UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CHANNEL DETECTION ON TWO-DIMENSIONAL MAGNETIC RECORDING

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ENFENG JIANG
Norman, Oklahoma
2013

CHANNEL DETECTION ON TWO-DIMENSIONAL MAGNETIC RECORDING

A DISSERTATION APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

_____
Dr. J. R. Cruz, Chair

_____
Dr. Ralf Schmidt

_____
Dr. Joseph Havlicek

_____
Dr. Tian-You Yu

_____
Dr. Lei Ding

Acknowledgements

The completion of this dissertation has never been like walking in the park, it required years of diligent working. Here, I would like to express my sincere appreciation to the individuals who have been a part of my doctoral career.

First of all, I would like to express my greatest gratitude to my advisor Dr. J. R. Cruz for his brilliant advice, encouragement during my years of study and research. I would not have this dissertation completed without his guidance.

I would like to thank Dr. Joseph Havlicek, Dr. Tian-You Yu, Dr. Lei Ding and Dr. Ralf Schmidt for their years of serving and support on my doctoral committee.

I would like to thank Hitachi Global Storage Technologies (HGST) for providing financial support for the TDMR project, and Dr. Richard Todd for his kind cooperation.

Also I would like to thank Dr. Wu Chang, Dr. Han Wang and Dr. Clint Keele, who are all former students of Dr. Cruz, for sharing their valuable experience and giving me helpful advises.

I would also like to give credit to the ECE department staff, especially Lynn Hall, Lisa Wilkins and Debbie Manning for their diligent help.

Finally, I would like to give my deepest appreciation to my dear parents for their endless love over the years, as well as my friends in the Trinity Baptist Church for their spiritual support.

Table of Contents

List of Tables

List of Figures

Abstract

Two-dimensional magnetic recording (TDMR) coupled with shingled-magnetic recording (SMR) is one of next generation techniques for increasing the hard disk drive (HDD) capacity up to 10 Tbit/in^2 in order to meet the growing demand of mass storage.

We focus on solving the tough problems and challenges on the detection end of TDMR. Since the reader works on the overlapped tracks, which are even narrower than the read head, the channel detector works in an environment of low signal-to-noise ratio (SNR), two-dimensional (2-D) inter-symbol interference (ISI) and colored noise, therefore it requires sophisticated detection techniques to provide reliable data recovery. Given that the complexity of optimal 2-D symbol detection is exponential on the data width, we had to choose suboptimal solutions.

To build our research environment, we use an innovative Voronoi grain based channel model which captures the important features of SMR, such as squeezed tracks, tilted bit cells, 2-D ISI, electronic and media noise, etc. Then we take an in-depth exploration of channel detection techniques on the TDMR channel model. Our approaches extend the conventional 1-D detection techniques, by using a joint-track equalizer to optimize the 2-D partial-response (PR) target followed by the multi-track detector (MTD) for joint detection, or using the inter-track interference (ITI) canceller to estimate and cancel the ITI from side tracks, followed by a standard BCJR detector. We used the single-track detector (STD) for pre-detecting the side tracks to lower the overall complexity. Then we use

pattern-dependent noise prediction (PDNP) techniques to linearly predict the noise sample, so as to improve the detection performance under colored media noise, and especially the data dependent jitter noise. The results show that our 2-D detectors provide significant performance gains against the conventional detectors with manageable complexity.

Chapter 1 Introduction to Two-Dimensional Magnetic Recording

The hard-disk drive (HDD) capacity has been growing over the years at a rate of 30% to 50% per year to meet the demand of mass storage [1], but the current generation of hard-disk drives using perpendicular magnetic recording (PMR) is approaching the areal density limit of reliable data storage at about 1 Tb/in^2 [2]. At the same time, HDDs are challenged by the solid-state drives (SSDs) in the storage market, while SSDs boast faster access speed without rotating platters, HDDs can still provide much bigger capacity at an affordable price. The hybrid drive, which is a combination of these two, has a small amount of flash memory for most frequently used data and HDD for mass storage.

