

Multihead Multitrack Detection in Shingled Magnetic Recording with ITI Estimation

Bing Fan, Hemant K. Thapar, and Paul H. Siegel

Center for Magnetic Recording Research

University of California, San Diego, La Jolla, CA 92093, USA

bifan@ucsd.edu, hemantkthapar@gmail.com, psiegel@ucsd.edu

Abstract—Multitrack detection for shingled magnetic recording (SMR) using a two-head array system is considered. The channel suffers from intersymbol interference (ISI) in the down-track direction and intertrack interference (ITI) in the crosstrack direction. We propose a practical multihead/multitrack detector that provides a low-complexity approach to adaptive estimation of time-varying ITI. The performance of the proposed detection algorithm is analyzed in terms of its minimum distance parameter, and simulation results show that the proposed detector offers a performance advantage in settings where complexity constraints limit the maximum-likelihood two-track detector to use a static ITI estimate.

I. INTRODUCTION

Shingled magnetic recording (SMR) is one of the leading technologies proposed to achieve ultra-high areal storage density in next generation hard disk drives (HDD) [1]. In SMR, tracks are squeezed to increase the track density, measured in terms of tracks per inch (TPI). This goal can be achieved by using a large writing pole-tip to sequentially overwrite a portion of the previous track. In the readback process, because the tracks are narrower, the read head can sense signals from adjacent tracks when reading from the target track, causing intertrack interference (ITI).

Several techniques, based on different practical requirements, have been proposed to resolve the ITI problem. The performance of a single-head/single-track (SHST) system is studied in [2]. Iterative ITI cancellation, which removes ITI from each single-track readback signal before detection, is explored in [3] and [4]. These SHST techniques achieve acceptable performance when ITI is low, but suffer as ITI becomes severe. Multihead/multitrack (MHMT) schemes have attracted considerable attention because of their ability to better combat ITI. In MHMT, a group of tracks are read back by an array of heads and jointly processed. The performance of MHMT systems is analyzed in [5] and [6]. An iterative detection/decoding scheme for a two-track channel model with two heads is simulated in [7].

In this paper we study a linear and symmetric two-head/two-track (2H2T) model such as that used in [4] [5] [6] [7]. One problem associated with ITI is how to estimate the response from an adjacent track. The authors of [2] propose a least mean square (LMS) adaptive algorithm to estimate the off-track interference for the SHST system. For the 2H2T case, we reformulate this parameter estimation problem as a gain control model, and propose a novel detector – the weighted

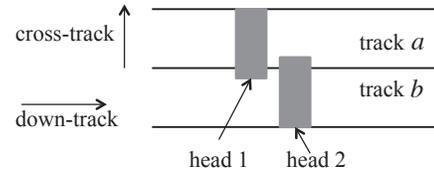


Fig. 1. Schematic of a two-head/two-track recording system

sum-subtract joint detector (WSSJD) – along with a gain loop to adaptively estimate the ITI level.

Another important issue associated with an optimal maximum-likelihood (ML) MHMT detector is its high computational complexity, which is proportional to $2^{M\nu}$, where M is the number of tracks jointly processed, and ν is the channel memory. For a system that jointly detects many tracks or that has a long channel impulse response, an ML detector will be impractical. The proposed WSSJD algorithm is amenable to a reduced-complexity implementation. Due to space limitations, this will be presented in a subsequent paper [8].

The paper is organized as follows. Section II introduces the 2H2T system model and reviews the optimal detector. In Section III we present the WSSJD and analyze its performance in terms of a minimum distance parameter. In Section IV we describe a gain loop structure to adaptively estimate the ITI level for use by the WSSJD algorithm. We then present performance simulation results in Section V and conclude the paper in Section VI.

II. TWO HEAD/TWO TRACK SYSTEM

We consider a linear and symmetric 2H2T system as shown in Fig. 1. Track a and track b are two adjacent tracks with no guard band between them. Let $x^a(D)$, $x^b(D)$ be the data sequences recorded on tracks a and b , with $x^i(D) = \sum_{k=0}^N x_k^i D^k$ and $x_k^i \in \{-1, +1\}$ for $i \in \{a, b\}$. We assume $x^a(D)$, $x^b(D)$ are both i.i.d. and equiprobable, and $x^b(D)$ is independent of $x^a(D)$. We also assume that there is no phase offset during the writing, i.e., the written patterns $x^a(D)$, $x^b(D)$ are perfectly aligned. Head 1 and head 2 have the same dimensions, are placed symmetrically over track a and track b , and move together in the down-track direction.

