{a}_n\}_{n=-\infty}^{\infty}$ and $\boldsymbol{b}=\{\overline{b}_n\}_{n=-\infty}^{\infty}$ be two allowable recorded sequences, i.e., $\boldsymbol{a},\boldsymbol{b}\in\mathcal{C}$, where $\mathcal{C}$ is the set of all possible recorded sequences as defined by (3). We refer to $\boldsymbol{\epsilon}=(\boldsymbol{a}-\boldsymbol{b})/2$ as the normalized error sequence corresponding to $\boldsymbol{a}$ and $\boldsymbol{b}$. The set of all admissible error sequences, i.e., the set of all error sequences that correspond to the difference between two allowable recorded sequences is given by

$$\mathcal{E}=\left\{\boldsymbol{\epsilon}=\{\overline{e}_n\in\{-1,0,1\}^{N_t}\}=(\boldsymbol{a}-\boldsymbol{b})/2,\boldsymbol{a},\boldsymbol{b}\in\mathcal{C}\right\}. \quad (8)$$

A normalized error sequence $\boldsymbol{\epsilon}\in\mathcal{E}$ is by definition (8) a sequence of normalized error vectors $\overline{e}_n\in\{-1,0,1\}^{N_t}$ containing bit errors of the $N_t$ tracks at time $n$. It can also be represented as a vector of normalized error sequences $\overline{e}_l\in\mathcal{E}_l$ containing bit errors of the $l$th track for all $n$, i.e., $\boldsymbol{\epsilon}=(\overline{e}_1,\cdots,\overline{e}_{N_t})$, where

$$\overline{e}_l\in\mathcal{E}_l\subseteq\{-1,0,1\}^{\infty},\qquad 1\le l\le N_t.$$

Suppose that the optimum detector is told by a genie that the true sequence is either $\boldsymbol{a}$ or $(\boldsymbol{a}-2\boldsymbol{\epsilon})$, for some $\boldsymbol{a},\boldsymbol{a}-2\boldsymbol{\epsilon}\in\mathcal{C}$. In this case, the detection reduces to a binary hypothesis testing $\Omega(\boldsymbol{a})\gtrless\Omega(\boldsymbol{a}-2\boldsymbol{\epsilon})$. The conditional probability of error given that $\boldsymbol{a}$ was recorded, denoted by $P[\Omega(\boldsymbol{a}-2\boldsymbol{\epsilon})>\Omega(\boldsymbol{a})|\boldsymbol{a}]$, is easily derived, using (5) and (7), to be

$$P[\Omega(\boldsymbol{a}-2\boldsymbol{\epsilon})>\Omega(\boldsymbol{a})|\boldsymbol{a}]=Q(d(\boldsymbol{\epsilon})\cdot\sqrt{\mathrm{SNR}})$$

where $Q(\cdot)$ is the error-function and $d^2(\boldsymbol{\epsilon})$ is the squared distance between sequences $\boldsymbol{a}$ and $\boldsymbol{a}-2\boldsymbol{\epsilon}$, given by

$$d^2(\boldsymbol{\epsilon})=\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\overline{\epsilon}_n\mathcal{A}'\mathcal{A}\overline{\epsilon}'_m x_{n-m}. \quad (9)$$

For large SNR's, the probability of an error event in the system is well approximated by $Q(d_{\min}\sqrt{\mathrm{SNR}})$, where

$$d_{\min}=\min_{\boldsymbol{\epsilon}\in\mathcal{E},\boldsymbol{\epsilon}\neq\boldsymbol{0}} d(\boldsymbol{\epsilon}). \quad (10)$$

### D. Single-Track Detection

Let there be $N_t$ interfering tracks, but only one head detecting the sequence $\overline{a}^l=\{a_{nl}\}_{n=-\infty}^{\infty}$ recorded on the $l$th track. The signal read by a single head positioned over the $l$th track is given by (1) as

$$r_l(t)=\sqrt{E}\sum_{p=1}^{N_t}\alpha_{lp}\left(\sum_{n=-\infty}^{\infty}a_{np}h(t-nT)\right)+n_l(t). \quad (11)$$

When the detector knows the ITI parameters, it can perform optimal single-head, multitrack detection, which is a special case of (5) described above for $N_h=1$. The performance of this optimal single-head detector is determined by $d_{\min}$ as given in (10) for $N_h=1$.

A suboptimal single-track detector is one that ignores (or is not aware of) the intertrack interference, assuming the received signal to be

$$r_l(t)=\sqrt{E}\sum_{n=-\infty}^{\infty}a_{nl}h(t-nT)+n_l(t). \quad (12)$$

The log-likelihood function for this detector is the well-known log-likelihood function for ISI channels [19] given by

$$\Omega(\overline{a}^l)=\sum_{n=-\infty}^{\infty}a_{nl}y_{nl}-\frac{\sqrt{E}}{2}\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}a_{nl}a_{ml}x_{n-m}. \quad (13)$$

In the above equation $y_{nl}$ are as given by (6) for $k=l$, and $E[\nu_{nl}\nu_{ml}]=x_{n-m}\sigma^2$. Following similar steps as for the multiple-head case, the performance of the suboptimal single-head detector can be approximated for large SNR by $Q(\delta_{\min}\cdot\sqrt{\mathrm{SNR}})$, where

$$\delta_{\min}=\min_{\overline{e}_l\in\mathcal{E}_l,\boldsymbol{a}\in\mathcal{C}}\delta(\overline{e}_l,\boldsymbol{a})$$

and

$$\delta(\overline{e}_l,\boldsymbol{a})=\frac{\displaystyle\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\epsilon_{nl}(\overline{a}_m\overline{\alpha}'_l-a_{ml}+\epsilon_{ml})x_{n-m}}{\left(\displaystyle\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\epsilon_{nl}\epsilon_{ml}x_{n-m}\right)^{1/2}}. \quad (14)$$

We analyze next the performance of single-head and multihead detectors. To illustrate the advantages of using the latter, we consider below several examples of some systems likely to be encountered in practice. In the following section we show how the results obtained for these specific examples can be generalized.