The HDDs still play an important part in the storage industry, but increasing the capacity while maintaining its reliability is the top priority. It has been known that a magnetic grain is the smallest unit for recording information on magnetic media. To increase the areal density, one has to decrease the bit size but keep the number of grains in each bit constant in order to maintain the system signal-to-noise ratio (SNR), so the grains have to be downsized correspondingly. However, the grain size cannot be too small due to the super-paramagnetic limit [3]. To be specific, there is a tradeoff among three parameters: system SNR, write capability of the head and thermal stability of the media. If we continue shrinking the grain size, the media becomes thermally unstable and would not provide reliable data storage.

Under these circumstances, the areal density of PMR is reaching the limit. Researchers have proposed several different approaches to address this problem. So

far the most likely techniques are bit-patterned magnetic recording (BPMR) [4], energy-assisted magnetic recording (EAMR), including heat-assisted magnetic recording (HAMR) [5] and microwave-assisted magnetic recording (MAMR) [6], and two-dimensional magnetic recording (TDMR) [1]. Before we discuss these techniques in detail, let us first review conventional PMR technology.

1.1. Conventional Magnetic Recording Technology

Conventional HDDs use PMR techniques to record the data on continuous media. The direction of the magnetization of the bits is perpendicular to the media plane, rather than parallel to the tracks as in longitudinal magnetic recording (LMR).

The channel response of a magnetic recording channel is related to the data transitions, in other words the change of magnetization direction. The transition response of the PMR channel is usually modeled as an error function or hyperbolic tangent function [7], denoted by

$$h(t) = V \tanh\left(\frac{\ln 3}{T_{50}} t\right), \tag{1.1}$$

where V is the amplitude of the response and T_{50} is the time interval between $-V/2$ and $V/2$ of h(t) in a single transition.

The readback waveform can be expressed as

$$\begin{aligned} r(t) &= \sum_k (a_k - a_{k-1}) h(t - kT) + n(t) \\ &= \sum_k b_k h(t - kT) + n(t), \end{aligned} \tag{1.2}$$

where $\{a_k\}$ denotes the written bits encoded as {-1,+1}, $b_k = a_k - a_{k-1}$ is the transition sequence, $n(t)$ is the additive white Gaussian noise (AWGN), and T is the bit interval. If we replace the data transition to the difference of the responses, (1.2) can be written in another form:

$$\begin{aligned} r(t) &= \sum_k a_k \left(h(t-kT) - h(t-(k-1)T) \right) + n(t) \\ &= \sum_k a_k \, p(t-kT) + n(t), \end{aligned} \tag{1.3}$$

where $p(t) = h(t) - h(t-T)$ is called the dibit response.

The PMR system also suffers from transition noise. A major part of the transition noise is position jitter noise, which happens in the data transitions because of the irregularity of the bit boundaries, such that the sampling position is moving back and forth around the ideal position.

The jitter noise at time k can be modeled as a Gaussian random variable Δt_k such that the transition response becomes $h(t-kT+\Delta t_k)$, which can be expanded by Taylor series:

$$h(t-kT+\Delta t_k) = h(t-kT) + \Delta t_k h'(t-kT) + \Delta t_k{}^2 h''(t-kT)\cdots, \tag{1.4}$$

As Δt_k is relatively small compared to the overall channel output, we choose to ignore the part beyond the second order and use the first-order approximation:

$$r(t) = \sum_k b_k h(t-kT) + \sum_k b_k \Delta t_k h'(t-kT) + n(t), \tag{1.5}$$

A typical PMR channel model is shown in Fig 1.1.

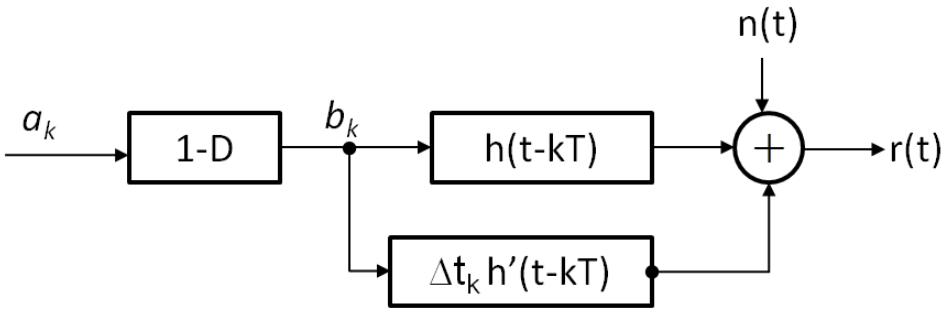

Fig 1.1. Block diagram of a PMR channel model.