During readback, the signal from each head is passed through a matched filter, a sampler, and then equalized

to the target dipulse response represented by polynomial $h(D) = h_0 + h_1D + \dots + h_\nu D^\nu$ of degree ν . The interference from the side track is additive and formulated as a scaled output from the ISI channel $h(D)$. The noiseless outputs of the 2H2T channel are given by

$$\begin{aligned} y^a(D) &= x^a(D)h(D) + \epsilon x^b(D)h(D) \\ y^b(D) &= \epsilon x^a(D)h(D) + x^b(D)h(D) \end{aligned} \quad (1)$$

where ϵ represents the ITI level determined by the overlap between the head and the side track. In this section, we assume that the ITI level ϵ is static and known to the receiver when analyzing the detector performance. Later, in Section IV, we relax this assumption and incorporate adaptive ITI estimation into our detection architecture.

The received signals from head 1 and head 2 are further corrupted by the electronic noise, i.e.

$$\begin{aligned} r^a(D) &= y^a(D) + n^a(D) \\ r^b(D) &= y^b(D) + n^b(D) \end{aligned} \quad (2)$$

where $n^a(D), n^b(D)$ are uncorrelated and i.i.d sequences, with $n_k^a, n_k^b \sim \mathcal{N}(0, \sigma^2)$.

At the receiver, the ML detector makes a decision by finding the input pair $\hat{x}^a(D), \hat{x}^b(D)$ that maximize the log likelihood of the received signals, i.e.

$$\begin{aligned} &\hat{x}^a(D), \hat{x}^b(D) \\ &= \arg \max_{x^a, x^b} \log \Pr(r^a(D), r^b(D) | x^a(D), x^b(D)) \\ &= \arg \min_{x^a, x^b} \|r^a(D) - y^a(D)\|^2 + \|r^b(D) - y^b(D)\|^2 \end{aligned} \quad (3)$$

where $\|\cdot\|^2$ denotes the squared Euclidean norm,

$$\|x(D)\|^2 = \sum_k x_k^2.$$

In other words, the received sequences are jointly decoded to the sequence pair whose noiseless channel outputs are closest to the received signals in the output space. This can be done by passing the received signals through a two-track Viterbi detector. The trellis is designed to simultaneously recover both tracks. Each trellis edge goes from an initial state $s(k-1) = [x_{k-\nu}^a \dots x_{k-1}^a, x_{k-\nu}^b \dots x_{k-1}^b]$ to a terminal state $s(k) = [x_{k-\nu+1}^a \dots x_k^a, x_{k-\nu+1}^b \dots x_k^b]$ with input label $\mathcal{L}_{\text{in}} = (x_k^a, x_k^b)$ and output label $\mathcal{L}_{\text{out}} = (y_k^a, y_k^b)$. For a channel with memory ν , the trellis contains $2^{2\nu}$ states each of which is associated with 4 incoming and outgoing edges.

The ML detector needs to know the value ϵ to calculate the noiseless output label (y_k^a, y_k^b) given by equation (1). Therefore, the conventional ML detector works efficiently only when ϵ is static. For varying ϵ , the conventional ML detector has to recalculate the output label (y_k^a, y_k^b) whenever the value of ϵ changes. If the channel trellis has a large number of branches or if ϵ changes continuously, this adaptation process incurs considerable delay. On a real hard drive, however, ϵ generally varies spatially due to mechanical effects such as head skew

and flying height variation. Thus, adaptive estimation of ϵ will be necessary, introducing significant detection latency.

In the next two sections, we present a novel detection architecture that makes it possible to adaptively estimate ϵ while keeping the efficiency of ML detection. We show that the proposed approach achieves ML performance with static ITI, but has the flexibility to efficiently work with an adaptive estimator for the ITI level ϵ . The proposed detector uses a different trellis diagram than the conventional two-track ML detector. For convenience, we refer to the latter as the ‘‘ML trellis’’ even though both detectors produce the ML output sequences.

It is well known that the error event probability of the trellis-based ML detector can be approximated as $\text{Pe} \propto Q(\frac{d_{\text{min}}}{2\sigma})$, where the Q-function is the tail probability of the standard Gaussian distribution, d_{min} is the minimum distance parameter of the channel, and σ is the standard deviation of the additive Gaussian channel noise. Therefore, the performance of the detector can be accurately predicted by analyzing the minimum distance. As given in [6], the minimum distance parameter of the ML detector on the 2H2T channel is

$$d_{\text{min}}^2 = \begin{cases} (1 + \epsilon^2)d_0^2 & \text{if } 0 \leq \epsilon \leq 2 - \sqrt{3} \\ 2(1 - \epsilon)^2d_0^2 & \text{if } 2 - \sqrt{3} \leq \epsilon \leq 1/2 \end{cases} \quad (4)$$

where d_0 is the minimum distance of a single track with channel polynomial $h(D)$ when there is no ITI. When ITI is low, the single track error events are the minimum distance error patterns. When ITI increases, the double track error events become the dominant error events. The operating point that gives the highest minimum distance, or the best performance of the ML detector, is at $\epsilon = 2 - \sqrt{3}$.