## III. Some Special Cases of Multihead, Multitrack Recording Channels

We consider disc recording systems in which only adjacent tracks interfere. The reading heads are centered over their information tracks and no head straddles two tracks. However, when a reading head is positioned over one of the tracks, it responds to the magnetization of an adjacent track as if it were positioned over that track but with an amplitude modified by a weighting parameter $\alpha$. Thus the signal read by the head positioned over the $l$th track is given by

$$\begin{aligned}r_l(t)=\sqrt{E}&\sum_{n=-\infty}^{\infty}a_{nl}h(t-nT)\\&+\alpha\cdot\sqrt{E}\left(\sum_{n=-\infty}^{\infty}a_{n(l-1)}h(t-nT)\right.\\&\qquad\left.+\sum_{n=-\infty}^{\infty}a_{n(l+1)}h(t-nT)\right)+n_l(t).\end{aligned}$$

For the examples that follow, we assume that each track can be modeled as a dicode channel, i.e.,

$$\{x_k\}\overset{\mathrm{DFT}}{\longleftrightarrow}X(D)=\frac{1}{2}F(D)F^*(D^{-1}),\qquad F(D)=(1-D).$$

This assumption is removed in the next section. We analyze five different detection systems and compare their performance on the basis of the minimum-distance parameters. We first consider the optimal single-head detector, i.e., the single-head detector that knows the interference and accounts for

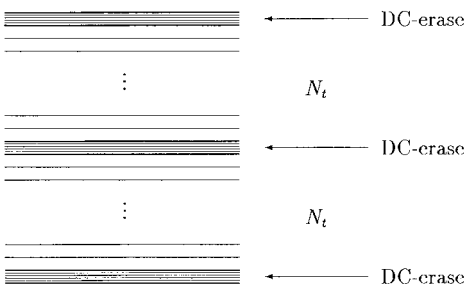



Fig. 1. Disk configuration.

it. We then consider multihead detectors for a particular disc configuration where the tracks are written and read in groups of $N_t$ tracks at a time. Adjacent groups of $N_t$ tracks are separated by a track with constant magnetization in one direction known to the detector, as shown in Fig. 1.

### A. The Optimal Single-Head Detector

The optimal single-head (multitrack) detector that knows ITI is a special case of the optimum multitrack, multihead detector, whose performance is determined to a large extent by the distance metric in (9). The parameter $d^2_{\min}$ is obtained by minimization of

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) = &\sum_n (\epsilon_{nl} - \epsilon_{(n-1)l})^2 \\ &+ \alpha^2 \sum_n \left[(\epsilon_{n(l-1)} - \epsilon_{(n-1)(l-1)})^2 \right. \\ &\qquad \left. + (\epsilon_{n(l+1)} - \epsilon_{(n-1)(l+1)})^2\right] \\ &+ 2\alpha \sum_n (\epsilon_{nl} - \epsilon_{(n-1)l}) \\ &\cdot \left[(\epsilon_{n(l-1)} - \epsilon_{(n-1)(l-1)}) + (\epsilon_{n(l+1)} - \epsilon_{(n-1)(l+1)})\right] \\ &+ 2\alpha^2 \sum_n (\epsilon_{n(l-1)} - \epsilon_{(n-1)(l-1)}) \\ &\cdot (\epsilon_{n(l+1)} - \epsilon_{(n-1)(l+1)}) \end{aligned}$$

over all admissible error sequences $\boldsymbol{\epsilon} = (\overline{e}_{l-1}, \overline{e}_l, \overline{e}_{l+1})$ $\mathbf{0}$.

### B. The Suboptimal Single-Head Detector

The single-head detector, as previously stated, ignores the ITI, and the performance parameter $\delta_{\min}$ is obtained by minimization of (obtained as a special case of (14))

$$\begin{aligned} \delta(\overline{e}_l, \overline{a}_{l-1}, \overline{a}_{l+1}) = &\frac{1}{\sqrt{2}\left[\sum_n [\epsilon_{nl} - \epsilon_{(n-1)l}]^2\right]^{1/2}} \\ &\cdot \left[\sum_n [\epsilon_{nl} - \epsilon_{(n-1)l}]^2 + \alpha \sum_n [\epsilon_{nl} - \epsilon_{(n-1)l}] \right. \\ &\left. \cdot \left[a_{n(l-1)} - a_{(n-1)(l-1)} + a_{n(l+1)} - a_{(n-1)(l+1)}\right]\right] \end{aligned}$$

over all admissible error sequences $\overline{e}_l \neq \overline{0}$ and all recorded sequences $\overline{a}_{(l-1)}$ and $\overline{a}_{(l+1)}$.

### C. A Two-Head, Two-Track Detector

Tracks are written and read in groups of $N_t = 2$ tracks at a time. There are two reading heads $(N_h = 2)$ flying over these tracks. In this case

$$\mathcal{A} = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix} \tag{15}$$

and the distance parameter $d^2_{\min}$ is obtained by minimization of

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) = &(1+\alpha^2) \sum_n \left[(\epsilon_{n1} - \epsilon_{(n-1)1})^2 + (\epsilon_{n2} - \epsilon_{(n-1)2})^2\right] \\ &+ 4\alpha \sum_n (\epsilon_{n1} - \epsilon_{(n-1)1})(\epsilon_{n2} - \epsilon_{(n-1)2}) \end{aligned}$$

over all admissible error sequences $\boldsymbol{\epsilon} \neq \mathbf{0}$.

### D. Three-Head, Three-Track Detector

In this case, tracks are written and read in groups of $N_t = 3$ tracks at a time, and there are three reading heads flying over these tracks $(N_h = 3)$. Thus

$$\mathcal{A} = \begin{bmatrix} 1 & \alpha & 0 \\ \alpha & 1 & \alpha \\ 0 & \alpha & 1 \end{bmatrix}$$

and parameter $d^2_{\min}$ is obtained by minimization of

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) = &(1+\alpha^2) \sum_n \left[(\epsilon_{n1} - \epsilon_{(n-1)1})^2 + (\epsilon_{n3} - \epsilon_{(n-1)3})^2\right] \\ &+ (1+2\alpha^2) \sum_n (\epsilon_{n2} - \epsilon_{(n-1)2})^2 \\ &+ 4\alpha \sum_n \left[(\epsilon_{n1} - \epsilon_{(n-1)1})(\epsilon_{n2} - \epsilon_{(n-1)2}) \right. \\ &\qquad \left. + (\epsilon_{n2} - \epsilon_{(n-1)2})(\epsilon_{n3} - \epsilon_{(n-1)3})\right] \\ &+ 2\alpha^2 \sum_n (\epsilon_{n1} - \epsilon_{(n-1)1})(\epsilon_{n3} - \epsilon_{(n-1)3}) \end{aligned}$$

over all admissible error sequences $\boldsymbol{\epsilon} \neq \mathbf{0}$.

