

# Multihead Multitrack Detection for Next Generation Magnetic Recording, Part II: Complexity Reduction — Algorithms and Performance Analysis

Bing Fan, *Student Member, IEEE*, Hemant K. Thapar, *Fellow, IEEE*, and Paul H. Siegel, *Fellow, IEEE*

**Abstract**—To achieve large storage capacity on magnetic hard disk drives, very high track density is required, causing severe intertrack interference (ITI). Multihead multitrack (MHMT) detection has been proposed to better combat the effects of ITI. Such detection, however, has prohibitive implementation complexity. Reduced-state sequence estimation (RSSE) is a promising technique for significantly reducing the complexity, while retaining good performance. In this paper, several different MHMT models are considered, including symmetric and asymmetric 2H2T systems, and a symmetric 3H3T system. By carefully evaluating the effective distance between two input symbols, we propose optimized set partition trees for each channel model. Different trellis configurations for RSSE are constructed based on the desired performance/complexity tradeoff. Simulation results show that the reduced MHMT detector can achieve near maximum-likelihood (ML) performance with a small fraction of the original number of trellis states. We also use error event analysis to explain the behavior of RSSE. The proposed algorithm could be potentially applied to next generation magnetic recording systems, especially when the ML detector is infeasible due to the high computational complexity.

**Index Terms**—Shingled magnetic recording, multitrack multithread detection, intertrack interference, reduced-state sequence estimation.

## I. INTRODUCTION

INTERTRACK interference (ITI), caused by aggressively shrinking the track pitch, is one of the more severe impairments in next generation hard disk drives (HDDs) [1], [2]. The use of an array reader to simultaneously read and process multiple tracks has recently drawn intensive interest because of its capability to handle ITI as well as electronic noise [3], [4]. The associated maximum likelihood (ML) detector complexity is, however, drastically increased.

Let  $\mathbf{x}_i = [x_i^1, x_i^2, \dots, x_i^n]^\top$ ,  $x_i^j \in \{-1, +1\}$ , be a column vector of the input symbols written on  $n$  adjacent tracks at time  $i$ . Let  $\mathbf{x}(D) = [x^1(D), x^2(D), \dots, x^n(D)]$  denote the

collection of sequences recorded on  $n$  tracks, where  $D$  is the delay unit. Assume all the tracks are equalized to the same channel polynomial  $h(D) = h_0 + h_1D + \dots + h_vD^v$ . An  $n$ -head,  $n$ -track ( $nHnT$ ) system is generally modeled as

$$\mathbf{r}_i = A_n \mathbf{y}_i + \boldsymbol{\omega}_i, \quad (1)$$

where  $\mathbf{r}_i = [r_i^1, r_i^2, \dots, r_i^n]^\top$  is the vector of received signals from  $n$  heads,  $\mathbf{y}_i = [y_i^1, y_i^2, \dots, y_i^n]^\top$  is a vector of noiseless channel outputs,  $y_i^j = \sum_{k=0}^v h_k x_{i-k}^j$ , and  $\boldsymbol{\omega}_i = [\omega_i^1, \omega_i^2, \dots, \omega_i^n]^\top$ ,  $\omega_i \sim \mathcal{N}(0, \sigma^2)$  is a vector of independent Gaussian electronic noise samples. We assume the noise samples are uncorrelated across tracks, i.e.  $E[\boldsymbol{\omega}_i \boldsymbol{\omega}_i^\top] = \sigma^2 I_n$ , where  $I_n$  is the  $n \times n$  identity matrix. The ITI effect is characterized by an  $n \times n$  matrix  $A_n$ . We will be primarily interested in the situation when

$$A_n = \begin{bmatrix} 1 & \epsilon & & & \\ \epsilon & 1 & \ddots & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \ddots & & 1 & \epsilon \\ & & & \epsilon & 1 \end{bmatrix}, \quad (2)$$

where  $\epsilon \in [0, 0.5]$  represents the ITI level. Such a system is symmetric, and the ITI only comes from the immediately adjacent tracks.  $\epsilon$  is i.i.d. with  $n_k^a, n_k^b \sim \mathcal{N}(0, \sigma^2)$ .

The ML detector decodes  $n$  tracks by simultaneously processing readback signals from  $n$  heads [5], [6]. The resulting joint trellis is composed of  $2^{nv}$  states, each associated with  $2^n$  incoming and outgoing edges. Constructing the trellis requires knowledge of  $\epsilon$ , which is generally time varying, and unknown to the receiver. This problem is resolved in Part I. More specifically, in Part I we propose a novel ML-equivalent detection method - weighted sum subtract joint detector (WSSJD) - along with a gain loop structure that can estimate the ITI as well as adapt itself to the new estimates. This work was partially presented in [7].

In Part II, we explore ways to reduce the complexity of MHMT detection, which is another challenging problem that needs to be solved to make MHMT practical. Compared to the traditional single-head single-track (SHST) detector with complexity  $O(2^v)$ , the  $nHnT$  ML detector has complexity  $O(2^{nv})$ . For  $v > 3$ , which is typical in practical recording channels, direct implementation of the  $nHnT$  ML detector could become infeasible even for small  $n$ . On the other hand, reduced-state sequence estimation (RSSE) [8], which was first proposed for

Manuscript received May 27, 2016; revised September 30, 2016; accepted November 18, 2016. Date of publication January 16, 2017; date of current version April 14, 2017. This work was supported in part by the National Science Foundation under Grant CCF-1405119, and the Center for Memory and Recording Research (formerly, Center for Magnetic Recording Research) at UC San Diego. The associate editor coordinating the review of this paper and approving it for publication was L. Dolecek.

B. Fan and P. H. Siegel are with the Center for Memory and Recording Research, Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093, USA (e-mail: bifan@ucsd.edu; psiegel@ucsd.edu).

H. K. Thapar is with OmniTier, Inc., Santa Clara, CA 95054, USA (e-mail: hemantkthapar@gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCOMM.2017.2652456

transmitting quadrature amplitude modulation (QAM) symbols through a partial response channel with long memory, is a good candidate to mitigate the complexity issue of MHMT detector. The RSSE trellis, originally constructed based on the Ungerboeck set partition tree, has fewer states. We note that the M-Viterbi algorithm [9] offers similar performance and complexity to RSSE, but it uses an ad hoc larger trellis which complicates its analysis and hardware implementation.

Efforts have been made to develop similar algorithms for use in MHMT detection. In particular, Kurtas *et al.* [10] presented a way to apply RSSE to MHMT system, but their construction generally suffers from high performance loss.

We propose a different approach, which is based on our work in Part I. We find that the channel transformation in WSSJD decomposes the system in (1) into  $n$  parallel sub-channels, which naturally leads to a set partition rule on the MHMT input constellation. For the simplest 2H2T system, the channel after transformation becomes a QAM-type model, and the reduced-state trellis can be constructed by redefining the distance measure on the transformed input constellation. The resulting four-level set partition tree provides better flexibility in performance/complexity tradeoffs. Our simulation results show that, with fewer than half the number of the full ML trellis states, RSSE can achieve near-ML performance on many channels. The concept of using RSSE in 2H2T case was partially presented in [11] and [12], and some of that discussion is briefly restated in this paper for the sake of completeness and better understanding. Further details about implementation issues and more thorough performance evaluations are also provided here.

We further show that the evaluation of RSSE performance is tractable through error event analysis. In contrast to ML detection, some error events in RSSE merge early due to the reduced-state trellis structure. We introduce an early-merging condition to identify these error events, and a modified error state diagram is used to search for the dominant early-merged error events. The search results for several reduced-state trellis configurations at different ITI levels are presented. When the minimum distance parameter of the early-merged error events is larger than that of the ML detector, the performance loss of the RSSE trellis is almost negligible. An asymmetric 2H2T system is also considered because of its potential practical interest. The error event analysis shows that the proposed set partitioning rule is also applicable in the asymmetric case.

Finally, we investigate a more complex 3H3T model. The effective distance between symbols shows different monotonicity behavior as the ITI level changes. Therefore, we propose two types of set partition trees, one suitable for the low ITI environment, and the other better suited for high ITI levels. Simulation results are provided for both cases, and they show that RSSE can significantly reduce the computational complexity of 3H3T detection while retaining acceptable performance.

The paper is organized as follows. In Section II we briefly review the original RSSE algorithm developed for the QAM system, and the WSSJD algorithm. Next, in Section III, we show how to construct a reduced-state trellis for the symmetric 2H2T channel by redefining the distance measure in the

input constellation and designing proper set partitioning trees. Performance simulation results for RSSE on 2H2T systems with different channel models and trellis configurations are provided in Section IV. Section V gives the early-merging condition and error event analysis. The dominant error events for several reduced-state trellises on different channels are also tabulated. In Section VI we consider the applicability of the RSSE algorithm to the asymmetric 2H2T model. Finally, set partition trees of the 3H3T system and their performance is analyzed and simulated in Section VII. Conclusions are presented in Section VIII.

## II. BACKGROUND

### A. Review of RSSE

The RSSE algorithm proposed in [8] was primarily designed for transmitting QAM symbols through an ISI channel with memory  $\nu$ . Recall that the trellis state in the ML detector is represented as

$$\mathbf{p}_i = [\mathbf{x}_{i-1}, \mathbf{x}_{i-2}, \dots, \mathbf{x}_{i-\nu}], \quad (3)$$

where each symbol  $\mathbf{x}_{i-k}$  is selected from a complex-valued signal set  $\mathcal{C}$  whose size is  $M$ . In RSSE, the reduced complexity trellis is constructed by grouping several ML states into a **subset state**. To do this, for the  $k$ th element  $\mathbf{x}_{i-k}$  in  $\mathbf{p}_i$ , a set partition  $\Omega(k)$  of  $\mathcal{C}$  is defined, and  $\mathbf{x}_{i-k}$  is represented by its subset index  $a_{i-k}(k)$  in  $\Omega(k)$ . Notice that  $\Omega(k)$  can be different for  $k = 1, \dots, \nu$ . Let  $J_k = |\Omega(k)|$  be the number of subsets in partition  $\Omega(k)$ ,  $1 \leq J_k \leq M$ . Then the subset index  $a_{i-k}(k)$  can take its value from  $0, 1, \dots, J_k - 1$ . The corresponding subset state of  $\mathbf{p}_i$  is denoted by

$$\mathbf{s}_i = [a_{i-1}(1), a_{i-2}(2), \dots, a_{i-\nu}(\nu)]. \quad (4)$$

The trellis constructed from all possible  $\mathbf{s}_n$  is called the **subset trellis**. To obtain a well-defined trellis structure, the partition  $\Omega(k)$  is restricted to be a further partition of the subsets in  $\Omega(k+1)$ , for  $1 \leq k \leq \nu - 1$ . This condition guarantees that for a given state  $\mathbf{s}_i$  and current input  $\mathbf{x}_i$ , the next subset state is uniquely determined and represented as

$$\mathbf{s}_{i+1} = [a_i(1), a_{i-1}(2), \dots, a_{i-\nu+1}(\nu)], \quad (5)$$

where  $a_i(1)$  is the subset index of  $\mathbf{x}_i$  in  $\Omega(1)$ ,  $a_{i-1}(2)$  is the index of  $\mathbf{x}_{i-1}$  in  $\Omega(2)$ , and so on. The number of states in the subset trellis is  $\prod_{k=1}^{\nu} J_k$ . The complexity of an RSSE trellis can be controlled by specifying  $\nu$  parameters  $J_k$ ,  $1 \leq k \leq \nu$ . We define the **configuration** of a subset trellis to be a vector  $\mathbf{J} = [J_1, J_2, \dots, J_\nu]$ . A valid configuration satisfies  $J_1 \geq J_2 \geq \dots \geq J_\nu$ .