1.2. Next Generation Recording Techniques

A. BPMR

BPMR [4] is a promising technology that uses patterned media, which is a totally different media than the conventional perpendicular magnetic media. The patterned media platter consists of well-defined isolated magnetic island separated by non-magnetized area or spacing. Each individual island records a single bit, which contains at least one magnetic grain. The layout of the islands can be rectangular or hexagonal arrays, giving different channel specifications, as shown in Fig. 1.2.

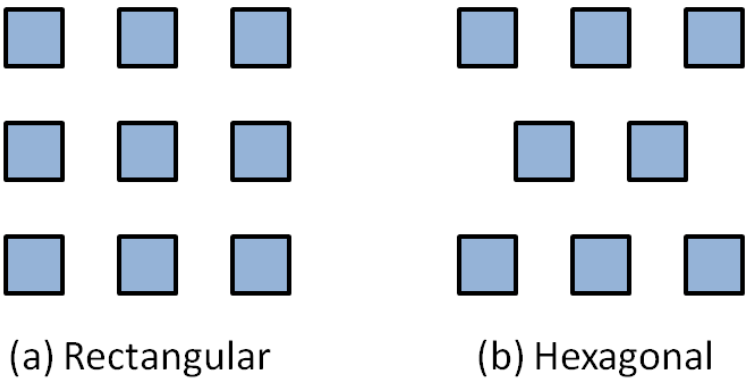

Fig 1.2. Examples of bit-patterned media.

Such architecture largely reduces the number of grains per bit without sacrificing the grain size and SNR, as the adjacent bits are separated. The media noise in BPMR is quite different from PMR on continuous media. As the spacing between islands is not magnetized, the transition noise characteristics of PMR can be ignored. But the media noise mainly comes from material fabrication, such as size, shape and position variation of the islands.

Obviously such delicate patterned media is quite expensive for mass production, compared with the traditional PMR media [8]. Also the issue of writing synchronization is essential to the system, because the size of the island is so small, it is easy to introduce write errors.

B. EAMR

EAMR basically increases the write capability while maintaining the thermal stability, either by the assistance of heat (HAMR) or microwave (MAMR) energy, such that the size of the thermally stable grains can be reduced. During the writing process of HAMR, a tiny area of the media is temporarily heated by a laser beam to lower the threshold of reversing the magnetization polarity. As soon as the magnetization is done, the heat is removed and the medium quickly cools down to become thermally stable [5]. On the other hand in MAMR, the energy source is not operating directly on the media, but using a radio frequency to enable writing saturation below the coercivity of the media while writing [6].

HAMR and MAMR coupled with BPMR have the potential of reaching the areal density at 10 Tbit/in^2 [9]. At the same time, there are some challenges in practice. The head must be carefully designed to integrate the heat or microwave source. In HAMR, the material should sustain frequent laser heated writing process so the life span of the material is also an important element to consider.

C. TDMR

Unlike the techniques mentioned above, TDMR proposed by Wood in [1] still operates on the conventional PMR media without reducing the grain size or using external energy assistance, but maintains the write-ability and thermal stability of the media.

The unique feature of TDMR is using shingled magnetic recording (SMR), in the writing process for SMR, the tracks are sequentially written and partially overlap one after another by a certain percentage, which leaves the non-overlapped portion as the actual tracks. Fig. 1.3 simply illustrates a shingled writing process.