III. WEIGHTED SUM-SUBTRACT JOINT DETECTION

The weighted sum-subtract joint detection (WSSJD) algorithm differs from the conventional ML detector in two respects. First, it adds a ‘‘sum-subtract’’ preprocessor before the Viterbi detector. Second, it uses weighted branch metrics in the Viterbi detector. When we introduce the algorithm, we assume ϵ to be known. This condition will be relaxed in Section IV where we show that ϵ acts as a gain factor that can be estimated by means of a first-order gain loop.

Instead of directly passing the received sequences $r^a(D)$ and $r^b(D)$ to the Viterbi detector, the WSSJD first calculates their sum $r^+(D)$ and difference $r^-(D)$, normalized by $\frac{1}{1+\epsilon}$ and $\frac{1}{1-\epsilon}$, respectively, i.e.,

$$\begin{aligned} r^+(D) &= \frac{1}{1+\epsilon} (r^a(D) + r^b(D)) \\ r^-(D) &= \frac{1}{1-\epsilon} (r^a(D) - r^b(D)). \end{aligned} \quad (5)$$

Defining the sum and difference input signals by

$$z^+(D) = x^a(D) + x^b(D), \quad z^-(D) = x^a(D) - x^b(D) \quad (6)$$

and the corresponding noiseless output signals by

$$y^+(D) = z^+(D)h(D), \quad y^-(D) = z^-(D)h(D) \quad (7)$$

x_k^a	x_k^b	z_k^+	z_k^-
1	1	2	0
1	-1	0	2
-1	1	0	-2
-1	-1	-2	0

TABLE I: mapping between (x_k^a, x_k^b) and (z_k^+, z_k^-)

we can rewrite equation (5) as

$$\begin{aligned} r^+(D) &= y^+(D) + n^+(D) \\ r^-(D) &= y^-(D) + n^-(D) \end{aligned} \quad (8)$$

where the Gaussian noise components

$$\begin{aligned} n^+(D) &= \frac{1}{1+\epsilon} (n^a(D) + n^b(D)), \\ n^-(D) &= \frac{1}{1-\epsilon} (n^a(D) - n^b(D)) \end{aligned} \quad (9)$$

satisfy $n_k^+ \sim \mathcal{N}(0, \frac{2\sigma^2}{(1+\epsilon)^2})$, $n_k^- \sim \mathcal{N}(0, \frac{2\sigma^2}{(1-\epsilon)^2})$. Furthermore,

$$E(n_k^+ n_k^-) = \frac{1}{1-\epsilon^2} (E(n_k^{a2}) - E(n_k^{b2})) = 0 \quad (10)$$

which implies that $n^+(D)$ and $n^-(D)$ are uncorrelated and, therefore, independent.

We can think of $r^+(D)$ and $r^-(D)$ as the noisy outputs obtained by passing each of $z^+(D)$ and $z^-(D)$ through a channel $h(D)$, but with different SNRs. These channels are called the ‘‘sum channel’’ and the ‘‘subtract channel,’’ respectively. Notice that corresponding input sequences $z^+(D)$ and $z^-(D)$ have a three-level alphabet, $\mathcal{B} = \{-2, 0, 2\}$. There is a one-to-one mapping between (z_k^+, z_k^-) and (x_k^a, x_k^b) , as shown in Table I.

Since $r^+(D)$ and $r^-(D)$ are obtained from separate channels, one can independently detect $z^+(D)$ and $z^-(D)$, and then map (z_k^+, z_k^-) to (x_k^a, x_k^b) according to Table I. This method corresponds to solving two detection problems

$$\begin{aligned} \hat{z}^+(D) &= \arg \max_{z^+} \log \Pr(r^+(D) | z^+(D)) \\ &= \arg \min_{z^+} \|r^+(D) - z^+(D)\|^2 \\ \hat{z}^-(D) &= \arg \max_{z^-} \log \Pr(r^-(D) | z^-(D)) \\ &= \arg \min_{z^-} \|r^-(D) - z^-(D)\|^2. \end{aligned} \quad (11)$$

However, this approach is not optimal. From Table I we see that $z^+(D)$ and $z^-(D)$ are not independent, e.g., $z_k^+ = 2$ forces z_k^- to be 0. Independent detection ignores this correlation and produces some undecodable $(\hat{z}_k^+, \hat{z}_k^-)$ pairs. Optimal detection must jointly consider both the sum channel and the subtract channel, determining

$$\begin{aligned} \hat{z}^+(D), \hat{z}^-(D) \\ = \arg \max_{z^+, z^-} \log \Pr(r^+(D), r^-(D) | z^+(D), z^-(D)). \end{aligned} \quad (12)$$