The Viterbi algorithm (VA) can be applied to the reduced-state trellis without incurring additional computational complexity. To track the survivor ML state at each time instant, a modified path history is used to store the survivor symbol  $\hat{\mathbf{x}}_{i-1}$  that leads to state  $\mathbf{s}_i$ . Then the actual survivor ML state  $\hat{\mathbf{p}}_i$  is obtained by tracing back  $\nu$  steps in the path history. We say that  $\hat{\mathbf{p}}_i$  is the unique survivor ML state at time  $i$  among all possible  $\mathbf{p}_i$ 's whose corresponding subset state is  $\mathbf{s}_i$ . Error propagation may occur, but its effect is negligible [8], [13].

The underlying idea of RSSE is to drop less likely paths early in the detection process. Since each subset state contains multiple ML states, certain paths will merge earlier in the subset trellis than in the ML trellis. To minimize the performance loss, proper set partitions  $\Omega(k)$  should be selected carefully to guarantee that sufficient distance differences have been accumulated to reliably distinguish between merging paths. For the  $M$ -QAM system, it has been suggested [8] that good performance can generally be obtained by maximizing the minimum intrasubset Euclidean distance for each partition  $\Omega(k)$ ,  $k = 1, \dots, \nu$ . The Ungerboeck set partition tree [14] was shown to have this property and was adopted to make the selection of  $\Omega(k)$ . However, such a set partition tree cannot be directly applied to the  $nHnT$  system because of the ITI. In the following sections, we redefine the way to measure the distance between input symbols, and propose several set partition trees that perform well under different circumstances.

### B. Weighted Sum Subtract Joint Detector (WSSJD)

WSSJD is proposed to resolve the problem of ITI estimation. It works on the transformed model of (1), given by

$$\Lambda_n^{-1} V_n^\top \mathbf{r}_i = V_n^\top \mathbf{y}_i + \Lambda_n^{-1} V_n^\top \boldsymbol{\omega}_i, \quad (6)$$

where  $V_n$  and  $\Lambda_n$  are the induced matrices from the eigen-decomposition of  $A_n$ ,  $A_n = V_n \Lambda_n V_n^\top$ . Since  $A_n$  is a symmetric tridiagonal Toeplitz matrix, it has the property that,  $V_n$  is a constant matrix, and  $\Lambda_n$  is a diagonal matrix whose diagonal elements,  $\lambda_j$ ,  $j = 1, \dots, n$ , are functions of  $\epsilon$ . For example, the 2H2T system has  $A_2 = V_2 \Lambda_2 V_2^\top$ , where

$$\Lambda_2 = \begin{bmatrix} 1+\epsilon & 0 \\ 0 & 1-\epsilon \end{bmatrix}, \quad V_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}. \quad (7)$$

Let  $\mathbf{z}_i = V_n^\top \mathbf{x}_i$ ,  $\bar{\mathbf{r}}_i = \Lambda_n^{-1} V_n^\top \mathbf{r}_i$  and  $\bar{\boldsymbol{\omega}}_i = \Lambda_n^{-1} V_n^\top \boldsymbol{\omega}_i$  be the input, received sample and noise of the transformed system given by (6). Then the system consists of  $n$  parallel sub-channels, each of which has the input-output relationship

$$\bar{r}_i^j = \sum_{k=0}^{\nu} h_k z_{i-k}^j + \bar{\omega}_i^j, \quad j = 1, \dots, n. \quad (8)$$

In the joint trellis constructed for (6), a state  $(\mathbf{z}_{i-\nu}, \dots, \mathbf{z}_{i-1})$  associated with the input  $\mathbf{z}_i$  will have output  $\bar{y}_i^j = \sum_{k=0}^{\nu} h_k z_{i-k}^j$ . Notice that the trellis labels are independent of  $\epsilon$ . Moreover, the noise samples,  $\bar{\omega}_i^j$ , are independent and have different powers,  $E[\bar{\omega}_i^j \bar{\omega}_i^{j\top}] = \sigma^2 (\Lambda_n^{-1})^2$ .

The use of WSSJD is summarized as follows:

- 1) Calculate  $\bar{\mathbf{r}}_i$  by  $\bar{\mathbf{r}}_i = \Lambda_n^{-1} V_n^\top \mathbf{r}_i$ .
- 2) To retain the ML property, the branch metrics are weighted when applying the Viterbi algorithm (VA),

$$m(\mathbf{s}_{i-1}, \mathbf{s}_i) = \sum_{j=1}^n \lambda_j^2 (\bar{r}_i^j - \bar{y}_i^j)^2. \quad (9)$$

WSSJD can work with a gain loop structure which adaptively estimates  $\epsilon$ . The new estimate is then fed back to WSSJD to update the weights in calculating the branch metric. The discussion of ITI estimation is beyond the scope of this

paper; therefore, we assume that  $\epsilon$  is known. For more details on WSSJD, the reader is referred to Part I.

Throughout this paper, the complexity reduction techniques are developed based on WSSJD. This is motivated by the fact that WSSJD has ML-equivalent performance. Moreover, the coordinate transformations in WSSJD lead to a better measure of the distance between input symbols, which plays an important role in designing the set partition tree. Further, as we will see, the structure of parallel channels can provide additional complexity reduction in selecting survivor paths. We emphasize, however, that the RSSE techniques described here are also applicable to the standard MHMT ML detector trellis. Henceforth, with a slight abuse of terminology, when we say the ‘‘ML trellis’’, we refer to the WSSJD trellis.

### III. SET PARTITION TREE FOR 2H2T SYSTEM

We first consider the symmetric 2H2T system

$$\begin{bmatrix} r_i^1 \\ r_i^2 \end{bmatrix} = \begin{bmatrix} 1 & \epsilon \\ \epsilon & 1 \end{bmatrix} \begin{bmatrix} y_i^1 \\ y_i^2 \end{bmatrix} + \begin{bmatrix} \omega_i^1 \\ \omega_i^2 \end{bmatrix}. \quad (10)$$

which is also studied in [5] and [6]. After the WSSJD transformation, two parallel sub-channels are formed, given by

$$\bar{r}_i^1 = \sum_{k=0}^{\nu} h_k z_{i-k}^1 + \bar{\omega}_i^1, \quad \bar{r}_i^2 = \sum_{k=0}^{\nu} h_k z_{i-k}^2 + \bar{\omega}_i^2 \quad (11)$$

where

$$\begin{bmatrix} z_i^1 \\ z_i^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_i^1 \\ x_i^2 \end{bmatrix} \quad (12)$$

$$\begin{bmatrix} \bar{r}_i^1 \\ \bar{r}_i^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{1+\epsilon} & 0 \\ 0 & \frac{1}{1-\epsilon} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} r_i^1 \\ r_i^2 \end{bmatrix} \quad (13)$$

$$\begin{bmatrix} \bar{\omega}_i^1 \\ \bar{\omega}_i^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{1+\epsilon} & 0 \\ 0 & \frac{1}{1-\epsilon} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \omega_i^1 \\ \omega_i^2 \end{bmatrix}. \quad (14)$$

are the new input symbol, received symbol, and noise component of the transformed system (11), respectively. The noise samples satisfy  $\bar{\omega}_i^1 \sim \mathcal{N}(0, \frac{2\sigma^2}{(1+\epsilon)^2})$ ,  $\bar{\omega}_i^2 \sim \mathcal{N}(0, \frac{2\sigma^2}{(1-\epsilon)^2})$ .

In this new system,  $z^1(D)$  and  $z^2(D)$  are transmitted separately through  $h(D)$ . If we treat  $z_i^1$  and  $z_i^2$  as the real and imaginary components of a complex symbol, the resulting system is QAM-like, where the only difference is that the two subchannels have different signal-to-noise ratios (SNRs). Considering this dimensional asymmetry, we define the **effective symbol pair distance** (ESPD) between two input symbols  $z_i$  and  $\tilde{z}_i$  as

$$d_e^2(z_i, \tilde{z}_i) = \frac{(1+\epsilon)^2}{2} (z_i^1 - \tilde{z}_i^1)^2 + \frac{(1-\epsilon)^2}{2} (z_i^2 - \tilde{z}_i^2)^2. \quad (15)$$

The ESPDs between different pairs of inputs are listed in Table I. Notice that the ESPDs,  $\Delta_1^2$ ,  $\Delta_2^2$ , and  $\Delta_3^2$ , show different monotonicity behavior when  $\epsilon$  changes. For  $\epsilon \in [0, 0.5]$ ,  $\Delta_1^2$  and  $\Delta_3^2$  are increasing functions, while  $\Delta_2^2$  decreases. The changes in ESPDs affect the performance of a reduced-state trellis. Therefore, even with the same subset trellis configuration, the RSSE performs differently at various ITI levels.

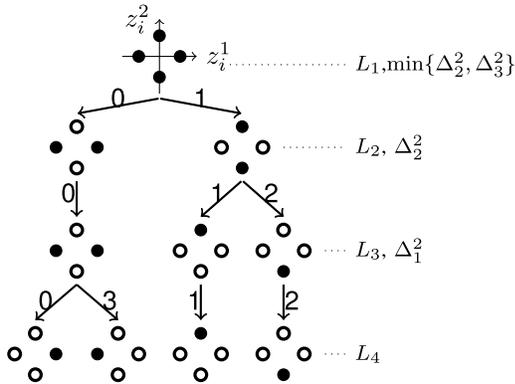


Fig. 1. The set partition tree constructed for 2H2T system. The horizontal and vertical axes in the constellation correspond to  $z_i^1$  and  $z_i^2$  dimension, respectively. This tree contains 4 levels,  $\{L_1, L_2, L_3, L_4\}$ , each of which is a set partition of the constellation. The minimum ESPD on each level is specified on the right side. The number labeled on each branch is the index of the subset in the corresponding set partition.

TABLE I  
THE ESPDS BETWEEN DIFFERENT INPUT SYMBOLS

$(z_i, \tilde{z}_i)$	$d_c(z_i, \tilde{z}_i)$
$([+2, 0], [-2, 0])$	$\Delta_1^2 = 8(1 + \epsilon)^2$
$([0, +2], [0, -2])$	$\Delta_2^2 = 8(1 - \epsilon)^2$
$([+2, 0], [0, +2])$	$\Delta_3^2 = 4(1 + \epsilon^2)$
$([+2, 0], [0, -2])$	
$([-2, 0], [0, +2])$	
$([-2, 0], [0, -2])$	

Based on Table I, we propose a set partition tree shown in Fig. 1. The minimum intrasubset ESPD increases from the top level to the bottom level. Compared to the Ungerboeck set partition tree, the additional level  $L_3$  comes from the asymmetric distance measure in the  $z^1$  and  $z^2$  dimensions, and it provides better flexibility in performance/complexity tradeoff. We emphasize that, although the proposed set partition tree is motivated by the WSSJD transformation, it can be implemented with the standard ML detectors.

The subset trellis is constructed by choosing  $\Omega(k)$  from the levels of the set partition tree for each  $k = 1, \dots, v$ , and to guarantee a well-defined trellis structure,  $\Omega(k)$  should always be at the same level or at a higher level than  $\Omega(k-1)$ . During the detection process, only one ML state can survive inside each subset state at each time slot. Consider the example given in Fig. 2 and assume  $\epsilon = 0.1$ . Once the survivor state  $p_i$  is decided, the output labels are also determined. A look-up table can be stored to facilitate the process of finding the corresponding output labels given the survivor ML states. Assume the transformed received signals are  $\tilde{\mathbf{r}} = (\tilde{r}^1, \tilde{r}^2) = (5, 3)$ . The metric comparison shows that the subset state  $[0, 1]$  with survivor ML state  $[(+2, 0), (0, +2)]$  gives the smallest path metric, so the survivor ML state of  $s_{n+1} = [0, 0]$  is updated to be  $[(+2, 0), (+2, 0)]$ , which will be used in the next time slot.