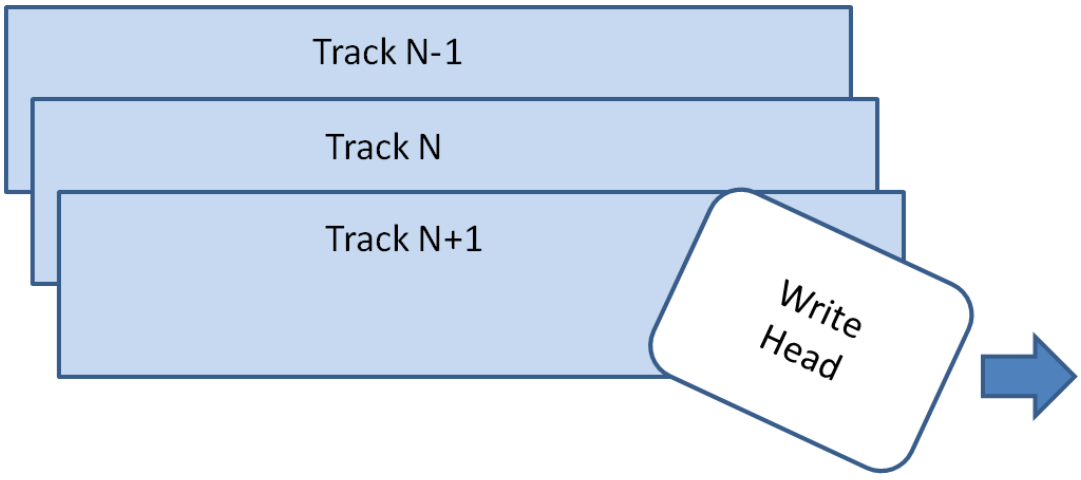

Fig 1.3. Illustration of shingled writing process.

One of the drawbacks of SMR is that we cannot simply re-write a single track because it will overwrite the other existing tracks. Suppose we want to write another track exactly on track N-1, as shown in Fig. 1.3, then the portion of track N overlapping track N-1 will be erased. So if we need to rewrite a certain track, the whole section must be rewritten.

The reason of using SMR is two-fold: firstly the size of the head is maintained, which provides enough magnetic field for writing and so external energy assistance is not required. Secondly, the shingled writing can make the track pitch much smaller than usual in order to reach a much higher channel density. According to [1] TDMR with SMR has the potential of increasing the areal density up to $10Tb/in^2$.

In this dissertation we are interested in exploring TDMR with SMR, because it operates on the continuous media, and some techniques can be ported from PMR.

1.3. TDMR Channel Modeling

Having a good TDMR channel model is very important for further research. So There are currently several TDMR models that have been applied, such as the 4-grain model, which consists of 4 types of grains (1x1, 1x2, 2x1, and 2x2 pixels) [8], microcell model to create irregular bit boundaries [10], and the grain flipping model based on micro-magnetic simulations [11].

The above models are either too simple or too complicated. We choose to use the Voronoi-based channel model proposed by Todd, Jiang *et al.* [12] , which is an

innovative method to incorporate the important features of SMR process in the model, but not extremely complicated to implement.

The modeling problem can be divided into three parts: creating the Voronoi grain layout, getting the grains magnetized and generating the readback waveform. In the following subsections, we will discuss the steps of modeling the Voronoi-based channel described in [12].

1.3.1. Creating the grain layout

The first step is to generate magnetic grains. The so-called Voronoi grains are generated by randomly distributed points (also called nuclei) on a plane, and the region closer to one point than any other points represents a magnetic grain. Once the areal density and average number of grains in a bit cell are determined, the average size of the grain can be determined. But the problem is: as the points are quite randomly distributed, the sizes of the grains will have large variations. In this case we take several steps to cope this problem.

1) Replace each grain nucleus with the grain centroid.
2) Remove small grains with size less than a threshold.
3) Split large grains size above a threshold, adding a new grain nucleus besides the original nucleus.
4) Repeat steps 1-3 until the standard deviation of grain sizes are sufficiently close to the desired target.

5) Finally, to create the non-magnetic grain boundaries, the grains are shrunk by moving each corner in towards the nucleus by 10%.