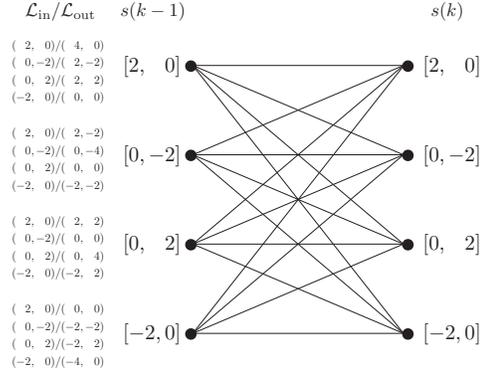


Fig. 2. WSSJD trellis for channel $h(D) = 1 + D$

The WSSJD provides a practical trellis-based algorithm for solving this problem. The WSSJD trellis has the same number of states as the ML trellis. Each branch connects an initial state $s(k-1) = [z_{k-\nu}^+ \dots z_{k-1}^+, z_{k-\nu}^- \dots z_{k-1}^-]$ to a terminal state $s(k) = [z_{k-\nu+1}^+ \dots z_k^+, z_{k-\nu+1}^- \dots z_k^-]$ with input label $\mathcal{L}_{in} = (z_k^+, z_k^-)$ and output label $\mathcal{L}_{out} = (y_k^+, y_k^-)$. Fig. 2 shows a WSSJD trellis for the channel $h(D) = 1 + D$. The text to the left of each state lists the branch labels in the form of input/output. Note that, unlike the ML trellis, the WSSJD trellis is independent of ϵ .

Since the sum channel and the subtract channel have different noise powers, the WSSJD computes a weighted sum of their individual distance metrics, $\|r^+(D) - y^+(D)\|^2$ and $\|r^-(D) - y^-(D)\|^2$. The optimal choice of the weights is found by evaluating equation (12):

$$\begin{aligned} \hat{z}^+(D), \hat{z}^-(D) \\ &= \arg \max_{z^+, z^-} \log \Pr(r^+(D), r^-(D) | z^+(D), z^-(D)) \\ &= \arg \max_{z^+, z^-} \log \Pr(r^+(D) | z^+(D)) + \log \Pr(r^-(D) | z^-(D)) \\ &= \arg \min_{z^+, z^-} \frac{\|r^+(D) - y^+(D)\|^2}{2\sigma^2/(1+\epsilon)^2} + \frac{\|r^-(D) - y^-(D)\|^2}{2\sigma^2/(1-\epsilon)^2} \\ &= \arg \min_{z^+, z^-} (1+\epsilon)^2 \|r^+(D) - y^+(D)\|^2 \\ &\quad + (1-\epsilon)^2 \|r^-(D) - y^-(D)\|^2. \end{aligned} \quad (13)$$

Let $M_{k-1}(s)$ denote the survivor path metric for state s at time $k-1$. Then equation (13) suggests that the path metric corresponding to the extension along a branch from state s to state s' at time k is

$$\begin{aligned} M_k(s') &= M_{k-1}(s) + (1+\epsilon)^2 (r_k^+ - y_k^+)^2 \\ &\quad + (1-\epsilon)^2 (r_k^- - y_k^-)^2 \end{aligned} \quad (14)$$

where (y_k^+, y_k^-) is the output label of the branch. The term $m(s, s') = (1+\epsilon)^2 (r_k^+ - y_k^+)^2 + (1-\epsilon)^2 (r_k^- - y_k^-)^2$ is called the weighted branch metric.

Since the transformation in the sum-subtract preprocessing

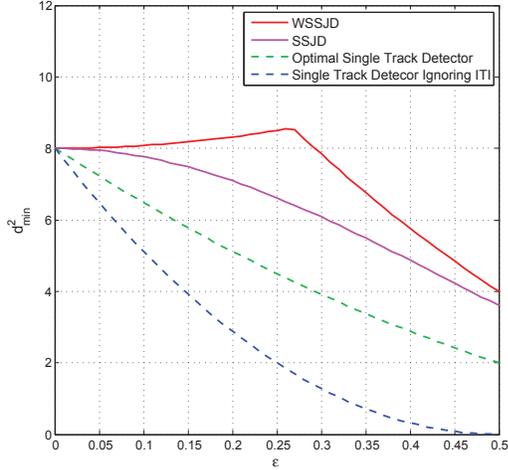


Fig. 3. Minimum squared distance comparison of different detectors on channel $h(D) = 1 + D$ with $d_0^2 = 8$

is bijective, we have

$$\Pr(r^+(D), r^-(D) | z^+(D), z^-(D)) = \Pr(r^a(D), r^b(D) | x^a(D), x^b(D)). \quad (15)$$

Therefore WSSJD gives the ML solution, so the minimum distance parameter governing its estimated performance is the same as the ML detector, shown in equation (4). The detector that ignores the weighting factors, i.e., that uses

$$m(s, s') = (r_k^+ - y_k^+)^2 + (r_k^- - y_k^-)^2 \quad (16)$$

as the branch metric, is suboptimal. We refer to this as sum-subtract joint detection (SSJD). The performance loss incurred by SSJD is reflected in its minimum distance parameter, which is given by

$$d_{\min}^2(\text{SSJD}) = \frac{(1 + \epsilon)^2(1 - \epsilon)^2}{1 + \epsilon^2} d_0^2. \quad (17)$$

In contrast to WSSJD, the minimum distance of SSJD is dominated by single-track error events for all $\epsilon \in [0, 0.5]$. Fig. 3 shows the squared minimum distance as a function of ϵ for WSSJD and SSJD, as well as for two single track detectors [9] included for comparison purposes.