For a configuration with  $J_1 < 4$ , the subset trellis contains parallel branches. A pre-selection between the parallel branches is required during detection. Due to the symmetric property of WSSJD trellis labels, this pre-selection can be done without explicitly calculating the branch metric, for

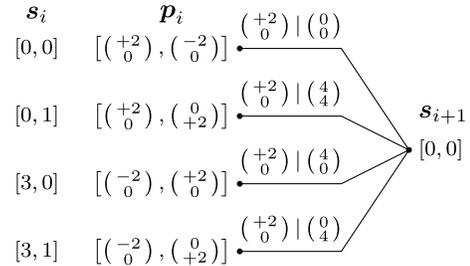
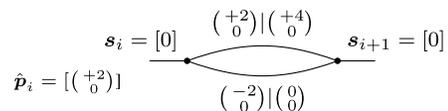
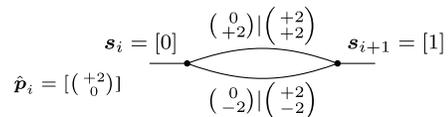


Fig. 2. An illustration of detection on subset trellis. The subset trellis is constructed for PR2 channel  $1 + 2D + D^2$ , with configuration  $[4, 2]$ . The left columns list the subset states  $s_i$  and their associated survivor ML states  $p_i$ . The label on each branch is the channel input|output. All of the branches terminate at subset state  $s_{i+1} = [0, 0]$ .



(a) parallel branches from subset state 0 to 0



(b) parallel branches from subset state 0 to 1

Fig. 3. Sample parallel branches for subset trellis with 2 states constructed for channel  $1 + D$ .

$J_1 = 2$  or 3. For instance, consider the two scenarios illustrated in Fig. 3. In both cases, the survivor ML state at the starting stage is assumed to be  $\hat{p}_i = [(+2, 0)]$ . The input and output labels are marked on the branches. In Fig. 3(a), both the input symbols  $(+2, 0)$  and  $(-2, 0)$  lead the paths to subset state 0. Notice that the input symbols  $(+2, 0)$  and  $(-2, 0)$  have the same value in the  $z^2$  dimension, producing the same output on the subtract channel. Instead of calculating metrics from equation (9), the pre-selection performs a thresholding on the sum channel output and makes the decision. In this example, the threshold is  $+2$ , obtained by averaging  $+4$  and  $0$ . If  $\tilde{r}_i^1 > +2$ , the strategy is to pick  $(+2, 0)$  as the survivor symbol, while for the case  $\tilde{r}_i^1 < +2$ ,  $(-2, 0)$  should be the survivor. Similarly for another case shown in Fig. 3(b), the thresholding is conducted on the subtract channel output, since the two input symbols produce the same output in the sum channel. By comparing  $\tilde{r}_i^2$  with the threshold  $0$ , the detector picks  $(0, +2)$  if  $\tilde{r}_i^2 > 0$ , or  $(0, -2)$  if  $\tilde{r}_i^2 < 0$ . This symmetry property renders the WSSJD formulation preferable over the traditional ML detector.

#### IV. PERFORMANCE OF RSSE ON 2H2T SYSTEM

We examine the RSSE performance on various subset trellises constructed from the proposed set partition tree. Several types of channels at different ITI levels are considered. The SNR is defined as

$$\text{SNR(dB)} = 10 \log \frac{\|h(D)\|^2}{2\sigma^2} \quad (16)$$

where  $\|h(D)\|^2 = \sum_i h_i^2$ .

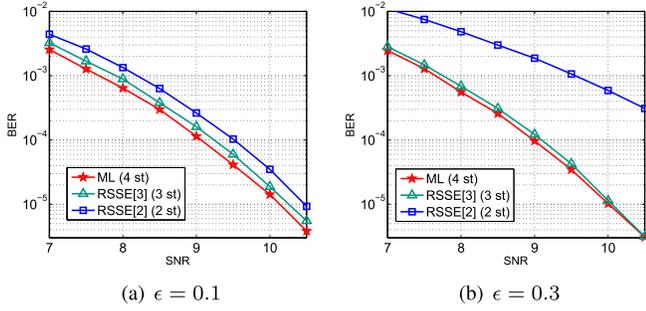


Fig. 4. Performance comparison between RSSE and ML detector on dicode channel at different ITI levels. The legend shows the RSSE subset trellis configuration and the corresponding number of trellis states.

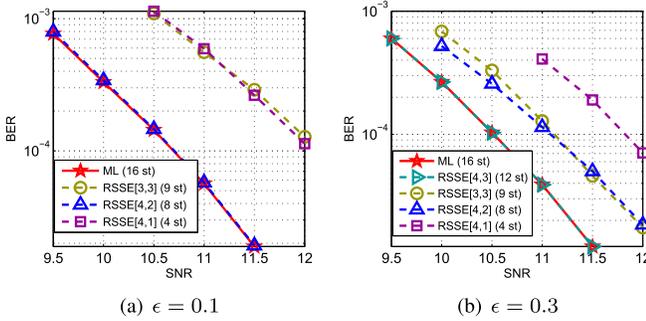


Fig. 5. Performance comparison between RSSE and ML detector on PR2 channel at different ITI levels. The legend shows the RSSE subset trellis configuration and the corresponding number of trellis states.

### A. Dicode Channel

This simple example helps us understand how the pre-selection between parallel branches affects the system. Although early-merging happens at every time step, it does not seriously degrade the performance. From Fig. 4 we see that the performance loss of the 3-state subset trellis is less than 0.1dB. Moreover, the 3-state RSSE has better performance at the higher ITI level ( $\epsilon = 0.3$ ), while the 2-state RSSE performs better at the lower ITI level ( $\epsilon = 0.1$ ). In Section V-A, we explain this observation by analyzing the length-1 error events.

### B. Channel with Higher Memory

Higher channel memory provides more flexibility in constructing the subset trellis. PR2 and EPR4 are two commonly used PR targets to approximate magnetic recording channels. For the PR2 channel,  $h(D) = 1 + 2D + D^2$ , the bit error rate (BER) performance as a function of SNR at different ITI levels is plotted in Fig. 5. The comparison between Fig. 5(a) and Fig. 5(b) shows that even using the same subset trellis, RSSE performs differently at different ITI levels, and its performance correlates with the minimum intrasubset ESPDs of the set partitions configured in the subset trellis. At a low ITI level ( $\epsilon = 0.1$ ), the performance of RSSE on the [4, 2] subset trellis coincides with that of the ML detector. The performance of trellis [4, 3] is not plotted, but can be predicted to be close to the ML curve. The other two trellises, [4, 1] and [3, 3], lose approximately 1.25dB. When the ITI level becomes higher ( $\epsilon = 0.3$ ), the subset trellis [4, 2] cannot provide reliable early path merging because the minimum

TABLE II  
THE SNR LOSS OF DIFFERENT SUBSET TRELLIS CONFIGURATIONS TO ACHIEVE BER =  $10^{-4}$ .

(a) PR2 channel

Trellises	$\epsilon$			
	0.1	0.2	0.3	0.4
RSSE [4, 1]	1.25dB	1.3dB	1.35dB	1.2dB
RSSE [4, 2]	$\ll 0.1$ dB	0.15dB	0.6dB	1.1dB
RSSE [3, 3]	1.4dB	0.9dB	0.6dB	0.2dB
RSSE [4, 3]	$\ll 0.1$ dB			

(b) EPR4 channel

Trellises	$\epsilon$			
	0.1	0.2	0.3	0.4
RSSE [4, 3, 3]	0.1dB	0.1dB	0.05dB	$\ll 0.1$ dB
RSSE [4, 4, 2]	$\ll 0.1$ dB	0.15dB	0.7dB	$> 1$ dB
RSSE [4, 3, 2]	0.1dB	0.25dB	0.7dB	$> 1$ dB
RSSE [4, 2, 2]	0.3dB	$> 1$ dB		
RSSE [3, 3, 3]	$> 1$ dB	0.7dB	0.4dB	0.05dB
RSSE [4, 3, 1]	$> 1$ dB			

intrasubset ESPD  $\Delta_2^2$  in  $\Omega(3) = L_2$  is substantially reduced. However, [4, 3] can achieve near-ML performance. The trellis [4, 1] still has a 1.35dB loss, while the increase of  $\Delta_1^2$  brings [3, 3] closer to the ML performance.

In [12] we also present simulation results for the EPR4 channel,  $h(D) = 1 + D - D^2 - D^3$ . We show that the RSSE[4, 3, 2] and RSSE[4, 3, 3] trellises can provide near-ML performance (less than 0.1dB loss) at  $\epsilon = 0.1$  and  $\epsilon = 0.3$ , respectively.

Tables II summarize the performance loss in dB for several subset trellis configurations compared to an ML detector at BER =  $10^{-4}$  on the PR2 and EPR4 channels, respectively. Several conclusions can be drawn from these tables. First, a trellis with fewer states does not necessarily have worse performance than one with more states. For example, in Table II(b) for the EPR4 channel, when  $\epsilon = 0.1$ , the [4, 4, 2] configuration with 32 states outperforms the [4, 3, 3] configuration with 36 states. Second, the performance of a configuration may change drastically at different ITI levels. One example is the [4, 4, 2] trellis, which essentially achieves ML performance at  $\epsilon = 0.1$ , but loses over 1dB for  $\epsilon = 0.4$ . Finally, not all configurations suffer further performance losses at higher ITI. It is interesting to observe that the RSSE [3, 3, 3] tellis with parallel branches can have near-optimal performance at  $\epsilon = 0.4$ . Therefore, the pre-selection between parallel branches at every stage is quite reliable. In Section V we give an explanation of these observations from the point of view of error event analysis.

### C. Minimum Phase Channels

Minimum phase channels can better model the real channel on a disk drive. Assume the transition response of a perpendicular magnetic recording (PMR) disk is  $s(t) = V_{\max} \tanh(\frac{2t}{0.579\pi\delta})$ , where  $V_{\max}$  is the writing voltage and  $\delta$  indicates the linear density on one data track. Using the whitened matched filter structure in [15], we derive two minimum phase channel polynomials: channel 1,

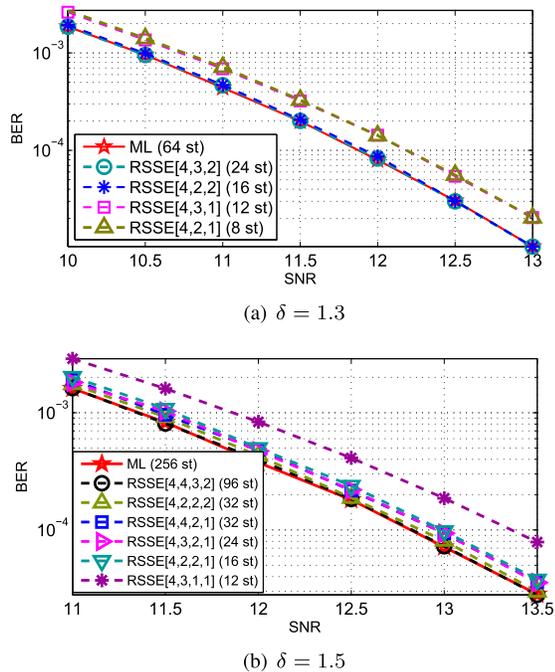


Fig. 6. Performance comparison between RSSE and ML detector on minimum phase channels at  $\epsilon = 0.1$ . The polynomials are (a)  $h(D) = 1 + 1.6D + 1.1D^2 + 0.4D^3$ , (b)  $h(D) = 1 + 1.9D + 1.6D^2 + 0.8D^3 + 0.3D^4$ .