### E. Five-Head, Three-Track Detector

Here $N_t = 3, N_h = 5$, and

$$\mathcal{A} = \begin{bmatrix} \alpha & 0 & 0 \\ 1 & \alpha & 0 \\ \alpha & 1 & \alpha \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{bmatrix}. \tag{16}$$

The parameter $d_{\min}$ is obtained by minimization of

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) = &(1+2\alpha^2) \sum_n \left[(\epsilon_{n1} - \epsilon_{(n-1)1})^2 \right. \\ &\qquad \left. + (\epsilon_{n2} - \epsilon_{(n-1)2})^2 + (\epsilon_{n3} - \epsilon_{(n-1)3})^2\right] \\ &+ 4\alpha \sum_n \left[(\epsilon_{n1} - \epsilon_{(n-1)1})(\epsilon_{n2} - \epsilon_{(n-1)2}) \right. \\ &\qquad \left. + (\epsilon_{n2} - \epsilon_{(n-1)2})(\epsilon_{n3} - \epsilon_{(n-1)3})\right] \\ &+ 2\alpha^2 \sum_n (\epsilon_{n1} - \epsilon_{(n-1)1})(\epsilon_{n3} - \epsilon_{(n-1)3}) \end{aligned}$$

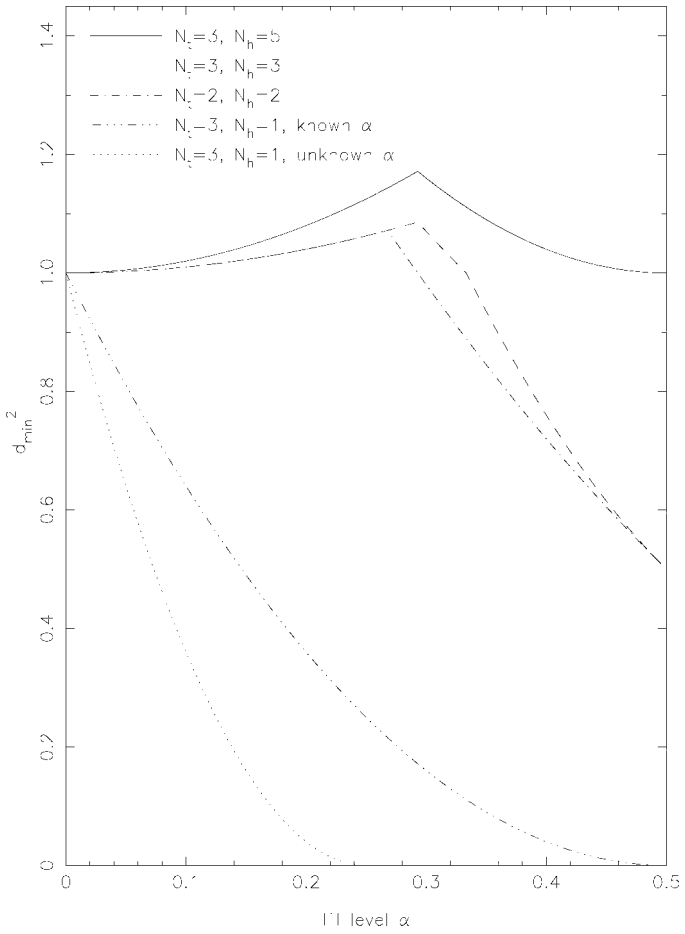



Fig. 2. Performance of five different detection systems for a channel of three interfering tracks.

over all admissible error sequences $\boldsymbol{\epsilon} \neq \mathbf{0}$. Existence of the heads above the bounding tracks makes all tracks equally susceptible to errors.

The performance of the five detection systems described above are compared on the basis of their minimum distance parameters (labeled as $d^2_{\min}$ in the figure). Fig. 2 plots $d^2_{\min}$ for each of the five cases as a function of the interference parameter $\alpha$. From the figure we see the strong dependence of the performance of the single-head, single-track receiver on the ITI level. While the performance of the optimal two- and three-head receivers increases slightly over a range of ITI levels, the performance of the single-head detector degrades rapidly with increasing ITI level. Intuitively, one would expect that ITI, as a noise of specific kind, causes performance degradation. However, multihead systems also benefit from ITI because each head contributes $\alpha^2$ in energy to the joint detector related to an adjacent track. An insight into this phenomenon can be gained by considering the analytical expressions for $d^2_{\min}$ derived in the following section.

Performance loss due to ITI can, as we saw in the examples, be partly recovered by the use of multihead systems. Performance loss due to SNR decrease in single-track single-head systems can be recovered by the use of modulation codes. We, therefore, investigate next the possibility of using single-track codes in multitrack, multihead systems.

## IV. Distance Properties and Single-Track Coding in Multitrack, Multihead Systems

In the previous section we saw that multitrack, multihead detectors for a specific class of channels, can effectively combat the ITI incurred, for example, by track narrowing. In this section, we analytically show that this is also the case for a more general class of channels. The performance of multihead detectors should be further enhanced by appropriate coding to improve noise immunity, thus compensating for a possible SNR loss. There exist modulation codes designed to improve noise immunity in single-track recording systems, but with the exception of work reported in [9]–[14], their multidimensional counterparts are still to be investigated. In this section, we investigate the possibility of using existing codes designed for single-track systems in multitrack, multihead systems. We first look at some of the special cases given as examples in the previous section, but with a general model of ISI, and then we consider the general case.

To examine the error-probability performance of multitrack, multihead detection systems as well as the applicability of existing single-track codes in these systems, we consider their minimum distance properties. The multitrack minimum-distance parameter $d^2_{\min}$ is obtained by minimization over certain sets of $d^2(\boldsymbol{\epsilon})$ given in (9). We shall use another form of this expression which we obtain next. As mentioned above, a normalized error sequence $\boldsymbol{\epsilon} \in \mathcal{E}$ can be represented as a vector of normalized error sequences $\overline{e}_l \in \mathcal{E}_l$ containing bit-errors of the $l$th track, $1 \leq l \leq N_t$, i.e., $\boldsymbol{\epsilon} = (\overline{e}_1, \cdots, \overline{e}_{N_t})$, where

$$\overline{e}_l \in \mathcal{E}_l \subseteq \{-1, 0, 1\}^{\infty}, \qquad 1 \leq l \leq N_t.$$

Thus the above expression (9) for $d^2(\boldsymbol{\epsilon})$ can be rewritten as

$$d^2(\boldsymbol{\epsilon}) = \sum_{l=1}^{N_t} \sum_{i=1}^{N_t} \overline{e}_l X \overline{e}_i' \gamma_{li} \tag{17}$$

where $X = \{\chi_{nm}\}_{(2M+1)\times(2M+1)}$ is a matrix with elements $\chi_{nm} = x_{n-m}, -\infty \leq n, m \leq \infty$, and $\gamma_{li}$ are the elements of matrix $\mathcal{A}'\mathcal{A}$, i.e., $\mathcal{A}'\mathcal{A} = \{\gamma_{lj}\}_{N_t \times N_t}$. Note that

$$d^2_{0l} = \min_{\overline{e} \in \mathcal{E}_l, \overline{e} \neq \overline{0}} \overline{e}_l X \overline{e}_l' \tag{18}$$

is the well-known single-track minimum-distance parameter, corresponding to the $l$th track, $1 \leq l \leq N_t$. Since $X$ is positive-definite, we can define an inner product in $R^{\infty}$ as $\langle \overline{u}, \overline{v} \rangle_X = \overline{u} X \overline{v}', \overline{u}, \overline{v} \in R^{\infty}$. The norm of $\overline{u} \in R^{\infty}$ is then $||\overline{u}||^2_X = \langle \overline{u}, \overline{u} \rangle_X$, and the Cauchy–Schwarz inequality implies $|\langle \overline{u}, \overline{v} \rangle_X|^2 \leq ||\overline{u}||^2_X ||\overline{v}||^2_X$.