The properties of WSSJD are summarized as follows. First, WSSJD is ML equivalent. Second, the WSSJD trellis is independent of ϵ , which only affects the noise components in the independent sum and subtract channels and is taken into account by suitably weighting their respective branch metrics. This independence is the key to combining WSSJD with adaptive estimation of ϵ .

IV. ADAPTIVE ITI LEVEL ESTIMATION

A. ITI Sensitivity

To evaluate the sensitivity of the various detectors to a small change in the ITI level, we introduce a small offset into our

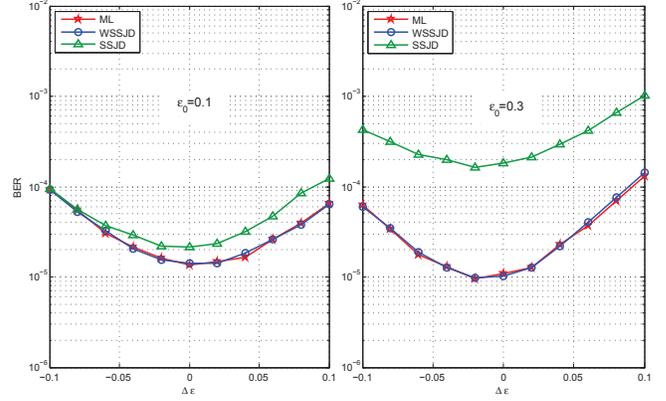


Fig. 4. ITI sensitivity simulation results for different detectors with $\epsilon_0 = 0.1$ (left) and $\epsilon_0 = 0.3$ (right).

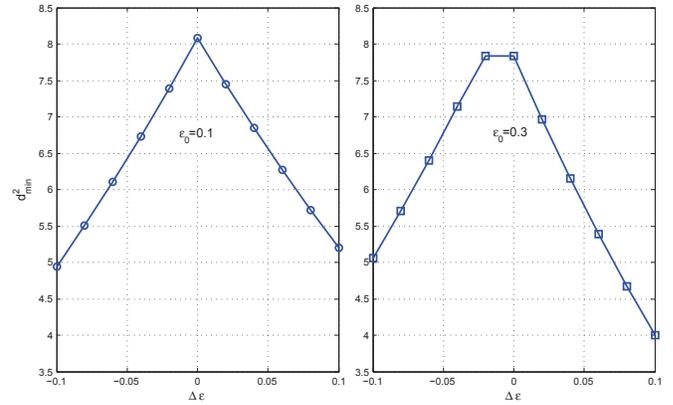


Fig. 5. The effective minimum distance of the ML detector as a function of mismatch $\Delta\epsilon$ with $\epsilon_0 = 0.1$ (left) and $\epsilon_0 = 0.3$ (right).

performance simulations. Suppose the nominal level is ϵ_0 , with offset $\Delta\epsilon$. The noiseless readback signals are then given by

$$\begin{aligned} y^a(D) &= x^a(D)h(D) + (\epsilon_0 + \Delta\epsilon)x^b(D)h(D) \\ y^b(D) &= x^b(D)h(D) + (\epsilon_0 + \Delta\epsilon)x^a(D)h(D). \end{aligned} \quad (18)$$

Fig. 4 shows the simulated bit error rate (BER) as a function of the ITI mismatch $\Delta\epsilon$ for the ML, WSSJD, and SSJD detectors on the channel $h(D) = 1 + D$ at SNR = 10dB, with $\epsilon_0 = 0.1$ and $\epsilon_0 = 0.3$, respectively. We also notice that the BER curves are not symmetric about $\Delta\epsilon = 0$. Furthermore, the minimum BER points occur at offsets with opposite polarity for $\epsilon_0 = 0.1$ and $\epsilon_0 = 0.3$.