$h(D) = 1 + 1.6D + 1.1D^2 + 0.4D^3$  for  $\delta = 1.3$ , and channel 2,  $h(D) = 1 + 1.9D + 1.6D^2 + 0.8D^3 + 0.3D^4$  for  $\delta = 1.5$ . These are two commonly used densities in current commercial HDDs. Since the minimum phase condition implies that most of the channel energy is distributed over the most recent samples, the early merge in RSSE can be more reliable for these channels compared to linear phase channels, such as PR2 and EPR4. It is interesting to compare channel 1 and EPR4, both of which have memory  $\nu = 3$ . As shown in Fig. 6(a), ML, RSSE[4, 3, 2], and RSSE[4, 2, 2] have essentially identical performance. Therefore, RSSE can achieve near-ML performance with only 16 states, if the [4, 2, 2] trellis is used, whereas the ML detector requires 64 states. The performance of other, more aggressive configurations is also plotted. As can be seen, RSSE with only 8 states can achieve performance that is within 0.3dB of ML detection.

The simulation results for channel 2 are plotted in Fig. 6(b). They show that RSSE[4, 2, 2, 2] with 32 states can achieve near-ML performance. In contrast, the ML trellis requires 256 states. If 0.1dB loss is permissible, the RSSE [4, 2, 2, 1] trellis can be used, reducing the number of states to only 16.

In summary, the simulation results on linear phase channels and minimum phase channels show that RSSE can achieve near optimal performance with significantly reduced number of states. It could potentially substitute the ML detector when the channel interference becomes much more severe.

## V. ERROR EVENT ANALYSIS

We will use error event analysis to study the performance-complexity tradeoff among different subset trellis configurations.

The detector makes errors if the survivor path diverges from the correct one. Let

$$\begin{aligned} e(D) &= [e^1(D), e^2(D)] \\ &= [x^1(D) - \hat{x}^1(D), x^2(D) - \hat{x}^2(D)] \end{aligned} \quad (17)$$

denote an error event of the original 2H2T system, and let

$$\begin{aligned} \bar{e}(D) &= [\bar{e}^1(D), \bar{e}^2(D)] \\ &= [z^1(D) - \hat{z}^1(D), z^2(D) - \hat{z}^2(D)] \end{aligned} \quad (18)$$

be the transformed error event of the WSSJD system. We also use  $e_i = [e_i^1, e_i^2]^\top$  and  $\bar{e}_i = [\bar{e}_i^1, \bar{e}_i^2]^\top$  to represent the original and transformed error symbols at time  $i$ , respectively. It is easy to see that

$$\begin{bmatrix} \bar{e}_i^1 \\ \bar{e}_i^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e_i^1 \\ e_i^2 \end{bmatrix}. \quad (19)$$

It is well-known [15] that at high SNR, the error event probability of a trellis-based detector can be approximated by  $P_e \approx c \cdot Q(\frac{d_{\min}^2}{2\sigma})$ , where  $Q(\cdot)$  is the area under the tail of the standard Gaussian distribution,

$$d_{\min}^2 = \min_{e(D)} d^2(e(D)) \quad (20)$$

is the **minimum distance parameter**, and  $c$  is a coefficient indicating the average number of error events at distance  $d_{\min}^2$ . Due to the exponential nature of the  $Q$  function, the performance comparison between two detectors can be easily conducted by considering their minimum distance parameter. The error events that lead to  $d_{\min}^2$  are the **dominant error events**.

For the WSSJD detector, an effective measure of the distance associated with  $\bar{e}(D)$  is defined by

$$\begin{aligned} d_{\text{W}}^2(\bar{e}(D)) &= \frac{(1+\epsilon)^2}{2} \|\bar{e}^1(D)h(D)\|^2 \\ &\quad + \frac{(1-\epsilon)^2}{2} \|\bar{e}^2(D)h(D)\|^2. \end{aligned} \quad (21)$$

Comparing (15) to (21), we can see that ESPD is proportional to the distance associated with a single error symbol  $\bar{e}_i$ . Recall that the set partition tree is constructed based on ESPDs, and  $\Delta_1^2, \Delta_2^2, \Delta_3^2$  are varying with respect to  $\epsilon$ . Therefore, the minimum distance parameter of the reduced-state trellis configuration also changes with  $\epsilon$ , and its trend can be roughly predicted by analyzing the change of minimum ESPD in each  $\Omega(k)$ . We will give more detailed insights into this behavior in the following subsections. The minimum value of  $d_{\text{W}}^2(\bar{e}(D))$  is abbreviated to  $d_{\min}^2$ , which is the minimum distance parameter of the 2H2T ML detector. It serves as a benchmark for evaluating the performance of the RSSE algorithm.

### A. Parallel Branches

For the subset trellis with parallel branches, early merge happens at every time instant. Ignoring the error propagation effect, we assume that at time  $i$ , both the correct and the estimated sequences are at state  $s_i$ , and  $z_{i-k} = \hat{z}_{i-k}$  for all  $k = 1, \dots, \nu$ . At time  $i+1$ , if  $z_i, \hat{z}_i \in a_i(1)$ , the detector

needs to decide a survivor symbol, and discard the other one. Once the correct symbol is discarded, this wrong decision can not be reversed in the remaining steps. The probability of making a wrong decision in the parallel branch selection is  $Q(\frac{h_0 d(z_i, \hat{z}_i)}{2\sigma})$ , where  $d(z_i, \hat{z}_i)$  is the square-root of ESPD. Let  $E_1$  be the set of all such length-1 error events due to the parallel branches. Then,  $\bar{e}_i \in E_1$  if and only if there exist two inputs  $z_i, \hat{z}_i \in a_i(1)$  such that  $\bar{e}_i = z_i - \hat{z}_i$ . It can be shown that

$$\begin{aligned} d_{\min}^2(E_1) &= \min_{\bar{e}_i \in E_1} \frac{(1+\epsilon)^2}{2} (h_0 \bar{e}_i^1)^2 + \frac{(1-\epsilon)^2}{2} (h_0 \bar{e}_i^2)^2 \\ &= \begin{cases} 8h_0^2(1+\epsilon)^2 = h_0^2 \Delta_1^2 & J_1 = 3 \\ 8h_0^2(1-\epsilon)^2 = h_0^2 \Delta_2^2 & J_1 = 2. \end{cases} \end{aligned} \quad (22)$$

The existence of parallel branches will not significantly degrade the performance if it can achieve the same minimum distance as ML detection, i.e.,  $d_{\min}^2(E_1) \geq d_{\min}^2$ . For the dicode channel,

$$d_{\min}^2 = \begin{cases} 8(1+\epsilon)^2 & \text{if } 0 \leq \epsilon \leq 2 - \sqrt{3} \\ 16(1-\epsilon)^2 & \text{if } 2 - \sqrt{3} \leq \epsilon \leq 1/2. \end{cases} \quad (23)$$

Therefore, for the 3-state subset trellis,  $d_{\min}^2(E_1) \geq d_{\min}^2$  for all  $\epsilon$ , leading to performance close to the ML detector, as shown in Fig. 4. Moreover, as  $\epsilon$  increases,  $d_{\min}^2(E_1)$  becomes much larger than  $d_{\min}^2$ , making the effect of these length-1 error events negligible. So we observe that the BER curve of the 3-state RSSE trellis almost overlaps with that of the ML detector at  $\epsilon = 0.3$ . In contrast to the 3-state trellis, in the 2-state trellis  $d_{\min}^2(E_1) < d_{\min}^2$  for all  $\epsilon$ , resulting in worse performance.

As for the PR2 and EPR4 channels, their  $d_{\min}^2$  is given by

$$d_{\min}^2 = \begin{cases} 16(1+\epsilon^2) & \text{if } 0 \leq \epsilon \leq 2 - \sqrt{3} \\ 32(1-\epsilon)^2 & \text{if } 2 - \sqrt{3} \leq \epsilon \leq 1/2. \end{cases} \quad (24)$$

Consider a subset trellis with  $J_1 = 3$ . For  $\epsilon > \frac{1}{3}$ ,  $d_{\min}^2(E_1)$  is strictly larger than  $d_{\min}^2$ . Therefore the error events in  $E_1$  are not the dominant ones. As shown in Tables II and II, at  $\epsilon = 0.4$ , the RSSE [3, 3] and RSSE [3, 3, 3] configurations perform very close to their corresponding ML detectors, respectively.

### B. Early Merging Condition

We next try to identify longer RSSE error events. Suppose the decoding paths of  $[z^1(D), z^2(D)]$  and  $[\hat{z}^1(D), \hat{z}^2(D)]$  are merged at times  $i_1$  and  $i_2$  and unmerged in between. Let  $E$  denote the set of all error events ending at time  $i_2$ , where the starting position  $i_1$  is arbitrary. According to [8], an error event  $\bar{e}(D) \in E$  if and only if the following hold.

- 1)  $\bar{e}_{i_1}$  is non-zero.
- 2) The last  $\nu$  elements,  $[\bar{e}_{i_2-\nu}, \dots, \bar{e}_{i_2-1}]$ , should satisfy the **merging condition**, i.e.,  $\bar{e}_{i_2-k} = z_{i_2-k} - \hat{z}_{i_2-k}$  where  $z_{i_2-k}$  and  $\hat{z}_{i_2-k}$  belong to the same subset in the partition  $\Omega(k)$  for all  $k = 1, \dots, \nu$ .
- 3) No earlier  $\nu$  elements satisfy the merging condition.

In MLSE, the merging condition requires  $\bar{e}_{i_2-k} = 0$  for  $k = 1, \dots, \nu$ . However, this is not the case in RSSE. We call

TABLE III  
ERROR SYMBOLS BY INDEX

index	$[\bar{e}_i^1, \bar{e}_i^2]$	$[e_i^1, e_i^2]$
0	[0, 0]	[0, 0]
1	[4, 0]	[2, 2]
2	[-4, 0]	[-2, -2]
3	[0, 4]	[2, -2]
4	[0, -4]	[-2, 2]
5	[2, 2]	[2, 0]
6	[-2, -2]	[-2, 0]
7	[2, -2]	[0, 2]
8	[-2, 2]	[0, -2]

the error events  $\bar{e}(D) \in E$  whose last  $\nu$  elements are not all zero the **early merged error events**, denoted by  $E^r$ . Clearly  $E_1 \subseteq E^r$ .

We now present a necessary and sufficient condition for an error event  $\bar{e}(D)$  to belong to  $E^r$ . We refer to this as the **early merging condition**. We first introduce some terminology.

For a partition  $\Omega$  of the input constellation, the set of **intrasubset errors**, denoted by  $\mathcal{E}_a(\Omega)$ , is a collection of error symbols such that if there exist two input symbols  $z, \hat{z}$  satisfying the condition that  $\bar{e} = z - \hat{z}$  and  $z, \hat{z}$  belong to the same subset in  $\Omega$ , then  $\bar{e} \in \mathcal{E}_a(\Omega)$ . Similarly, the set of **intersubset errors**, denoted by  $\mathcal{E}_b(\Omega)$ , is a collection of error symbols such that if there exist two inputs  $z, \hat{z}$  satisfying the condition that  $\bar{e} = z - \hat{z}$  and  $z, \hat{z}$  belong to two different subsets in  $\Omega$ , then  $\bar{e} \in \mathcal{E}_b(\Omega)$ .