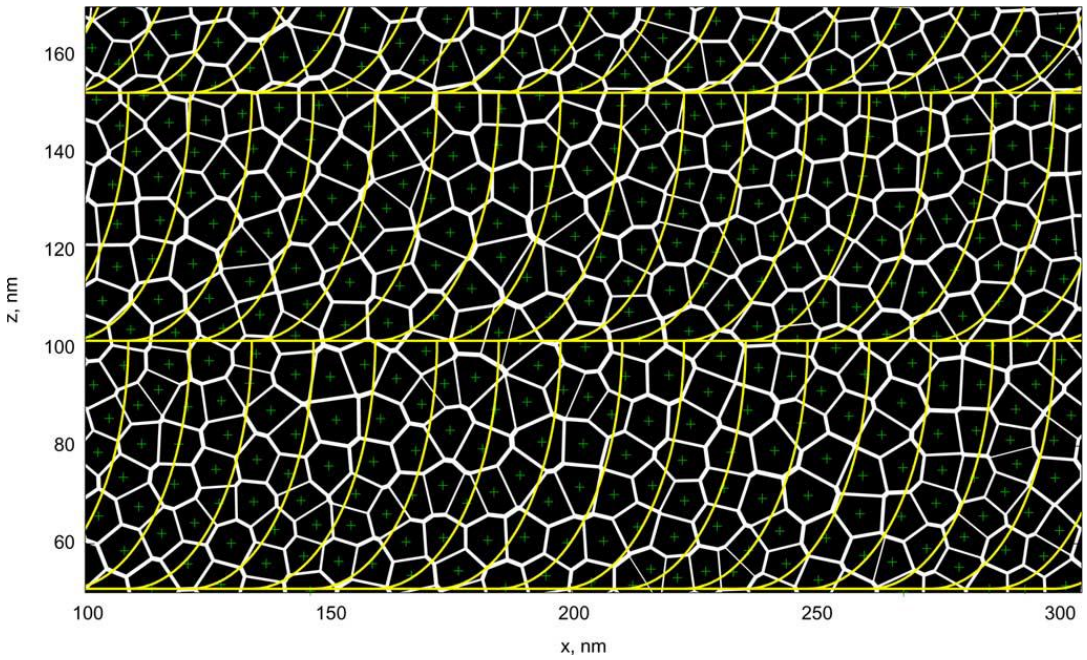

Fig 1.4. Voronoi grains and bit cells at 1Tb/in^2 [12].

An example plot of grains and bit cells generated from our TDMR model is shown in Fig. 1.4, where green dots are grain centroids, the white lines indicate the grain boundaries, and the yellow lines are boundaries of the bit cells. *x* and *z* represent coordinates on down-track and cross-track axis, respectively.

1.3.2. Write process

As shown in Fig 1.4, the bit cell is in a "curved-rectangular" shape instead of regular rectangular. According to the shingled writing process described in [1], the

writing head is tilted such that the especially designed corner can be used to reach a higher writing field. As the head is much wider than the actual track pitch, only one corner of the head is essential for the writing. The specific shape of a bit cell is determined by many parameters, such as the areal density, average grains per bit, and bit aspect ratio (BAR), which is defined as the ratio of cross-track length over the along-track length of each bit. In Fig 1.4, the BAR is equal to 4.

The magnetized grains give a two-dimensional array called the channel magnetization $m(x,z)$. At the beginning, the grains are randomly magnetized as +1 or -1. Once the written bits are determined, those grains whose centroids fall in the region of a bit cell are magnetized according to the polarity of the certain bit.

However, in the actual writing process some of the grains may be incorrectly magnetized, especially on the border because of the overlapped tracks and bits in the shingled writing. In order to include this important feature, we define a probability $p(g, i, j)$ of grain g being magnetized to the value of bit i of track j as a function of the grain centroid (x_g, z_g).

We define the function $A_{i,j}(x,z)$ to be the perfect writing probability function, where $A_{i,j}(x,z)=1$ when x and z are inside that bit cell and 0 outside.

And we convolve $A_{i,j}(x,z)$ with a two-dimensional Gaussian function $G(x,z)$ to get the the probability $p(g, i, j)$ of grain g being magnetized to the value of bit (i,j)

$$p(g,i,j) = (A_{i,j}(x,z) * G(x,z))|_{x=x_g, z=z_g}, \tag{1.6}$$

Hence, this probability is near one inside the (*i*, *j*) bit cell and drops off as one moves away from the bit cell borders. In this way our model thus implements a certain small probability that grains in bit cells adjacent to (*i*, *j*) will get incorrectly written to the data written in the (*i*, *j*) bit cell. Our model attempts to capture this important feature without requiring the extensive micro-magnetic computations.

1.3.3. Read process

The two-dimensional (2-D) head response is given by Wood in [13], which includes a positive main response and a negative response. The combined response is in the form:

$$h(x,z)=\frac{1}{2\pi\sigma_x\sigma_z}\exp\left(-\frac{x^2/\sigma_x{}^2+z^2/\sigma_z{}^2}{2}\right) -\frac{1}{2\pi l_x l_z}K_0\left(\sqrt{x^2/\sigma_x{}^2+z^2/\sigma_z{}^2+\varepsilon^2}\right). \tag{1.7}$$

The positive response is approximated as a bi-variant Gaussian function, where the standard deviation on each direction is proportional to the 50% width W_{50}, with the relationship $\sigma=\frac{W_{50}}{2\sqrt{2\ln 2}}$.