Fig. 3 suggests that the observed behaviors are due to minimum distance properties of the mismatched detectors. To confirm this, we conducted an exhaustive search for the minimum distance of the mismatched system at $\epsilon_0 = 0.1$ and $\epsilon_0 = 0.3$, with offsets $\Delta\epsilon \in [-0.1, 0.1]$. The results are plotted in Fig. 5. For $\epsilon_0 = 0.1$, a positive offset in this range tends to give higher minimum distance than a negative offset of the same magnitude. For $\epsilon_0 = 0.3$, this situation is reversed, and in a small range of negative offsets, $\Delta\epsilon \in [-0.02, 0]$, the mismatch

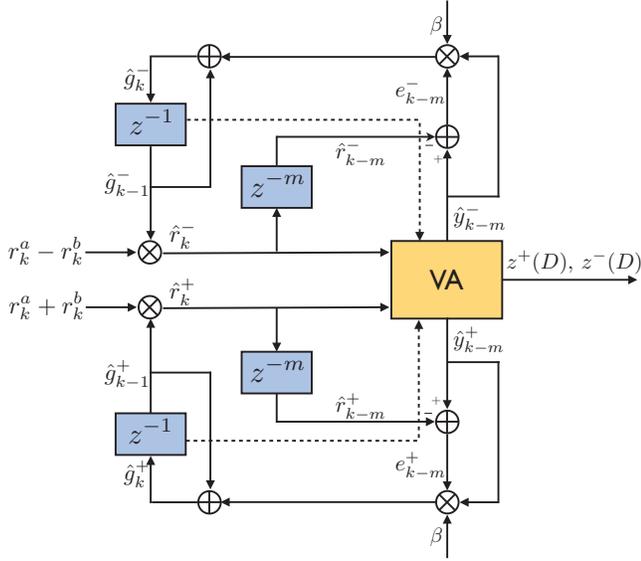


Fig. 6. WSSJD with gain control to adaptively estimate ITI level

doesn't reduce the minimum distance of the system. These results are consistent with the observed BER performance.

The explanation for the asymmetry about $\Delta\epsilon = 0$ is that the mismatch affects distance properties of the single-track error events and the double-track error events differently. For the ML detector on the $h(D) = 1 + D$ channel, the minimum distance of the single-track error events is

$$d_s^2 = \begin{cases} \frac{8(1+\epsilon_0^2-2\Delta\epsilon)^2}{1+\epsilon_0^2} & \text{if } \Delta\epsilon > 0 \\ \frac{8[1+\epsilon_0^2+(2+2\epsilon_0)\Delta\epsilon]^2}{1+\epsilon_0^2} & \text{if } \Delta\epsilon < 0 \end{cases} \quad (19)$$

while the minimum distance of the double-track error events is given by

$$d_d^2 = \begin{cases} 16[(1-\epsilon_0) - 2\Delta\epsilon]^2 & \text{if } \Delta\epsilon > 0 \\ 16(1-\epsilon_0)^2 & \text{if } \Delta\epsilon < 0. \end{cases} \quad (20)$$

The overall minimum distance of the system is

$$d_{\min}^2 = \min \{d_s^2, d_d^2\} \quad (21)$$

which matches the plot in Fig. 5.

B. Gain Loop

Recall that in the sum-subtract preprocessing, ϵ appears in the gain factors that normalize signals $r^+(D)$, $r^-(D)$. We rewrite equation (5) as

$$\begin{aligned} r^+(D) &= g^+ (r^a(D) + r^b(D)) \\ r^-(D) &= g^- (r^a(D) - r^b(D)) \end{aligned} \quad (22)$$

where $g^+ = \frac{1}{1+\epsilon}$, $g^- = \frac{1}{1-\epsilon}$ are the gain factors. We use the LMS adaptive algorithm to estimate these parameters. For \hat{g}^+ , the updating rule is given by

$$\hat{r}_k^+ = \hat{g}_{k-1}^+ (r_k^a + r_k^b) \quad (23)$$

$$e_k = \hat{y}_k^+ - \hat{r}_k^+ \quad (24)$$

$$\hat{g}_k^+ = \hat{g}_{k-1}^+ + \beta \hat{y}_k^+ e_k \quad (25)$$

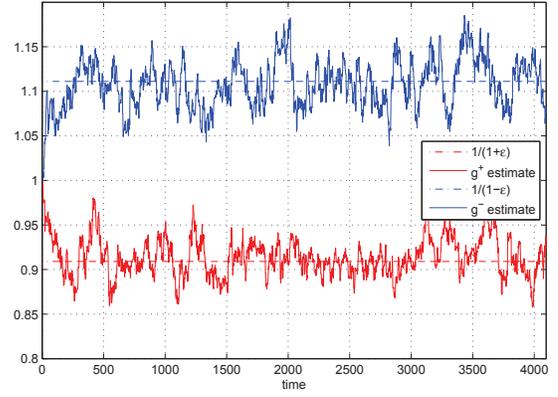


Fig. 7. Adaptive estimation of g^+ and g^- over one sector of 4096 bits on channel $h(d) = 1 + D$ at SNR = 10dB.

where β is the step-size parameter and \hat{y}_k^+ is the instantaneous decision fed back from the Viterbi detector. The step-size parameter β controls the convergence speed. A large β makes the loops converge faster, but also results in larger variance.