The following proposition gives the relationship between the intrasubset errors and intersubset errors for the set partition tree in Fig. 1. For convenience, the error symbols are indexed by the digits shown in Table III.

*Proposition 1:* For the proposed set partition tree in Fig. 1,  $\mathcal{E}_a(L_i) \cap \mathcal{E}_b(L_i) = \emptyset$  for  $i = 1, 2, 3, 4$ .

*Proof:* We prove the claim by enumeration.

- 1) If  $\Omega = L_1$ , all error symbols are intrasubset errors since there is only one subset.
- 2) If  $\Omega = L_2$ ,  $\mathcal{E}_a(L_2) = \{0, 1, 2, 3, 4\}$ ,  $\mathcal{E}_b(L_2) = \{5, 6, 7, 8\}$ .
- 3) If  $\Omega = L_3$ ,  $\mathcal{E}_a(L_3) = \{0, 1, 2\}$ ,  $\mathcal{E}_b(L_3) = \{3, 4, 5, 6, 7, 8\}$ .
- 4) If  $\Omega = L_4$ , all non-zero error symbols are intersubset errors, so  $\mathcal{E}_a(L_4) = \{0\}$ .  $\square$

We now present the early merging condition.

*Proposition 2:* (Early merging condition)

An error event  $\bar{e}(D) \in E^r$  if and only if the last  $\nu$  elements are not all zero symbols, and satisfy  $\bar{e}_{i_2-k} \in \mathcal{E}_a(\Omega(k))$  for all  $k = 1, \dots, \nu$ , and no previous  $\nu$ -tuple satisfies the condition.

*Proof:* Given  $\bar{e}(D) \in E^r$ , it is straightforward from the definition of “merging condition” that the last  $\nu$  elements must be intrasubset error symbols in the corresponding partition  $\Omega(k)$ . On the other hand, if  $\bar{e}_{i_2-k} \in \mathcal{E}_a(\Omega(k))$  for all  $k = 1, \dots, \nu$ , by Proposition 1, the sequences that produce  $\bar{e}(D)$  must satisfy that  $z_{i_2-k}$  and  $\hat{z}_{i_2-k}$  belong to the same subset in  $\Omega(k)$ . Therefore the decoding paths are merged at  $i_2$ , and  $\bar{e}(D) \in E^r$ .  $\square$

TABLE IV  
THE DOMINANT RSSE ERROR EVENTS FOR CHANNEL [1, 1.6, 1.1, 0.4]

	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.3$	$\epsilon = 0.4$
	$d_{\min}^2 = 9.1304$	$d_{\min}^2 = 9.4016$	$d_{\min}^2 = 8.8592$	$d_{\min}^2 = 6.5088$
RSSE[3, 3, 3] (27 st)	[1]/9.6800	[5 2 1 2]/10.7168	[5 2 1 2]/10.0028	[5 2 1 2]/9.5472
RSSE[3, 3, 2] (18 st)	[1]/9.6800	[5 2 1 2]/10.7168	[3 4 0 0]/8.2320	[3 4 0 0]/6.0480
RSSE[4, 3, 2] (24 st)	[3 4 0 0]/13.6080	[3 4 0 0]/10.7520	[3 4 0 0]/8.2320	[3 4 0 0]/6.0480
RSSE[4, 2, 2] (16 st)	[3 4 0]/10.4328	[3 4 0]/8.2432	[3 4 0]/6.3112	[3 4 0]/4.6368
RSSE[3, 2, 2] (12 st)	[1]/9.6800	[3 4 0]/8.2432	[3 4 0]/6.3112	[3 4 0]/4.6368
RSSE[4, 2, 1] (8 st)	[5 6 0 0]/8.4840	[5 6 0 0]/8.7360	[3 4 0]/6.3112	[3 4 0]/4.6368

*Remark 1:* Notice that Proposition 1 is also true for the QAM Ungerboeck set partition tree. So Proposition 2 also applies to the original RSSE formulation.

*Remark 2:* The single track error events are not affected by the RSSE algorithm if  $J_i > 1$  for all  $i = 1, \dots, \nu$ .

Assume  $\bar{e}(D) \in E^r$  and starts from  $k_1$ . The distance parameter of  $\bar{e}(D)$  is given by

$$d_r^2(\bar{e}(D)) = \frac{(1+\epsilon)^2}{2} \sum_{i=i_1}^{i_2} \left( \sum_{k=0}^{\nu} h_k \bar{e}_{i-k}^1 \right)^2 + \frac{(1-\epsilon)^2}{2} \sum_{i=i_1}^{i_2} \left( \sum_{k=0}^{\nu} h_k \bar{e}_{i-k}^2 \right)^2. \quad (25)$$

The distance parameter measured by (25) is always smaller than or equal to that measured by (21) [8]. The possible reduction represents the price paid for using the reduced-state trellis. An example is given to illustrate the difference.

*Example 1:* Consider the PR2 channel. Assume that  $\epsilon = 0.1$ . Assume a single error  $\bar{e}_i = [4, 0]^T$  happens at time  $i$ . In ML detection, the paths remerge at time  $i + 2$ , and the distance parameter contributed by  $\bar{e}_i$  is  $\frac{(1+0.1)^2}{2} \cdot (4^2 + 8^2 + 4^2) = 58.08$ . However, if RSSE[4, 3] is used, the paths will be early merged at time  $i + 1$ , since  $\bar{e}_i \in \mathcal{E}_a(\Omega(2))$ . Therefore the distance parameter of this error event is reduced to  $\frac{(1+0.1)^2}{2} \cdot (4^2 + 8^2) = 48.4$  in RSSE[4, 3].

Let

$$d_{\min}^2(E^r) = \min_{\bar{e}(D) \in E^r} d_r^2(\bar{e}(D)). \quad (26)$$

The early merged error events  $\bar{e}^*(D)$  that achieve (26) are referred to as the dominant RSSE error events. To obtain good performance, it is essential that  $d_{\min}^2(E^r) \geq d_{\min}^2$ .

### C. Error State Diagram

An error state diagram can be employed to search for the minimum distance and enumerate the dominant error events. Consider a labeled directed graph  $G = [V, E]$ . The vertex set  $V$  is the collection of all possible error states  $[\bar{e}_{i-1}, \dots, \bar{e}_{i-\nu}]$ , so  $|V| = 9^\nu$ . A state that satisfies the merging condition is called a **merging state**. For ML detection, the all-zero state is the only merging state, while for RSSE, additional early merging states are those which satisfy the early merging condition. If  $\mathcal{T}$  denotes the set of merging

states, then  $|\mathcal{T}| = \prod_{k=1}^{\nu} |\mathcal{E}_a(\Omega(k))|$ , which depends on the trellis configuration. An edge  $(u, v) \in E$  starts from initial state  $u = [\bar{e}_{i-1}, \dots, \bar{e}_{i-\nu}]$  and ends in terminal state  $v = [\bar{e}_i, \dots, \bar{e}_{i-\nu+1}]$ , with input/output label  $\bar{e}_i / \mathcal{L}_{\text{out}}$ . Here

$$\mathcal{L}_{\text{out}} = \frac{(1+\epsilon)^2}{2} \left( \sum_{k=0}^{\nu} h_k \bar{e}_{i-k}^1 \right)^2 + \frac{(1-\epsilon)^2}{2} \left( \sum_{k=0}^{\nu} h_k \bar{e}_{i-k}^2 \right)^2. \quad (27)$$

Notice that all the merging states except the all-zero state are sink nodes, which have no outgoing edges. A path starting from the all-zero state and terminating at the merging state defines a closed error event, and the sum of the output labels of all edges in the path gives the distance parameter of this error event. A closed error event that ends at a non-zero merging state is an early merged error event. As proposed in [16], a depth-first algorithm can be used to find all the error events that lead to a distance parameter smaller than a given threshold.

We are interested in the dominant RSSE error events, i.e., the error events that end at non-zero merging states and produce the distance  $d_{\min}^2(E^r)$ . Table IV summarizes the dominant RSSE error events and their induced distances for several trellis configurations for the minimum phase channel  $h(D) = 1 + 1.6D + 1.1D^2 + 0.4D^3$ . We simplify the table as follows: if  $d_{\min}^2(E^r) \geq d_{\min}^2$ , we only list the early merged error events that lead to  $d_{\min}^2(E^r)$ ; if  $d_{\min}^2(E^r) < d_{\min}^2$ , we list all the early merged error events whose distance parameters are smaller than or equal to  $d_{\min}^2$ . The table is also simplified by considering the symmetry of the error events, i.e.,  $\pm(e^1(D), e^2(D))$  will produce the same distance parameter, and if the error events of track 1 and 2 are switched, the distance remains the same. So we group them together and only list the one whose first error symbol has a positive  $e_i^1$  component. As shown in Table IV, the early merged error events in the RSSE [3, 3, 3] trellis always have distance parameter greater than  $d_{\min}^2$ , under all ITI levels. Specifically, when  $\epsilon = 0.1$ ,  $E_1$  are the dominant RSSE error events. As  $\epsilon$  increases,  $d_{\min}^2(E_1)$ , which is proportional to  $\Delta_1^2$ , also increases, and [5, 2, 1, 2] becomes the dominant one. For the RSSE [4, 3, 2] trellis, the error event [3, 4, 0, 0] is dominant, and its distance parameter decreases as  $\epsilon$  increases. In particular, for  $\epsilon = 0.3$  and  $0.4$ , its distance is strictly less than  $d_{\min}^2$ , so it can be predicted that RSSE [4, 3, 2] suffers greater performance loss compared to the ML detector at high ITI levels. One way to avoid this performance loss is

TABLE V  
THE DOMINANT RSSE ERROR EVENTS FOR PR2 CHANNEL

	$\epsilon = 0.1$	$\epsilon = 0.2$	$\epsilon = 0.3$	$\epsilon = 0.4$
	$d_{\min}^2 = 16.16$	$d_{\min}^2 = 16.64$	$d_{\min}^2 = 15.68$	$d_{\min}^2 = 11.52$
RSSE[4, 3] (12 st)	$[5 (2 1)^\infty 0]/24.24$ $[5 (2 1)^\infty 2 0]/24.24$	$[5 (2 1)^\infty 0]/24.96$ $[5 (2 1)^\infty 2 0]/24.96$	$[5 (2 1)^\infty 0]/26.16$ $[5 (2 1)^\infty 2 0]/26.16$	$[5 (2 1)^\infty 0]/27.84$ $[5 (2 1)^\infty 2 0]/27.84$
RSSE[4, 2] (8 st)	$[(3 4)^\infty 0]/19.44$ $[(3 4)^\infty 3 0]/19.44$	$[(3 4)^\infty 0]/15.36$ $[(3 4)^\infty 3 0]/15.36$	$[(3 4)^\infty 0]/11.76$ $[(3 4)^\infty 3 0]/11.76$	$[(3 4)^\infty 0]/8.64$ $[(3 4)^\infty 3 0]/8.64$
RSSE[3, 3] (9 st)	$[1]/9.68$ $[5 2 1]/14.56$ $[(5 6)^\infty 0 2]/16.16$ $[(5 6)^\infty 5 0 1]/16.16$ $[(5 6)^\infty 2 1]/16.16$ $[(5 6)^\infty 5 1 2]/16.16$	$[1]/11.52$ $[5 2 1]/13.44$ $[(5 6)^\infty 0 2]/16.64$ $[(5 6)^\infty 5 0 1]/16.64$ $[(5 6)^\infty 2 1]/16.64$ $[(5 6)^\infty 5 1 2]/16.64$	$[1]/13.52$ $[5 2 1]/12.64$	$[5 2 1]/12.16$

to sacrifice complexity reduction and use RSSE[4, 3, 3] which prevents the error event [3, 4, 0, 0] from being early merged. RSSE [4, 2, 2] has near-optimal performance at  $\epsilon = 0.1$  and performs much worse when  $\epsilon \geq 0.2$ . A more aggressive configuration, RSSE[4, 2, 1], cannot guarantee near-optimal performance since  $d_{\min}^2(E^r)$  is always smaller than  $d_{\min}^2$ . Therefore, to retain near optimal performance as well as reduce complexity, we may use RSSE[3, 2, 2] at  $\epsilon = 0.1$ , RSSE[3, 3, 2] at  $\epsilon = 0.2$ , RSSE[3, 3, 3] at  $\epsilon = 0.3$  and 0.4.