### A. Two-Track, Two-Head Systems

Systems with two interfering tracks simultaneously read by two heads and a general model of ISI were studied in [10]-[12]. For continuity, we briefly present without proof some of the results in this previous work here.

*Theorem IV.1 ([10]):* Let $d_0$ be a single-track minimum-distance parameter defined by (18) as

$$d_0^2 = \min_{\overline{e} \in \mathcal{E}_t, \overline{e} \neq \overline{0}} ||\overline{e}||_X^2$$

where for the two-head, two-track case under consideration $\mathcal{E}_t = \mathcal{E}_1 = \mathcal{E}_2$. Then

$$d_{\min}^2 = \begin{cases} (1+\alpha^2)d_0^2, & \text{if } 0 \le \alpha \le 2-\sqrt{3} \\ 2(1-\alpha)^2 d_0^2, & \text{if } 2-\sqrt{3} \le \alpha \le 1/2. \end{cases} \quad (19)$$

Note that, because of multiple heads, there is no performance loss due to ITI as long as $d_{\min}^2 \ge d_0^2$, i.e., $0 \le \alpha \le \sqrt{2}/2$.

*Corollary IV.2 ([10]):* A single-track code that provides an increase in the single-track minimum distance to $d_0^c = \sqrt{g} d_0$ when applied to each track, results in an increase in the two-track minimum distance to $d_{\min}^c = \sqrt{g} d_{\min}$.

Thus if, for example, the rate $1/2$ biphase code, which provides a coding gain of 4.8 dB on $(1-D)$ channel, is applied independently on each track, it will result in the same 4.8 dB coding gain for the two-track, two-head detector. In this case

$$d_{\min}^2 = \begin{cases} 3(1+\alpha^2), & \text{if } 0 \le \alpha \le 2-\sqrt{3} \\ 6(1-\alpha)^2, & \text{if } 2-\sqrt{3} \le \alpha \le 1/2. \end{cases}$$

Computer-simulated results for the biphase code applied to a two-track, two-head system for three different levels of ITI are shown in Fig. 3. The asymptotic performance gains are as predicted by Corollary IV.2. If the performance of an uncoded system with $\alpha = 0$ is used as the baseline, then the performance gain for $\alpha = 0$ is about 4.8 dB. Note that for $\alpha = 0.5$ this gain is offset by the loss of 3 dB due to ITI.

The above corollary shows that codes designed for improving noise immunity in single-track systems can provide a coding gain in two-track, two-head systems. However, these codes are not capable of recovering the loss in performance due to ITI, since they are not designed to account for it. A class of two-dimensional codes capable of recovering the SNR loss incurred by ITI for large interference levels was introduced in [10]-[12] for the dicode channel.

### B. A Three-Track, Five-Head System

We consider systems with three interfering tracks simultaneously read by five heads as in the preceding section. In this case, the ITI matrix is given by (16). We examine the distance properties of these systems through the following theorem

*Theorem IV.3:* Let $d_0$ be as previously defined. Let

$$\{x_k\} \overset{\text{DFT}}{\longleftrightarrow} X(D) = A^2 F(D) F^*(D^{-1})$$

where $F(D)$ is a polynomial with integer coefficients and $A$ is a normalizing constant. Then

$$d_{\min}^2 = \begin{cases} (1+2\alpha^2)d_0^2, & \text{if } 0 \le \alpha \le 1-\sqrt{2}/2 \\ 2(1+2\alpha^2-2\alpha)d_0^2, & \text{if } 1-\sqrt{2}/2 \le \alpha \le 1/2 \end{cases} \quad (20)$$

provided $d_0^2 \le 6A^2$.

*Proof:* See Appendix I. □

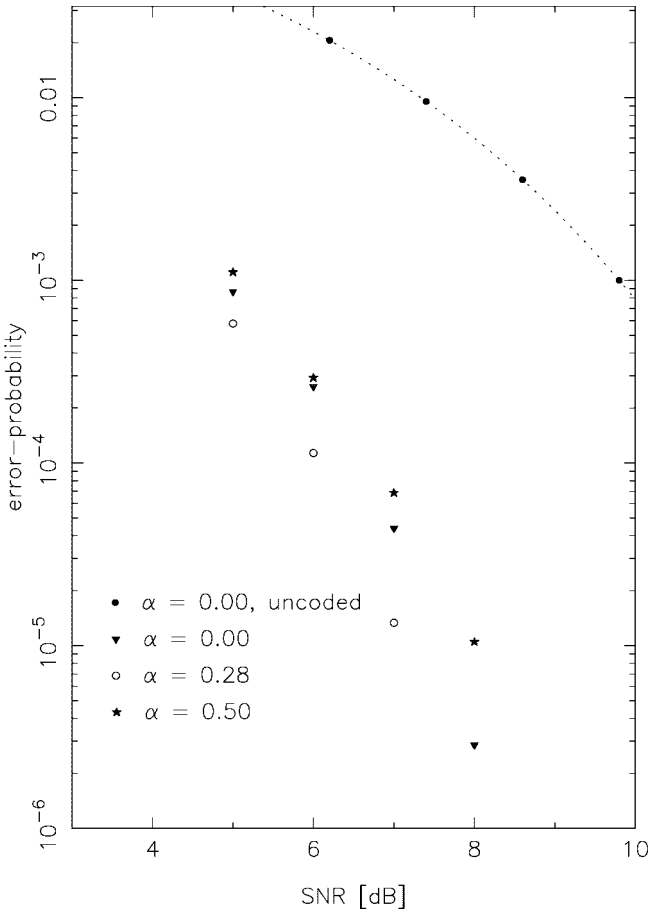



Fig. 3. Performance of a biphase coded two-track, two-head system for three different interference levels.

*Remarks:*

a) Recall that $F(D)$ for magnetic recording channels is usually modeled as $F(D) = (1-D)(1+D)^N$, and, therefore, is a polynomial with integer coefficients.

b) For the normalized $(1-D)$ channel, $A^2 = 1/2$, and, therefore, $d_0^2 \le 6A^2$ for $d_0^2 \le 3$, which is satisfied for uncoded systems ($d_0^2 = 1$), as well as for systems employing the biphase code ($d_0^2 = 3$) or higher rate MSN codes ($d_0^2 = 2$).

c) The assumptions of the theorem also hold for $F(D)$ given as above with $N = 1, 2, 3$, i.e., channel models most commonly considered in magnetic recording practice.

Note that there is no performance loss due to ITI as long as $d_{\min}^2 \ge d_0^2$, i.e., $0 \le \alpha \le 1/2$, which is the entire interval under consideration. As previously mentioned, coding in narrow-track recording systems should recover the loss in performance due to both ITI and decrease in SNR. However, in the three-track, five-head systems we considered, the loss in performance due to ITI is completely recovered by the multihead receiver and coding should recover the performance loss due to SNR decrease only.