One can also introduce a small delay $m \geq 1$ to get more accurate tentative decisions. In this case, equation (24) and equation (25) become

$$e_{k-m} = \hat{y}_{k-m}^+ - \hat{r}_{k-m}^+ \quad (26)$$

$$\hat{g}_k^+ = \hat{g}_{k-1}^+ + \beta \hat{y}_{k-m}^+ e_{k-m} \quad (27)$$

The estimates \hat{g}_k^+ , \hat{g}_k^- will be used in the next iteration, and also in the Viterbi detector path metric calculation equation (14), i.e.,

$$M_k(s') = M_{k-1}(s) + \hat{g}_k^+{}^2 (r_k^+ - y_k^+)^2 + \hat{g}_k^-{}^2 (r_k^- - y_k^-)^2. \quad (28)$$

Fig. 6 shows a complete block diagram for WSSJD with adaptive gain estimation. The system contains two separate gain loops for \hat{g}_k^+ and \hat{g}_k^- . While a combined loop for estimating \hat{g}_k^+ and \hat{g}_k^- can provide a better estimate for ϵ , using separate loops achieves similar performance and is more efficient.

In our simulations, \hat{g}_0^+ and \hat{g}_0^- are initially set to 1. At time k , $r_k^a + r_k^b$ and $r_k^a - r_k^b$ are normalized by the previously estimated gain factors \hat{g}_{k-1}^+ and \hat{g}_{k-1}^- , respectively. The resulting signals \hat{r}_k^+ and \hat{r}_k^- are sent to the Viterbi detector. The path metric of each trellis state is evaluated and scaled by \hat{g}_{k-1}^+ and \hat{g}_{k-1}^- . After comparing the path metrics, the Viterbi detector picks the most likely path, and feeds back its decision on \hat{y}_{k-m}^+ and \hat{y}_{k-m}^- . The error signal is calculated to update \hat{g}_k^+ and \hat{g}_k^- . Note that SSJD can also work with these gain loops, without feeding \hat{g}_k^+ and \hat{g}_k^- to the path metric evaluation.

Fig. 7 shows the behavior of the g_k^+ and g_k^- gain loops in one sector of length $N = 4096$ bits on the channel $h(D) = 1 + D$ at SNR = 10dB with step-size $\beta = 0.005$ and delay $m = 5$. For channels with longer memory, a larger delay m may be adopted.

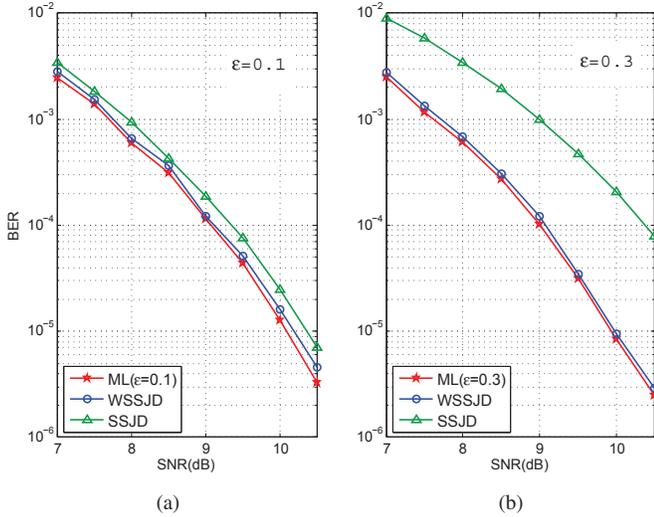


Fig. 8. BER vs. SNR of different detectors with (a) $\epsilon = 0.1$ and (b) $\epsilon = 0.3$.

V. SIMULATION RESULTS

We simulate WSSJD and SSJD with gain control on the 2H2T system with channel polynomial $h(D) = 1 + D$. In both cases we set $\beta = 0.008$ and $m = 5$. The initial values of gain factors g_0^+ and g_0^- are obtained by passing training samples through the system. The SNR is defined as

$$\text{SNR(dB)} = 10 \log \frac{\|h(D)\|^2}{2\sigma^2}$$

We first test the performance of the gain control loops when ϵ is fixed. Fig. 8 compares bit error rate (BER) vs. SNR of the ML detector, WSSJD, and SSJD, for $\epsilon = 0.1$ and $\epsilon = 0.3$. The frame size is 4096 bits. We assume that the ML detector knows the value ϵ , while WSSJD and SSJD adaptively estimate ϵ as in Fig. 6. The static ML detector provides a lower bound for optimal BER performance. It can be seen that adaptive WSSJD performs very close to the static ML detector. As expected from the minimum distance plots in Fig. 3, the performance of the SSJD is more severely degraded when $\epsilon = 0.3$ than when $\epsilon = 0.1$. The measures of frame error rate (FER) vs. SNR correlate well with the BER curves in the simulations.