For the PR2 and EPR4 channel, the error state diagrams contain zero cycles, leading to infinite recursive loops in the error event search. A zero cycle is a path that starts and ends at the same state, and accumulates zero path metric. The number of zero cycles depends on the reduced-state trellis configuration. In Examples 2 and 3, we summarize the zero cycles for PR2 and EPR4 channels. We follow the notations in [16] and let  $(e_1, \dots, e_k)^\infty$  represent an infinite periodic sequence with repeated pattern  $e_1, \dots, e_k$ . Notice that a periodic sequence of the shifted pattern  $(e_i, \dots, e_k, e_1, \dots, e_{i-1})^\infty$  is equivalent to  $(e_1, \dots, e_k)^\infty$ .

*Example 2:* For PR2 channel, if the ML detector is used, the zero cycles are  $0^\infty$ ,  $(1, 2)^\infty$ ,  $(3, 4)^\infty$ ,  $(5, 6)^\infty$ ,  $(7, 8)^\infty$ . If RSSE[3, 3] is used, both [2, 1] and [1, 2] becomes merging states, therefore  $(1, 2)^\infty$  will not be a zero cycle, while other zero cycles still exist.

*Example 3:* The zero cycles for the ML detector on EPR4 are  $(0)^\infty, \pm(0, 1)^\infty, \pm(0, 3), \pm(0, 5), \pm(0, 7), \pm(1)^\infty, \pm(1, 2)^\infty, \pm(1, 3)^\infty, \pm(1, 4)^\infty, \pm(1, 5)^\infty, \pm(1, 6)^\infty, \pm(1, 7)^\infty, \pm(1, 8)^\infty, \pm(3)^\infty, \pm(3, 4)^\infty, \pm(3, 5)^\infty, \pm(3, 6)^\infty, \pm(3, 7)^\infty, \pm(3, 8)^\infty, \pm(5)^\infty, \pm(5, 6)^\infty, \pm(5, 7)^\infty, \pm(5, 8)^\infty, \pm(7)^\infty, \pm(7, 8)^\infty$ . Here  $-(\cdot)^\infty$  represents taking the additive inverse of all symbols inside  $(\cdot)$ .

*Remark 3:* The zero cycles do not intersect, so each state can only be visited by at most one zero cycle. We use  $\gamma(s)$  to denote the zero cycle which starts and ends at state  $u$ , and  $\gamma(u, v)$  to be the fragment of the zero cycle from state  $u$  to  $v$ . By an abuse of notation, we also use  $\gamma(u, v)$  to represent the sequence of input labels on the fragment. The meaning will be clear according to the context.

Let  $\mathcal{Z}$  denote the collection of all the states visited by zero-cycles, and let  $\mathcal{T}$  be the set of all merging states. A two-step algorithm introduced in [16] can be used to search for the dominant error events, with a slight modification that considers the additional early merging states in the RSSE trellis. The procedure is summarized below.

- 1) Given a threshold  $\mathcal{D}$ , apply the depth-first search algorithm to search for all the error fragments, whose path metric is no bigger than  $\mathcal{D}$ , and that start from some state  $u \in \mathcal{Z}$  and end up at some state  $v \in \mathcal{Z} \cup \mathcal{T}$  without having visited  $\mathcal{Z} \cup \mathcal{T}$  in between. The path metric of such an error fragment is denoted as  $d^2(\bar{e}(u, v))$ .
- 2) Construct a new graph  $F$  whose vertices are the states in  $\mathcal{Z} \cup \mathcal{T}$ . The edges in  $F$  are found as follows. If there is an error fragment  $\bar{e}(u, v)$  starting from state  $u$  and ending up at state  $v$ , then for each state  $v' \in \gamma(v)$ , there is an edge from state  $u$  to  $v'$ . The input label of the edge is  $\bar{e}(u, v) + \gamma(v, v')$ , and the output label is  $d^2(\bar{e}(u, v))$ , since the path metric from  $v$  to  $v'$  is zero. Parallel edges are allowed.
- 3) The same depth-first search on  $F$  can be used to search for and list all the closed error events whose distance parameters are less than  $\mathcal{D}$ .

Tables V and VI list the dominant RSSE error events for several trellis configurations on the PR2 and EPR4 channels, respectively. They are constructed in the same manner as Table IV. The tables show a good match with the simulation results in Tables II.

## VI. ASYMMETRIC 2H2T SYSTEM

The asymmetric 2H2T system is worth consideration because of its practical relevance. In this model, the ITI levels sensed by the two heads are different, i.e.,

$$\begin{bmatrix} r_i^1 \\ r_i^2 \end{bmatrix} = \begin{bmatrix} 1 & \epsilon - \Delta\epsilon \\ \epsilon + \Delta\epsilon & 1 \end{bmatrix} \begin{bmatrix} y_i^1 \\ y_i^2 \end{bmatrix} + \begin{bmatrix} \omega_i^1 \\ \omega_i^2 \end{bmatrix}. \quad (28)$$

Without loss of generality, we assume  $0 \leq \Delta\epsilon \leq \epsilon$ .

As in the discussion of the symmetric system, we analyze RSSE on the asymmetric channel by considering the transformed system. After the same coordinate transformation that

TABLE VI  
THE DOMINANT RSSE ERROR EVENTS FOR EPR4 CHANNEL

	$\epsilon = 0.1$ $d_{\min}^2 = 16.16$	$\epsilon = 0.2$ $d_{\min}^2 = 16.64$	$\epsilon = 0.3$ $d_{\min}^2 = 15.68$	$\epsilon = 0.4$ $d_{\min}^2 = 11.52$
RSSE[4, 3, 3] (36 st)	$[(5\ 0)^\infty\ 1\ 0]/16.16$	$[(5\ 0)^\infty\ 1\ 0]/16.64$	$[(5\ 0)^\infty\ 1\ 0]/17.44$	$[(5\ 0)^\infty\ 1\ 0]/18.56$
RSSE[4, 3, 2] (24 st)	$[(5\ 0)^\infty\ 1\ 0]/16.16$	$[(3\ 0)^\infty\ 0]/15.36$ $[(3\ 4)^\infty\ 3\ 0\ 0]/15.36$ $[3\ 4\ (3\ 4)^\infty\ 0\ 0]/15.36$ $[(5\ 0)^\infty\ 1\ 0]/16.64$	$[(3\ 0)^\infty\ 0]/11.76$ $[(3\ 4)^\infty\ 3\ 0\ 0]/11.76$ $[3\ 4\ (3\ 4)^\infty\ 0\ 0]/11.76$	$[(3\ 0)^\infty\ 0]/8.64$ $[(3\ 4)^\infty\ 3\ 0\ 0]/8.64$ $[3\ 4\ (3\ 4)^\infty\ 0\ 0]/8.64$
RSSE[3, 3, 3] (27 st)	$[1]/9.68$ $[5\ 6\ 1\ 2\ 1]/14.56$ $[5\ 2\ 1\ 2]/16.16$ $[(5\ 0)^\infty\ 0\ 1]/16.16$ $[(5\ 0)^\infty\ 1\ 0]/16.16$ $[(5\ 0)^\infty\ 1\ 2]/16.16$ $[5\ 6\ (5\ 6)^\infty\ 0\ 0\ 2]/16.16$ $[5\ (6\ 5)^\infty\ 0\ 0\ 1]/16.16$ $[5\ 6\ (5\ 6)^\infty\ 1\ 2\ 1]/16.16$ $[5\ (6\ 5)^\infty\ 2\ 1\ 2]/16.16$	$[1]/11.52$ $[5\ 6\ 1\ 2\ 1]/13.44$ $[5\ 2\ 1\ 2]/16.64$ $[(5\ 0)^\infty\ 0\ 1]/16.64$ $[(5\ 0)^\infty\ 1\ 0]/16.64$ $[(5\ 0)^\infty\ 1\ 2]/16.64$ $[5\ 6\ (5\ 6)^\infty\ 0\ 0\ 2]/16.64$ $[5\ (6\ 5)^\infty\ 0\ 0\ 1]/16.64$ $[5\ 6\ (5\ 6)^\infty\ 1\ 2\ 1]/16.64$ $[5\ (6\ 5)^\infty\ 2\ 1\ 2]/16.64$	$[1]/13.52$ $[5\ 6\ 1\ 2\ 1]/12.64$	$[5\ 6\ 1\ 2\ 1]/12.16$

was used before, the asymmetric 2H2T channel becomes

$$\begin{bmatrix} \bar{r}_i^1 \\ \bar{r}_i^2 \end{bmatrix} = \begin{bmatrix} 1 & \frac{\Delta\epsilon}{1+\epsilon} \\ \frac{\Delta\epsilon}{\epsilon-1} & 1 \end{bmatrix} \begin{bmatrix} \bar{y}_i^1 \\ \bar{y}_i^2 \end{bmatrix} + \begin{bmatrix} \bar{\omega}_i^1 \\ \bar{\omega}_i^2 \end{bmatrix}, \quad (29)$$

where  $\bar{y}_i^1 = \sum_{k=0}^v h_k z_{i-k}^1$  and  $\bar{y}_i^2 = \sum_{k=0}^v h_k z_{i-k}^2$ , and  $\bar{r}_i$ ,  $\bar{z}_i$ , and  $\bar{\omega}_i$  are obtained from equation (12), (13), and (14), respectively.

In the asymmetric system, the noiseless channel outputs become

$$f_i^1 = \bar{y}_i^1 + \frac{\Delta\epsilon}{1+\epsilon} \bar{y}_i^2, \quad f_i^2 = \bar{y}_i^2 + \frac{\Delta\epsilon}{\epsilon-1} \bar{y}_i^1. \quad (30)$$

A joint trellis can be constructed by using the new output formulas, and then WSSJD is applicable. The same set partition tree shown in Fig. 1 is used to construct the subset trellis. We investigate change in performance by means of both simulation and error event analysis.