*Corollary IV.4:* Under the assumptions of the preceding theorem, a single-track code that provides an increase in the single-track minimum distance to $d_0^c = \sqrt{g} d_0$ when applied to each track, results in an increase in the three-track, five-head minimum distance to $d_{\min}^c = \sqrt{g} d_{\min}$, as long as $d_0^c \le \sqrt{6}A$.

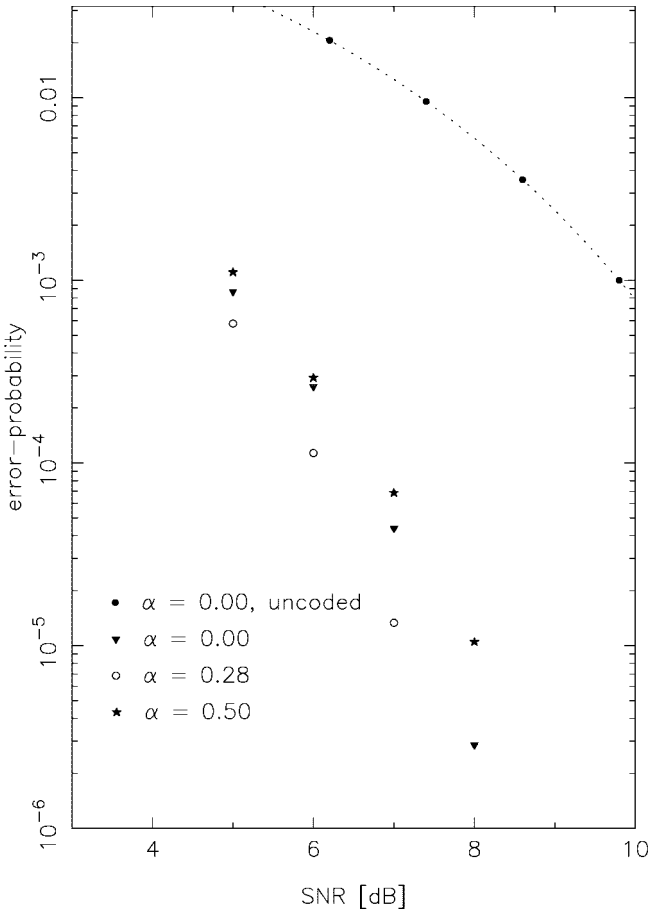



Fig. 4. Performance of a biphase-coded three-track, five-head system for three different interference levels.

As remarked above, the biphase coded $(1 - D)$ channel satisfies the assumptions of the corollary. Computer simulation results for this channel with three-track, five-head detection and for three different levels of ITI are shown in Fig. 4. The asymptotic performance gain is as predicted by Corollary IV.4. If the uncoded system with $\alpha = 0$ is used as the baseline, then the performance gain for $\alpha = 0$ is about 4.8 dB, as in Fig. 3, which shows the performance of the system with two-track, two-head detection. Note that now for $\alpha = 0.5$ this gain remains almost the same, as a result of five-head systems being better able to combat ITI.

### C. Multitrack, Multihead Systems

For the general class of multitrack recording channels defined in Section II, we prove that certain conditions imposed on the ITI parameters make existing, single-track codes applicable in the multitrack, multihead case.

*Theorem IV.5:* Let $d_0^2$ be as previously defined. Let

$$\mathcal{A}'\mathcal{A} = \{\gamma_{lj}\}_{N_t \times N_t}$$

be such that $\gamma_{li} = \gamma_{il} = s_{|l-i|}$ for $l \neq j, s_0 = \min_{1 \leq l \leq N_t} \gamma_{ll}$, and $s_l$ satisfy

$$s_0 > 2 \sum_{i=1}^{N_t - 1} |s_i|. \tag{21}$$

Then

$$d_{\min}^2 = s_0 d_0^2$$

i.e., the worst case error events are single-track error events.

*Proof:* See Appendix II. □

*Remark:* Note that condition (21) is a sufficient condition, i.e., it guarantees that the worst case error events (those for which $d_{\min}^2$ is achieved) are single-track error events. This condition is (not surprisingly) similar to that derived by Ungerboeck for ISI channels [20], which guarantees that the worst case error events are single-symbol error events.

*Corollary IV.6:* Under the assumptions and conditions of the preceding theorem, a single-track code that provides an increase in the single-track minimum distance to $d_0^c = \sqrt{g} d_0$ when applied to each track, results in an increase in the multitrack minimum distance to $d_{\min}^c = \sqrt{g} d_{\min}$.

Assumptions we made about the ITI in multitrack digital magnetic recording channels allow modeling, as we show later, with a matrix $\mathcal{A}'\mathcal{A}$ of the form described above. However, condition (21) can be made less restrictive.

*Example IV.7:* Consider our two-track, two-head system with ITI described by (16). The matrix $\mathcal{A}'\mathcal{A}$ is therefore given by

$$\mathcal{A}'\mathcal{A} = \begin{bmatrix} 1+\alpha^2 & 2\alpha \\ 2\alpha & 1+\alpha^2 \end{bmatrix}$$

and condition (21) gives $1 + \alpha^2 > 4\alpha$, which translates into $\alpha < 2 - \sqrt{3}$. Therefore, by Theorem IV.5, $\alpha < 2 - \sqrt{3}$ is a sufficient condition that the worst case error events be single-track errors. By Theorem IV.1, this condition is also necessary.

*Example IV.8:* Consider our three-track, five-head system with ITI described by (16). $\mathcal{A}'\mathcal{A}$ is therefore given by

$$\mathcal{A}'\mathcal{A} = \begin{bmatrix} 1+2\alpha^2 & 2\alpha & \alpha^2 \\ 2\alpha & 1+2\alpha^2 & 2\alpha \\ \alpha^2 & 2\alpha & 1+2\alpha^2 \end{bmatrix}$$

and condition (21) gives $1+2\alpha^2 > 4\alpha+2\alpha^2$, which translates into $\alpha < 1/4$. Therefore, by Theorem IV.5, $\alpha < 1/4$ is a sufficient condition that the worst case error events be single-track errors. By Theorem IV.3, a less restrictive condition $\alpha < 1 - \sqrt{2}/2$ is sufficient and necessary.