The negative response is in the form of an elliptical Bessel function, where l denotes the characteristic decay length of the flux in the soft underlayer (SUL), and ε is a small number introduced to avoid a singularity at the position (0,0).

The readback signal $r(x,z)$ is generated by the 2-D convolution of channel magnetization $m(x,z)$ with the head response $h(x,z)$, plus a certain amount of AWGN $n(x,z)$:

$$r(x,z) = m(x,z) * h(x,z) + n(x,z). \tag{1.8}$$

1.3.4. Model testing

Given the write data $w(x,z)$ and the computed readback waveform $r(x,z)$, one can calculate an estimate of the system response (do not confuse with head response). The system response is defined as the function $f(x,z)$ such that

$$r(x,z) = w(x,z) * f(x,z), \tag{1.9}$$

$f(x, z)$ can be estimated by taking fast Fourier transforms (FFTs) of $r(x, z)$ and $w(x, z)$, dividing them, and taking the inverse FFT, then averaging over multiple realizations.

$$f(x,z) = \mathrm{IFFT}\left\{\frac{\mathrm{FFT}[r(x,z)]}{\mathrm{FFT}[w(x,z)]}\right\}, \tag{1.10}$$

Once the system response is estimated, we can compute these parameters and compare against experimental data:

1. The T_{50} width of the down-track step response (bit response integrated along the line $z = 0$).

2. Cross-track full-width-half-magnitude (FWHM): the width at the half of the magnitude in the z-direction.

3. The intertrack interference (ITI) ratio: root-mean-square (RMS) magnitude of the response at $z = 0$ compared with that at $z = \pm T_p$. (T_p is the track pitch.)

The parameters of certain channel specification are given in [12]: with the model set for 1 Tb/in^2, BAR=4:1, and a density of ten teragrains per square inch, we got a T_{50} width of 14.5 nm and z-direction FWHM of 47.2 nm, with ITI ratios on each side of 7.4% and 8.5%, respectively. A plot of the 2-D bit response is shown in Fig. 1.5.

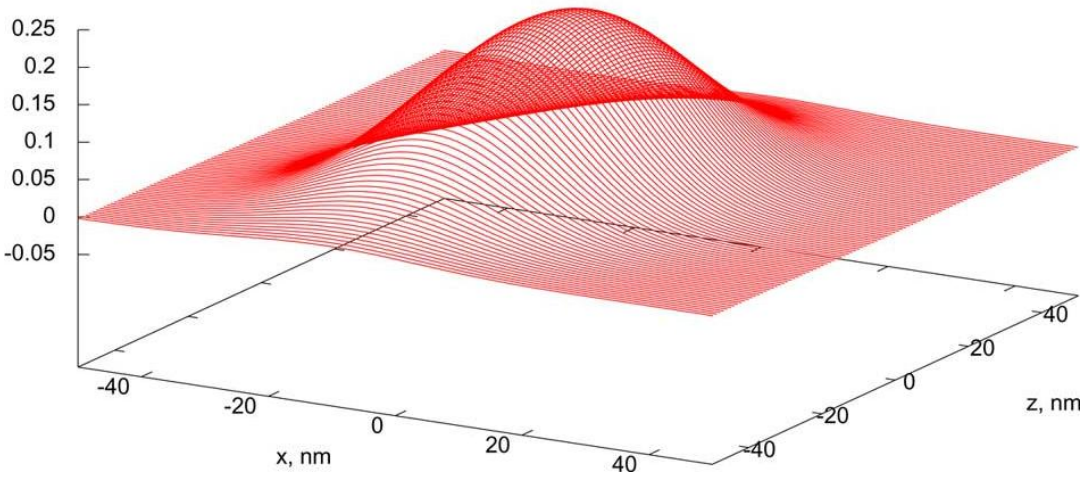

Fig 1.5. Example of estimated bit response, BAR = 4 [12].

One can also compute least-squares estimates of the so-called "track-to-track" responses, i.e., computing response $f(x)$ that is the best least-squares fit for $r(x,z_1)=f(x)*w(x,z_2)$ for a pair of track locations z_1, z_2. An example of track-to-track bit response is shown in Fig. 1.6 from the center of a track to itself. The T_{50} can be

computed from such a track-to-track response, and the resulting value is 14.6 nm, approximately the same result as we got from the FFT-based response.