We first consider the case of parallel branches. Assume  $J_1 > 1$ . The effective squared distance between two parallel branches coming from the same state is

$$\begin{aligned} d_r^2(\bar{e}_i \in E_1) &= \frac{(1+\epsilon)^2}{2} (h_0 \bar{e}_i^1 + \frac{\Delta\epsilon}{1+\epsilon} h_0 \bar{e}_i^2)^2 \\ &\quad + \frac{(1-\epsilon)^2}{2} (\frac{\Delta\epsilon}{\epsilon-1} h_0 \bar{e}_i^1 + h_0 \bar{e}_i^2)^2 \quad (31) \\ &= \frac{h_0^2 (1+\epsilon)^2}{2} \left[ (\bar{e}_i^1)^2 + \frac{\Delta\epsilon^2}{(1+\epsilon)^2} (\bar{e}_i^2)^2 \right] \\ &\quad + \frac{h_0^2 (1-\epsilon)^2}{2} \left[ \frac{\Delta\epsilon^2}{(\epsilon-1)^2} (\bar{e}_i^1)^2 + (\bar{e}_i^2)^2 \right]. \quad (32) \end{aligned}$$

Here  $\bar{e}_i$  is defined as in (18). The second equality follows from the fact that when  $J_1 = 2$  or  $J_1 = 3$ ,  $\bar{e}_i$  always has a zero component, so  $\bar{e}_i^1 \bar{e}_i^2 = 0$ .

$$\begin{aligned} d_{\min, \text{asy}}^2(E_1) &= \min_{\bar{e}_i \in E_1} d_r^2(\bar{e}_i \in E_1) \\ &= \begin{cases} h_0^2 \Delta_1^2 + 8\Delta\epsilon^2 h_0^2 & J_1 = 3 \\ h_0^2 \Delta_2^2 + 8\Delta\epsilon^2 h_0^2 & J_1 = 2. \end{cases} \quad (33) \end{aligned}$$

TABLE VII

$d_{\min}^2(E^r)$  FOR ASYMMETRIC 2H2T EPR4 CHANNEL UNDER VARIOUS  $\Delta\epsilon$  WHERE (A)  $\epsilon = 0.1$ , AND (B)  $\epsilon = 0.4$ .

(a)  $\epsilon = 0.1$

	$\Delta\epsilon = 0$ $d_{\min}^2 = 16.16$	$\Delta\epsilon = 0.05$ $d_{\min}^2 = 16.04$	$\Delta\epsilon = 0.1$ $d_{\min}^2 = 16$
RSSE[4, 4, 3]	22.64	22.70	22.88
RSSE[4, 4, 2]	19.44	19.50	19.68
RSSE[4, 4, 1]	12.12	12.03	12.00
RSSE[4, 3, 3]	16.16	16.20	16.32
RSSE[4, 3, 2]	16.16	16.20	16.32
RSSE[3, 3, 3]	9.68	9.70	9.76

(b)  $\epsilon = 0.4$

	$\Delta\epsilon = 0$ $d_{\min}^2 = 11.52$	$\Delta\epsilon = 0.05$ $d_{\min}^2 = 11.60$	$\Delta\epsilon = 0.1$ $d_{\min}^2 = 11.84$
RSSE[4, 4, 3]	21.44	21.50	21.68
RSSE[4, 4, 2]	8.64	8.7	8.88
RSSE[4, 4, 1]	8.64	8.7	8.88
RSSE[4, 3, 3]	18.56	18.60	18.72
RSSE[4, 3, 2]	8.64	8.7	8.88
RSSE[3, 3, 3]	12.16	12.20	12.32

Compared to the symmetric case,  $d_{\min, \text{asy}}^2(E_1)$  is increased both for  $J_1 = 2$  and  $J_1 = 3$ .

For a longer error event  $\bar{e}(D)$ , the induced squared distance is

$$\begin{aligned} d^2(\bar{e}(D)) &= \frac{(1+\epsilon)^2}{2} \|\bar{e}^1(D)h(D) + \frac{\Delta\epsilon}{1+\epsilon} \bar{e}^2(D)h(D)\|^2 \\ &\quad + \frac{(1-\epsilon)^2}{2} \|\frac{\Delta\epsilon}{\epsilon-1} \bar{e}^1(D)h(D) + \bar{e}^2(D)h(D)\|^2. \quad (34) \end{aligned}$$

Then the error state diagram and the error event search algorithm introduced in Section V can be applied to the asymmetric case, with the only modification being that the edge labels are calculated according to (34). Also notice that the zero cycles given in Example 2 and Example 3 remain the same in the asymmetric channel.

We search for  $d_{\min}^2(E^r)$  at two extreme values of  $\epsilon$  and various offsets  $\Delta\epsilon$  on the EPR4 channel. The results are

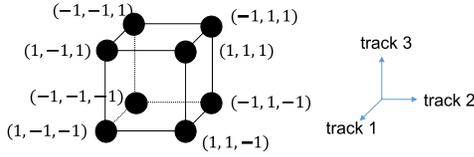


Fig. 7. The input constellation of 3H3T system. The transposed vector beside each node represents the input symbol  $\mathbf{x}_i = [x_i^1, x_i^2, x_i^3]^T$ . The corresponding dimensions are shown on the right.

listed in Tables VII. In each case,  $\Delta\epsilon$  could take values from  $\{0, 0.05, 0.1\}$ . For comparison, we also give the minimum distance parameter of the ML detector, denoted as  $d_{\min}^2$ , in each corresponding scenario. We find that  $d_{\min}^2(E^r)$  does not change much from the symmetric case ( $\Delta\epsilon = 0$ ). In addition, some trellis configurations tend to have increased  $d_{\min}^2(E^r)$  under severe asymmetry, while some do not. We see that the performance of a configuration is closely related to the distance parameters of the length-1 error events, which provides an approach to design the set partition tree for other MHMT channel models. The conclusion is that the proposed RSSE algorithm is applicable to the asymmetric channel.

### VII. 3H3T SYSTEM

The ITI interference matrix of the 3H3T system is given by

$$A_3 = \begin{bmatrix} 1 & \epsilon & 0 \\ \epsilon & 1 & \epsilon \\ 0 & \epsilon & 1 \end{bmatrix}.$$

The WSSJD transformation decomposes the 3H3T system into 3 parallel channels. Recall that the eigen-decomposition of  $A_3$  is  $A_3 = V_3 \Lambda_3 V_3^T$ , where

$$\Lambda_3 = \begin{bmatrix} 1 + \sqrt{2}\epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 - \sqrt{2}\epsilon \end{bmatrix}, \quad V_3 = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ \frac{1}{2} & -\frac{\sqrt{2}}{2} & \frac{1}{2} \end{bmatrix}.$$

The decomposed system is described by

$$\bar{\mathbf{r}}_i = \bar{\mathbf{y}}_i + \bar{\boldsymbol{\omega}}_i, \quad (35)$$

where

$$\bar{\mathbf{r}}_i = \Lambda_3^{-1} V_3^T \mathbf{r}_i, \quad \bar{\boldsymbol{\omega}}_i = \Lambda_3^{-1} V_3^T \boldsymbol{\omega}_i \quad (36)$$

are the transformed channel outputs and noises. The components of vector  $\bar{\mathbf{y}}_i$  are given by  $\bar{y}_i^j = \sum_{k=0}^v h_k z_{i-k}^j$ , for  $j = 1, 2, 3$ , where  $\mathbf{z}_i$  is the transformed channel input vector  $\mathbf{z}_i = V_3^T \mathbf{x}_i$ . Since  $V_3$  is independent of  $\epsilon$ , the joint trellis constructed according to the combination of  $\mathbf{z}_i$  is deterministic. Let  $\mathbf{e}(D) = [e^1(D), e^2(D), e^3(D)]^T$  be an error event of the system, where  $e^j(D)$  is the error event on track  $j$ ,  $e^j(D) = x^j(D) - \hat{x}^j(D)$ . An error symbol at time slot  $i$  is denoted as  $\mathbf{e}_i = \mathbf{x}_i - \hat{\mathbf{x}}_i$ . Then for the transformed 3H3T system,

$$\bar{\mathbf{e}}_i = V_3^T \mathbf{e}_i = V_3^T \bar{\mathbf{x}}_i - V_3^T \hat{\mathbf{x}}_i, \quad (37)$$

and the distance associated with error event  $\mathbf{e}(D)$  is

$$d^2(\mathbf{e}(D)) = \sum_{j=1}^3 \lambda_j^2 \|\bar{e}^j(D)h(D)\|^2, \quad (38)$$

TABLE VIII  
ESPDs OF 3H3T SYSTEM.  $\min d^2$  IS THE MINIMUM  
VALUE OF  $d^2(\mathbf{e}_i)$  ACHIEVED AT  $\epsilon^*$

index	$\mathbf{e}_i$	aci		
		$d^2(\mathbf{e}_i)$	$\min d^2$	$\epsilon^*$
1	$\begin{bmatrix} 2, 0, 0 \\ 0, 0, 2 \end{bmatrix}$	$4 + 4\epsilon^2$	4	0
2	$\begin{bmatrix} 0, 2, 0 \end{bmatrix}$	$4 + 8\epsilon^2$	4	0
3	$\begin{bmatrix} 2, 2, 0 \\ 0, 2, 2 \end{bmatrix}$	$8 + 16\epsilon + 12\epsilon^2$	8	0
4	$\begin{bmatrix} 2, -2, 0 \\ 0, -2, 2 \end{bmatrix}$	$8 - 16\epsilon + 12\epsilon^2$	3	0.5
5	$\begin{bmatrix} 2, 0, 2 \end{bmatrix}$	$8 + 16\epsilon^2$	8	0
6	$\begin{bmatrix} 2, 0, -2 \end{bmatrix}$	8	8	$[0, 0.5]$
7	$\begin{bmatrix} 2, -2, 2 \end{bmatrix}$	$12 - 32\epsilon + 24\epsilon^2$	2	0.5
8	$\begin{bmatrix} 2, -2, -2 \\ -2, -2, 2 \end{bmatrix}$	$12 + 8\epsilon^2$	12	0
9	$\begin{bmatrix} 2, 2, 2 \end{bmatrix}$	$12 + 32\epsilon + 24\epsilon^2$	12	0

where  $\lambda_1 = 1 + \sqrt{2}\epsilon$ ,  $\lambda_2 = 1$  and  $\lambda_3 = 1 - \sqrt{2}\epsilon$  are the eigenvalues on the diagonal of  $\Lambda_3$ .

To construct the reduced-state trellis, we first need to evaluate the distance between input symbols. Recall that the ESPD is proportional to the distance associated with length-1 error events, with the scaling factor  $h_0^2$ . Therefore, for symbols  $\mathbf{x}_i$  and  $\hat{\mathbf{x}}_i$  with difference  $\mathbf{e}_i = \mathbf{x}_i - \hat{\mathbf{x}}_i$ , their ESPD is calculated from

$$d(\mathbf{e}_i) = \sum_{j=1}^3 (\lambda_j \bar{e}_i^j)^2. \quad (39)$$

Table VIII lists the ESPDs for the symbol pair differences. The table is simplified by symmetry considerations, i.e.  $-\mathbf{e}_i$  produces the same distance as  $\mathbf{e}_i$ . The distances are functions of  $\epsilon$ . They display different monotonicity behavior over the range  $\epsilon \in [0, 0.5]$ . For instance, the error symbol 6 has the same distance for all the values of  $\epsilon$ . The distances of error symbols with index 4 and 7 decrease as  $\epsilon$  increases, while for other error symbols the distance functions are increasing functions. Therefore, at different ITI levels, the dominant ESPD is different, which should be taken into account when designing the set partition tree.