Next, we test the performance of the detectors with a dynamic ITI model in which ϵ changes slowly with respect to the location k in a sector. Specifically, we set

$$\epsilon(k) = \epsilon_0 + 0.1 \sin(4\pi(k/N))$$

where $N = 4096$ is the frame size and ϵ_0 is the mean ITI value. The ML detector again uses the static value ϵ_0 , while WSSJD and SSJD adaptively estimate $\epsilon(k)$. The simulation results, shown in Fig. 9, suggest that the adaptive algorithms outperform the static ML detector by about 0.3-0.5dB at high SNR.

VI. CONCLUSION

In this paper we propose a novel two-track detector on a two-head, two-track (2H2T) channel with intertrack interference (ITI). The proposed weighted sum-subtract joint detector

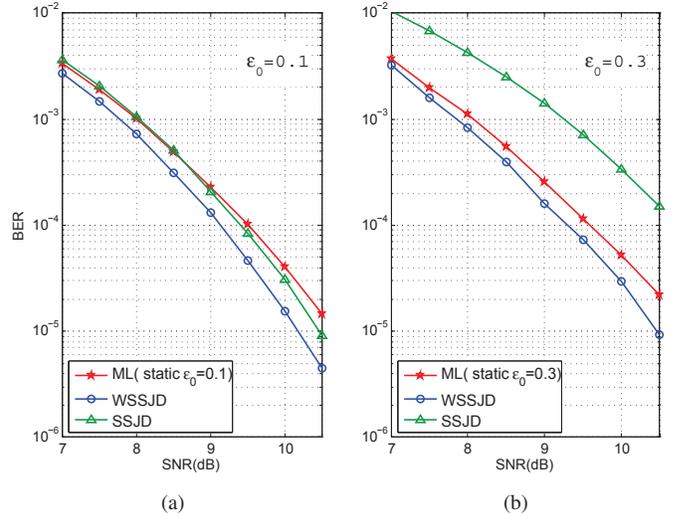


Fig. 9. BER vs. SNR of different detectors with ϵ slowly varying about the mean value (a) $\epsilon_0 = 0.1$ and (b) $\epsilon_0 = 0.3$.

(WSSJD) contains a sum-subtract preprocessing step and uses weighted branch metrics from the constituent sum and subtract channels. We use minimum distance analysis to explain the observed behavior of WSSJD under ITI mismatch conditions. The WSSJD separates the ITI level from its trellis structure, so, unlike the traditional joint-track ML detector, WSSJD can be efficiently combined with control loops that adaptively track the ITI level. Simulation results demonstrate the effectiveness of the adaptive WSSJD system with first-order loops in the presence of slowly varying ITI.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grant CCF-1405119.

REFERENCES

- [1] R. Wood, M. Williams, A. Kavcic, and J. Miles, "The feasibility of magnetic recording at 10 terabits per square inch on conventional media," *IEEE Trans. Magn.*, vol. 45, no. 2, pp. 917–923, Feb. 2009.
- [2] B. G. Roh, S.-U. Lee, J. Moon, and Y. Chen, "Single-head/single-track detection in interfering tracks," *IEEE Trans. Magn.*, vol. 38, no. 4, pp. 1830–1838, Jul. 2002.
- [3] N. Kumar, J. Bellorado, M. Marrow, and K. K. Chan, "Inter-track interference cancellation in presence of frequency offset for shingled magnetic recording," in *IEEE ICC 2013*, Jun. 2013, pp. 4342–4346.
- [4] M. Fujii and N. Shinohara, "Multi-track iterative iti canceller for shingled write recording," in *ISCIT*, Oct. 2010, pp. 1062–1067.
- [5] L. Barbosa, "Simultaneous detection of readback signals from interfering magnetic recording tracks using array heads," *IEEE Trans. Magn.*, vol. 26, no. 5, pp. 2163–2165, Sep. 1990.
- [6] E. Soljanin and C. Georghiades, "Multihead detection for multitrack recording channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2988–2997, Nov. 1998.
- [7] X. Ma and L. Ping, "Iterative detection/decoding for two-track partial response channels," *IEEE Communications Letters*, vol. 8, no. 7, pp. 464–466, Jul. 2004.
- [8] B. Fan, H. K. Thapar, and P. H. Siegel, "Multihead multitrack detection with reduced-state sequence estimation in shingled magnetic recording," to appear in the Proceedings of IEEE Intermag 2015, Beijing, China.
- [9] E. Soljanin and C. Georghiades, "On coding in multi-track, multi-head, disk recording systems," in *IEEE GLOBECOM*, vol. 4, Nov. 1993, pp. 18–22.