*Example IV.9:* We now consider a general case for the particular disc configuration described in the preceding section, where the tracks are written and read in groups of $N_t$ tracks at a time. Adjacent groups are separated by a track with constant magnetization in one direction known to the receiver, as in Fig. 1. The receiver consists of $N_h = N_t + 2$ reading heads flying over a group of $N_t$ tracks, and two boundary DC-erase tracks. Thus the loss in density due to the DC-erase tracks is $1/(N_t + 1)$, which can be made small by increasing $N_t$. We assume as above that only adjacent tracks interfere. The heads above the bounding DC-erase tracks makes all tracks equally susceptible to errors, which becomes significant with increase in servo positioning error. As in [4] and [5], we assume that the

TABLE I
POTENTIAL AREAL DENSITY INCREASE $\gamma$

| $W_{\text{old}}/W_{\text{new}}$ | $N_t$ | $N_h$ | coding scheme | $r_{\text{new}}$ | $\alpha_{\max}$ | $\gamma$ |
|---|---|---|---|---|---|---|
| 3 | 2 | 2 | biphase code on each track | 1/2 | 0.32 | 1.6875 |
| 2 | 2 | 2 | MSN code on each track | 4/5 | 0.32 | 1.8000 |
| 2 | 2 | 2 | two-track code | 3/4 | 0.42 | 1.6875 |
| 3 | 3 | 5 | biphase code on each track | 1/2 | 0.50 | 1.6875 |
| 2 | 3 | 5 | MSN code on each track | 4/5 | 0.50 | 1.8000 |

two outer heads are only half as wide as the $N_t$ inner heads. So, when the head array is positioned above a particular group of $N_t$ tracks, the outer heads do not pick up the magnetization of the neighboring groups.

The initial interference matrix for our model with the above assumptions is

$$\mathcal{A} = \begin{bmatrix} \alpha & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & \alpha & 0 & 0 & \cdots & 0 & 0 \\ \alpha & 1 & \alpha & 0 & \cdots & 0 & 0 \\ 0 & \alpha & 1 & \alpha & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \alpha & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \alpha \end{bmatrix}_{(N_t+2)\times N_t}. \tag{22}$$

In [16] the responses of a head to the magnetization of adjacent tracks are computed based on the system parameters. It seems that for an example given there, the above model is a good approximation with $\alpha$ about 0.1. To see for which $\alpha$ single-track errors are dominant, let $\overline{e}_l \in \{-1, 0, 1\}^{\infty}$, $1 \le l \le N_t$. For $\mathcal{A}$ given by (22), we have the matrix at the bottom of this page, and thus

$$d^2(\epsilon) = (1+2\alpha^2) \sum_{l=1}^{N_t} ||\overline{e}_l||_X^2 + 2 \cdot 2\alpha \sum_{l=1}^{N_t-1} \langle \overline{e}_l, \overline{e}_{l+1} \rangle_X + 2 \cdot \alpha^2 \sum_{l=1}^{N_t-2} \langle \overline{e}_l, \overline{e}_{l+2} \rangle_X.$$

The condition that $\alpha$ should satisfy for single-track errors to be dominant in this case is $1+2\alpha^2 > 4\alpha + 2\alpha^2$, which translates into $\alpha < 1/4$.

## V. CONCLUSIONS

A reduction in trackwidth in disc recording systems results in a desirable increase in areal density but also in the undesirable appearance of ITI and loss in SNR. We have found that the performance degradation incurred by the presence of ITI can be alleviated by using multihead receivers that simultaneously read data from multiple tracks. We have also found that the performance degradation incurred by the loss in SNR can be recovered by using single-track codes under some general assumptions. In the related paper [10] several two-dimensional codes were designed for the two-track, two-head, dicode channel, capable of compensating for the performance degradation due to both the SNR loss and the ITI for large interference levels.

We now discuss potential areal density increase achievable by track narrowing and detection and coding techniques presented here and in [10]. Consider a system employing a rate $r_{\text{old}}$ code on a nominal track width $W_{\text{old}}$. Suppose that the track width is decreased to $W_{\text{new}}$, causing an SNR loss of $10 \log_{10}(W_{\text{old}}/W_{\text{new}})$ decibels and no ITI. If a rate $r_{\text{new}}$ code is used on each track to compensate for the incurred SNR loss replacing the old code, the overall areal density increases by a factor of

$$\gamma = \frac{r_{\text{new}}}{r_{\text{old}}} \cdot \frac{W_{\text{old}}}{W_{\text{new}}}.$$

Consider now a benchmark $(1-D)$ PRML system employing a code of rate $r_{\text{old}} = 8/9$ on a nominal trackwidth $W_{\text{old}}$. If the trackwidth is reduced to $W_{\text{new}} = W_{\text{old}}/3$, the incurred loss in SNR can be recovered by applying the rate $1/2$ biphase code on each track replacing the old code. The overall areal density increases by a factor of $\gamma = 1.6875$.

Table I shows that multitrack, multihead, coded systems potentially provide a comparable areal density increase for these systems even in cases of high ITI. Parameter $\alpha_{\max}$ below denotes the maximum ITI level for which there will be no performance loss.

Although most of the results of this research are not confined to the $(1-D)$ channel model, other reasons, like complexity of implementation of multihead receivers, may favor this model. However, single-track magnetic recording channels that are easily equalized to $(1-D)$ channels are those having low linear density, and thus the overall areal density increase should be optimized by choosing the appropriate balance of linear and radial (track) density. The applicability of this research to actual magnetic recording systems is still to be investigated.

$$\mathcal{A}'\mathcal{A} = \begin{bmatrix} 1+\alpha^2 & 2\alpha & \alpha^2 & 0 & \cdots & 0 & 0 \\ 2\alpha & 1+2\alpha^2 & 2\alpha & \alpha^2 & \cdots & 0 & 0 \\ \alpha^2 & 2\alpha & 1+2\alpha^2 & 2\alpha & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1+2\alpha^2 & 2\alpha \\ 0 & 0 & 0 & 0 & \cdots & 2\alpha & 1+2\alpha^2 \end{bmatrix}_{N_t \times N_t}$$

## APPENDIX I
## PROOF OF THEOREM IV.3

The minimum distance $d^2_{\min}$ is obtained by minimizing $d^2(\boldsymbol{\epsilon})$ defined in (17). For three-track, five-head systems $d^2(\boldsymbol{\epsilon})$ is given by

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2 + \|\overline{e}_3\|_X^2) \\ &\quad + 2\cdot 2\alpha(\langle \overline{e}_1, \overline{e}_2\rangle_X + \langle \overline{e}_2, \overline{e}_3\rangle_X) + 2\cdot\alpha^2\langle \overline{e}_1, \overline{e}_3\rangle_X \end{aligned}$$

with $\boldsymbol{\epsilon} = (\overline{e}_1, \overline{e}_2, \overline{e}_3), \overline{e}_1, \overline{e}_2, \overline{e}_3 \in \mathcal{E}_t$. We partition the set of all possible error sequences $\boldsymbol{\epsilon} = (\overline{e}_1, \overline{e}_2, \overline{e}_3)$ into two subsets

1) Error vectors with exactly one nonzero component (single-track errors), i.e., either $\overline{e}_1 = \overline{e}_2 = \overline{0}$ or $\overline{e}_2 = \overline{e}_3 = \overline{0}$ or $\overline{e}_3 = \overline{e}_1 = \overline{0}$, for which

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2 + \|\overline{e}_3\|_X^2) \\ &\geq (1+2\alpha^2)\, d_0^2. \end{aligned}$$

Equality is achieved when $\|\overline{e}_l\|_X^2 = d_0^2$ for some (but exactly one) $\overline{e}_l \neq \overline{0}, l = 1,2,3$.