In Fig. 8 we propose two set partition trees optimized for low or high ITI levels. For the low ITI case, the single track error symbols, corresponding to the error symbols 1 and 2 in Table VIII, have smaller ESPDs. They are first removed from the level 1 to level 2 partitions in the type-1 construction, shown in Fig. 8(a). To further increase the intrasubset ESPD, the error symbols 4 and 5 are also removed in the level 3 partition, and error symbol 6 is avoided on level 4. Following a similar design rule, the type-2 set partition tree in Fig. 8(b) is constructed to handle the case of high ITI. It is necessary to first remove error symbols 4 and 7 from level 1 to level 2 since their induced distance is much smaller than others. Then the single track error symbols and the error symbol 6 are also avoided on level 3 and level 4, respectively. Notice that the

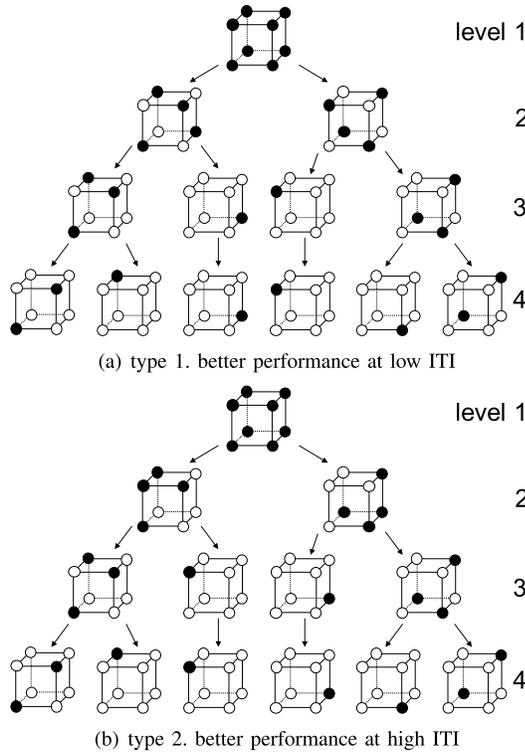


Fig. 8. Set partition trees designed for 3H3T system. Both of the trees have 5 level partitions. To save space, the level 5, where each symbol itself is a subset, is not shown on the pictures.

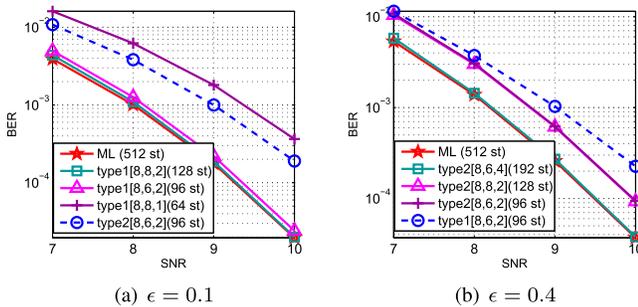


Fig. 9. Simulation results for 3H3T system with EPR4 channel polynomial. The prefix “type1” and “type2” indicate if the subset trellis is based on the type 1 or type 2 set partition tree, respectively.

type-1 and type-2 set partition trees differ only on level 2. Therefore, the subset trellis with  $J_k \neq 2$  for all  $k = 1, \dots, \nu$  will yield the same performance, no matter which set partition tree is used.

In Figs. 9 and Figs. 10 we plot the simulation results for RSSE on the EPR4 channel and a minimum phase channel, respectively. Two extreme cases are considered, corresponding to a relatively low ITI level,  $\epsilon = 0.1$ , and a high ITI level,  $\epsilon = 0.4$ . We construct several subset trellises based on the type-1 set partition tree for  $\epsilon = 0.1$  and the type-2 tree for  $\epsilon = 0.4$ . It can be observed from Fig. 9(a) that the type-1 RSSE [8, 8, 2] configuration and the type-1 RSSE [8, 6, 2] configuration have near-ML performance. For comparison, we also plot the performance curve for the type-2 RSSE[8, 6, 2] trellis, which is a subset trellis constructed using the type-2

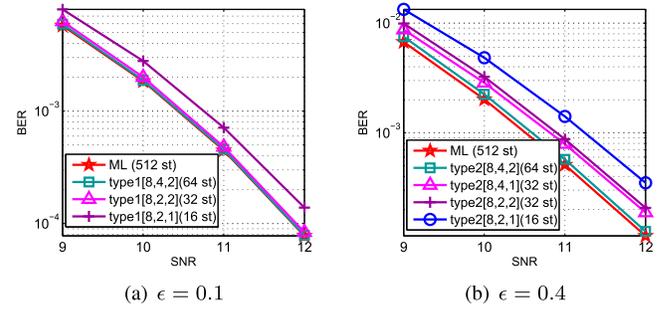


Fig. 10. Simulation results for 3H3T system with minimum phase channel  $h(D) = 1 + 1.6D + 1.1D^2 + 0.4D^3$ .

set partition tree. This trellis suffers from significant performance loss although it has the same configuration as type-1 RSSE [8, 6, 2]. In Fig. 9(b), we see that type-2 RSSE [8, 6, 4] can essentially achieve ML performance; the results for type-2 RSSE[8, 6, 2] and type-1 RSSE[8, 6, 2] are also plotted for comparison purposes.

For the minimum phase channel, whose performance results are shown in Fig. 10, the required computational complexity is further reduced. From Fig. 10(a), we see that the type-1 RSSE [8, 2, 2] trellis with 32 states has performance nearly equal to that of the ML detector, which requires 512 states. In Fig. 10(b), we see that the type-1 RSSE [8, 4, 2] trellis with 64 states provides essentially ML performance.

## VIII. CONCLUSION

Due to its capability of combating ITI, MHMT detection is expected to play an important role in next generation magnetic recording. The conventional ML detector, however, suffers from high computational complexity. In this work we address this problem by applying RSSE techniques with properly designed set partition trees. In particular, we define an alternative distance measure on the input constellation, based on which we propose a three-level set partition tree for the 2H2T model. The BER comparison shows that RSSE can achieve near optimal performance while significantly reducing the number of trellis states. Error event analysis is used to explain the performance variations observed for different trellises under various conditions. We also investigate the performance of RSSE on an asymmetric 2H2T system because of the practical relevance of the model. For the 3H3T model, since the effective distances between input symbols show different monotonicity behavior as  $\epsilon$  changes, we specifically design two set partition trees, where one is good for low ITI and the other is suitable for high ITI. Our work shows that the set partition tree plays a key role in applying RSSE to these channels. If the set partition tree can be properly designed, then the RSSE algorithm has the potential to be effectively applied to more general MHMT channels. This is a direction for future research.

## 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] S. Karakulak, P. H. Siegel, J. K. Wolf, and H. N. Bertram, "A new read channel model for patterned media storage," *IEEE Trans. Magn.*, vol. 44, no. 1, pp. 193–197, Jan. 2008.
- [3] G. Mathew, E. Hwang, J. Park, G. Garfunkel, and D. Hu, "Capacity advantage of array-reader-based magnetic recording (ARMR) for next generation hard disk drives," *IEEE Trans. Magn.*, vol. 50, no. 3, pp. 155–161, Mar. 2014.
- [4] H. Xia, L. Lu, S. Jeong, L. Pan, and J. Xiao, "Signal processing and detection for array reader magnetic recording system," *IEEE Trans. Magn.*, vol. 51, no. 11, pp. 1–4, Nov. 2015.
- [5] E. Soljanin and C. N. Georghiadis, "Multihead detection for multitrack recording channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2988–2997, Nov. 1998.
- [6] L. C. 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.
- [7] B. Fan, H. K. Thapar, and P. H. Siegel, "Multihead multitrack detection in shingled magnetic recording with ITI estimation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, U.K., Jun. 2015, pp. 425–430.
- [8] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 13–20, Jan. 1988.
- [9] A. Baier and G. Heinrich, "Performance of m-algorithm mlse equalizers in frequency-selective fading mobile radio channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Boston, MA, USA, Jun. 1989, pp. 281–285.
- [10] E. Kurtas, J. G. Proakis, and M. Salehi, "Reduced complexity maximum likelihood sequence estimation for multitrack high-density magnetic recording channels," *IEEE Trans. Magn.*, vol. 35, no. 4, pp. 2187–2193, Jul. 1999.
- [11] B. Fan, H. K. Thapar, and P. H. Siegel, "Multihead multitrack detection with reduced-state sequence estimation in shingled magnetic recording," in *Proc. IEEE Magn. Conf. (INTERMAG)*, Beijing, China, May 2015, p. 1.
- [12] B. Fan, H. K. Thapar, and P. H. Siegel, "Multihead multitrack detection with reduced-state sequence estimation in shingled magnetic recording," *IEEE Trans. Magn.*, vol. 51, no. 11, pp. 1–4, Nov. 2015.
- [13] W. Sheen and G. L. Stuber, "Error probability for reduced-state sequence estimation," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 3, pp. 571–578, Apr. 1992.
- [14] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.
- [15] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 18, no. 3, pp. 363–378, May 1972.
- [16] S. A. Altekari, M. Berggren, B. E. Moision, P. H. Siegel, J. K. Wolf, and J. K. W. Fellow, "Error-event characterization on partial-response channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 241–247, Jan. 1998.



**Bing Fan** (S'14) received the B.S. degree in electrical and computer engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. She is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of California, San Diego, La Jolla, CA, USA. She is currently with the Center for Memory and Recording Research. Her research interests include coding and signal processing in storage systems.



**Hemant K. Thapar** (M'78–SM'98–F'98) received the Ph.D. degree in electrical engineering from Purdue University. He was an Adjunct Lecturer with Electrical Engineering, Santa Clara University, from 1984 to 2004, and a Visiting Research Scientist with the Center for Memory and Recording Research, University of California, San Diego, from 2013 to 2014. He was previously a co-founder and the CEO of two successful startups, Link-A-Media Devices and DataPath Systems. He held various technical and management positions with Bell Telephone Laboratories, Holmdel, from 1979 to 1984, and with IBM Corporation, San Jose from 1984 to 1994, where he was involved in the areas of circuit-switched networking, data communications, and data storage. He is currently a founder and the CEO of OmniTier, Inc., a start-up company developing infrastructure application solutions for cloud data centers. He is a co-recipient of three best paper awards for his work on high-speed data transmission and high-density data storage. He served as a Guest Editor of two special issues of IEEE Transactions devoted to data storage technologies. He serves on the Board of Directors of Assia Corporation, a private company enabling next-generation broadband technology and products, and on the Advisory Boards of the School of Engineering, Santa Clara University, and the Ambala College of Engineering and Applied Sciences, India.



**Paul H. Siegel** (M'82–SM'90–F'97) received the S.B. and Ph.D. degrees in mathematics from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1975 and 1979, respectively. He held a Chaim Weizmann Post-Doctoral Fellowship with the Courant Institute, New York University, New York, NY, USA. He was with the IBM Research Division, San Jose, CA, USA, from 1980 to 1995. He joined the faculty at the University of California, San Diego, La Jolla CA, USA, in 1995, where he is currently a Professor of Electrical and Computer Engineering with the Jacobs School of Engineering. He is also with the Center for Memory and Recording Research, where he holds an Endowed Chair and served as the Director from 2000 to 2011. His research interests include information theory and communications, particularly coding and modulation techniques, with applications to digital data storage and transmission. He was a member of the Board of Governors of the IEEE Information Theory Society from 1991 to 1996 and from 2009 to 2014. He is a member of the National Academy of Engineering. He was the 2015 Padovani Lecturer of the IEEE Information Theory Society. He was a recipient of the 2007 Best Paper Award in Signal Processing and Coding for Data Storage from the Data Storage Technical Committee of the IEEE Communications Society. He was a co-recipient of the 1992 IEEE Information Theory Society Paper Award and the 1993 IEEE Communications Society Leonard G. Abraham Prize Paper Award. He served as a Co-Guest Editor of the 1991 Special Issue on Coding for Storage Devices of the IEEE TRANSACTIONS ON INFORMATION THEORY. He served as an Associate Editor of Coding Techniques of the IEEE TRANSACTIONS ON INFORMATION THEORY from 1992 to 1995, and as the Editor-in-Chief from 2001 to 2004. He was also a Co-Guest Editor of the 2001 two-part issue on The Turbo Principle: From Theory to Practice and the 2016 issue on Recent Advances in Capacity Approaching Codes of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.