2) Error vectors with exactly two nonzero components:
   a) $\overline{e}_1 \neq \overline{0}$, and $\overline{e}_2 \neq \overline{0}$, and $\overline{e}_3 = \overline{0}$, for which

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2) + 2\cdot 2\alpha\langle \overline{e}_1, \overline{e}_2\rangle_X \\ &\geq (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2) - 2\cdot 2\alpha|\langle \overline{e}_1, \overline{e}_2\rangle_X| \\ &\geq (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2) - 2\cdot 2\alpha\|\overline{e}_1\|_X \cdot \|\overline{e}_2\|_X \\ &\geq (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2) - 2\alpha(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2) \\ &= (1+2\alpha^2-2\alpha)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2) \\ &\geq 2(1+2\alpha^2-2\alpha)\, d_0^2. \end{aligned}$$

Equality is achieved when $\overline{e}_1 = -\overline{e}_2$ and $\|\overline{e}_1\|_X^2 = \|\overline{e}_2\|_X^2 = d_0^2$.

   b) $\overline{e}_1 = \overline{0}$ and $\overline{e}_2 \neq \overline{0}$ and $\overline{e}_3 \neq \overline{0}$, for which, similarly as above

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(\|\overline{e}_2\|_X^2 + \|\overline{e}_3\|_X^2) + 2\cdot 2\alpha\langle \overline{e}_2, \overline{e}_3\rangle_X \\ &\geq 2(1+2\alpha^2-2\alpha)\, d_0^2. \end{aligned}$$

Equality is achieved when $\overline{e}_2 = -\overline{e}_3$ and $\|\overline{e}_2\|_X^2 = \|\overline{e}_3\|_X^2 = d_0^2$.

   c) $\overline{e}_1 \neq \overline{0}$ and $\overline{e}_2 = \overline{0}$ and $\overline{e}_3 \neq \overline{0}$, for which

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_3\|_X^2) + 2\cdot\alpha^2\langle \overline{e}_1, \overline{e}_3\rangle_X \\ &\geq (1+2\alpha^2-\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_3\|_X^2) \\ &\geq 2(1+\alpha^2)\, d_0^2. \end{aligned}$$

Equality is achieved when $\overline{e}_1 = -\overline{e}_3$ and $\|\overline{e}_1\|_X^2 = \|\overline{e}_3\|_X^2 = d_0^2$.

3) Error vectors with three nonzero components, i.e., $\overline{e}_1 \neq \overline{0}$ and $\overline{e}_2 \neq \overline{0}$ and $\overline{e}_3 \neq \overline{0}$.
   a) $\overline{e}_2 = -\overline{e}_1$, for which

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(2\|\overline{e}_1\|_X^2 + \|\overline{e}_3\|_X^2) - 4\alpha\|\overline{e}_1\|_X^2 \\ &\quad - 4\alpha\langle \overline{e}_1, \overline{e}_3\rangle_X + 2\alpha^2\langle \overline{e}_1, \overline{e}_3\rangle_X \\ &\geq (3+8\alpha^2-8\alpha)\, d_0^2. \end{aligned}$$

Equality is achieved when $\overline{e}_1 = \overline{e}_3$ and $\|\overline{e}_1\|_X^2 = \|\overline{e}_2\|_X^2 = \|\overline{e}_3\|_X^2 = d_0^2$.

   b) $\overline{e}_2 = -\overline{e}_3$, for which, similarly as above

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + 2\|\overline{e}_3\|_X^2) - 4\alpha\|\overline{e}_3\|_X^2 \\ &\quad - 4\alpha\langle \overline{e}_1, \overline{e}_3\rangle_X + 2\alpha^2\langle \overline{e}_1, \overline{e}_3\rangle_X \\ &\geq (3+8\alpha^2-8\alpha)\, d_0^2. \end{aligned}$$

Equality is achieved when $\overline{e}_1 = \overline{e}_3$ and $\|\overline{e}_1\|_X^2 = \|\overline{e}_2\|_X^2 = \|\overline{e}_3\|_X^2 = d_0^2$.

   c) $\overline{e}_2 \neq -\overline{e}_1$ and $\overline{e}_2 \neq -\overline{e}_3$ and $\overline{e}_2 = -(\overline{e}_1 + \overline{e}_3)$, for which

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2 + \|\overline{e}_3\|_X^2) - 3\alpha\|\overline{e}_2\|_X^2 \\ &\quad - \alpha\|\overline{e}_1 + \overline{e}_3\|_X^2 + 2\alpha^2\langle \overline{e}_1, \overline{e}_3\rangle_X \\ &> (3+8\alpha^2-7\alpha)\, d_0^2. \end{aligned}$$

   d) $\overline{e}_2 \neq -\overline{e}_1$ and $\overline{e}_2 \neq -\overline{e}_3$ and $\overline{e}_2 \neq -(\overline{e}_1 + \overline{e}_3)$, for which

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= (1+2\alpha^2)(\|\overline{e}_1\|_X^2 + \|\overline{e}_2\|_X^2 + \|\overline{e}_3\|_X^2) + 2\alpha\langle \overline{e}_1, \overline{e}_2\rangle_X \\ &\quad + 2\alpha\langle \overline{e}_2, \overline{e}_3\rangle_X + 2\alpha\langle \overline{e}_2, \overline{e}_1 + \overline{e}_3\rangle_X + 2\alpha^2\langle \overline{e}_1, \overline{e}_3\rangle_X \\ &\geq (3+8\alpha^2-8\alpha)\, d_0^2 + \alpha(6A^2 - d_0^2). \end{aligned}$$

Equality is achieved when $\overline{e}_1 = \overline{e}_3$ and $\|\overline{e}_1\|_X^2 = \|\overline{e}_3\|_X^2 = d_0^2$.

Comparing the above bounds on $d(\boldsymbol{\epsilon})$ gives the minimum distance as in (20)

$$d^2_{\min} = \begin{cases} 1+2\alpha^2)d_0^2, & \text{if } 0 \leq \alpha \leq 1-\sqrt{2}/2 \\ 2(1+2\alpha^2-2\alpha)d_0^2, & \text{if } 1-\sqrt{2}/2 \leq \alpha \leq 1/2. \end{cases}$$

## APPENDIX II
## PROOF OF THEOREM IV.5

The minimum distance $d^2_{\min}$ is obtained by minimizing $d^2(\boldsymbol{\epsilon})$ defined in (17) as

$$d^2(\boldsymbol{\epsilon}) = \sum_{l=1}^{N_t}\sum_{i=1}^{N_t} \langle \overline{e}_l, \overline{e}_i\rangle_X \gamma_{li}.$$

We partition the set of all possible error vectors $\boldsymbol{\epsilon} = (\overline{e}_1, \cdots, \overline{e}_{N_t})$ into two subsets.

1) Error vectors with exactly one nonzero component (single-track errors), i.e.,

$$\boldsymbol{\epsilon} = (0, \cdots, 0, \overline{e}_N, 0, \cdots, 0), \quad 1 \leq N \leq N_t,\ \|\overline{e}_N\|_0^2 \geq d_0^2$$

for which

$$\begin{aligned} d^2(\boldsymbol{\epsilon}) &= \sum_{l=1}^{N_t}\sum_{i=1}^{N_t} \langle \overline{e}_l, \overline{e}_i\rangle_X \gamma_{li} \\ &= \gamma_{NN}\|\overline{e}_N\|_X^2 \\ &\geq s_0 d_0^2. \end{aligned}$$

Equality is achieved when $\|\overline{e}_N\|_X = d_0^2$ and $\gamma_{NN} = s_0$.

2) Error vectors with at least two nonzero components (multitrack errors), i.e.,

$$\boldsymbol{\epsilon} = \{0, \cdots, 0, \overline{e}_{N_1}, \cdots, \overline{e}_{N_2}, 0, \cdots, 0\},$$
$$1 \leq N_1 < N_2 \leq N_t,\ \|\overline{e}_{N_1}\|_X^2, \|\overline{e}_{N_2}\|_X^2 \geq d_0^2$$

for which

$$
\begin{aligned}
d^2(\boldsymbol{\epsilon}) &= \sum_{l=1}^{N_t}\sum_{i=1}^{N_t}\langle \overline{e}_l, \overline{e}_i\rangle_X s_{|l-i|} \\
&\geq s_0 \sum_{l=N_1}^{N_2} ||\overline{e}_l||_X^2 - \sum_{i=1}^{N_t-1} 2|s_i| \sum_{l=N_1}^{N_2-i} |\langle \overline{e}_l, \overline{e}_{l+i}\rangle_X| \\
&\geq s_0 \sum_{l=N_1}^{N_2} ||\overline{e}_l||_X^2 \\
&\quad - \sum_{i=1}^{N_t-1} |s_i| \sum_{l=N_1}^{N_2-i} (||\overline{e}_l||_X^2 + ||\overline{e}_{l+i}||_X^2) \\
&\geq s_0 \sum_{l=N_1}^{N_2} ||\overline{e}_l||_X^2 - \sum_{i=1}^{N_t-1} 2|s_i| \left( \sum_{l=N_1}^{N_2} ||\overline{e}_l||_X^2 - d_0^2 \right) \\
&= s_0 d_0^2 + \left( s_0 - \sum_{i=1}^{N_t-1} 2|s_i| \right) \left( \sum_{l=N_1}^{N_2} ||\overline{e}_l||_X^2 - d_0^2 \right) \\
&> s_0 d_0^2
\end{aligned}
$$

where the last inequality follows if (21) is satisfied and the definition of $d_0^2$. Therefore, under condition (21), the minimum distance is given by $d_{\min}^2 = s_0 d_0^2$. □

## REFERENCES


[1] M. W. Marcellin and H. J. Weber, "Two-dimensional modulation codes," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 254–266, Jan. 1992.

[2] R. E. Swanson and J. K. Wolf, "A new class of two-dimensional RLL recording codes," *IEEE Trans. Magn.*, vol. 28, pp. 3407–3416, Nov. 1992.

[3] L. Kee and M. W. Marcellin, "A new construction for $n$-track $(d, k)$ codes with redundancy, in *Proc. 1994 IEEE Int. Symp. Information Theory (ISIT'94)* (Trondheim, Norway, June 1994), p. 145.

[4] P. A. Voois and J. M. Cioffi, "A decision feedback equalizer for multiple-head digital magnetic recording," in *Proc. 1991 IEEE Int. Conf. Communications (ICC '91)* (Denver, CO, June 1991), pp. 26.4.1.–26.4.5.

[5] ——, "Multichannel digital magnetic recording," in *Proc. 1992 IEEE Int. Conf. Communications (ICC '92)* (Chicago, IL, June 1992), pp. 305.4.1–305.4.6.

[6] ——, "Achievable radial information densities in magnetic recording systems," in *Proc. 1992 IEEE Global Telecommunications Conf. (GLOBECOM '92)* (Orlando, FL, Dec. 1992), pp. 1067–1071.

[7] L. C. Barbosa, "Simultaneous detection of readback signals from interfering magnetic recording tracks using array heads," *IEEE Trans. Magn.*, vol. 26, pp. 2163–2165, Sept. 1990.

[8] E. Soljanin "Coding for improving noise immunity in multi-track, multihead recording systems," Ph.D. dissertation, Texas A&M Univ., College Station, TX, Aug. 1994.

[9] E. Soljanin and C. N. Georghiades, "On coding in multi-track, multihead, disk recording systems," in *Proc. 1993 IEEE Global Telecommunications Conf. (GLOBECOM '93)* (Houston, TX, Nov. 1993), pp. 18–22.

[10] ——, "Coding for two-head recording systems," *IEEE Trans. Inform. Theory*, vol. 41, pp. 747–755, May 1995.

[11] ——, "Sliding-block codes for two-track magnetic recording channels," in *Proc. 28th Annual Conf. Information Sciences and Systems* (Princeton, NJ, Mar. 1994), pp. 1–5.

[12] ——, "Two-track codes for magnetic recording channels," in *Proc. 1994 IEEE Int. Symp. Information Theory (ISIT'94)* (Trondheim, Norway, June 1994), p. 150.

[13] M. Siala and G. K. Kaleh, "A two-track matched spectral-null code with rate 3/4 for the binary dicode channel," *IEEE Trans. Magn.*, vol. 30, pp. 2778–2787, Sept. 1994.

[14] T. Guzeoglu and J. G. Proakis, "A modulation coding scheme for multichannel PRML," in *Proc. 1994 IEEE Int. Symp. Information Theory (ISIT'94)* (Trondheim, Norway, June 1994), p. 213.

[15] L. C. Barbosa, C. E. MacDonald, and J. K. Wolf, "Array heads for estimating track misregistration," *IEEE Trans. Magn.*, vol. 29, pp. 4033–35, Nov. 1993.

[16] M. P. Vea and J. M. F. Moura, "Magnetic recording channel model with intertrack interference," *IEEE Trans. Magn.*, vol. 27, pp. 4834–4836, Nov. 1991.

[17] D. A. Lindholm, "Spacing losses in finite track width reproducing systems," *IEEE Trans. Magn.*, vol. MAG-14, pp. 55–59, Mar. 1978.

[18] L. C. Barbosa, "Maximum likelihood sequence estimators: a geometric view," *IEEE Trans. Inform. Theory*, vol. 35, pp. 419–427, Mar. 1989.

[19] J. G. Proakis, *Digital Communications.* New York, NY: McGraw-Hill, 1983.

[20] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data transmission systems," *IEEE Trans. Commun.*, vol. COM-22, pp. 624–636, May 1974.