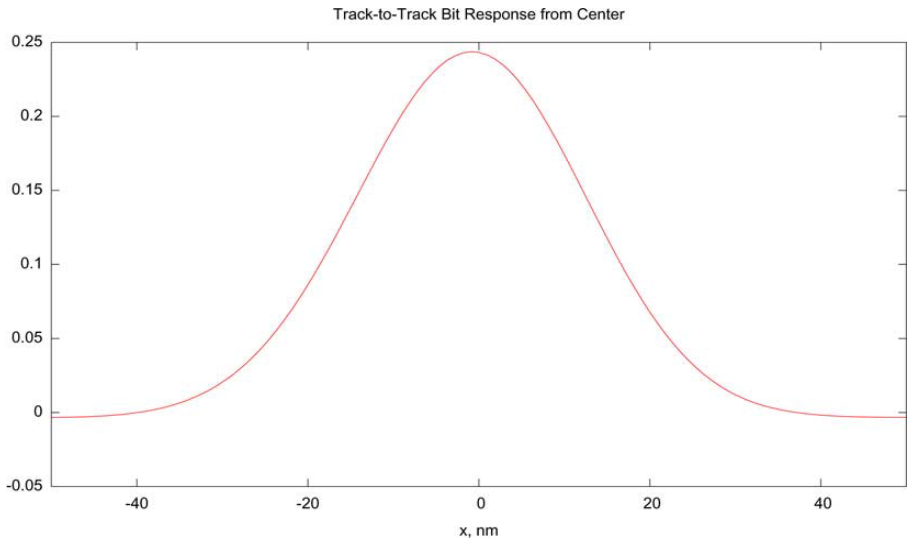

Fig 1.6. Track-to-track bit response [12].

1.4. Challenges in TDMR Channel Detection

By using SMR technology, TDMR is able to squeeze the size of the bits without sacrificing thermal stability, but the SNR is going down dramatically. Therefore the biggest challenge for TDMR system falls on the signal processing part, providing reliable detection schemes for data recovery, and we have several tough problems to solve.

1) 2-D inter-symbol interference

In traditional PMR, there are guarding bands between tracks, and track pitches are fairly wide. Although the head response is two-dimensional, very little response

is due to the adjacent tracks and so we treat PMR as a one-dimensional (1-D) system. But this situation never happens in a high-density TDMR system. The read head is wider than the shingled tracks, so the head picks up the signal on the main track as well as the signals on adjacent tracks. Therefore, the readback signal contains inter-symbol interference (ISI) not only from down-track direction but also from the cross-track direction, where the ISI on the cross-track is also called ITI cannot be ignored. The conventional 1-D detector simply treats ITI as additive noise and has poor performance. Aggressive 2-D signal processing is necessary for TDMR channel detection.

2) Media noise

In a TDMR system, media noise will be the dominate noise component, and the electronic noise (AWGN) is not important. Since we are still using the continuous magnetic recording media in TDMR, the channel still suffers from media noise. On the along-track direction, the transition noise mainly comes from the irregular grain boundaries. On the cross-track direction, there is also media noise existing on the border between the tracks, due to the shingled writing process.

3) Complexity

It is well known that maximum-likelihood sequence detection (MLSD) and maximum *a posteriori* (MAP) decision are two optimal solutions for the 1-D detection problem. However, it is not straightforward to generalize the 1-D

detection to the 2-D case. In 1-D channel detection, the complexity of optimal detectors is related to the states of the trellis, or essentially the ISI memory length *I*. For binary input/output trellis, the number of trellis states is 2^I, and two transitions for each state, so the total number of transitions of the whole data sequence of length *N* is $N2^{I+1}$.

Now let us expand the idea to the 2-D case. The two-dimensional data $a_{k,l}$ can be expressed in an $M \times N$ data block, and the 2-D ISI channel response $h_{i,j}$ has size $I_x \times I_y$. Then the input/output relationship of a two-dimensional ISI channel can be written as:

$$r_{k,l} = \sum_{i=0}^{Ix} \sum_{j=0}^{Iy} h_{i,j} a_{k-i,l-j} + n_{k,l} \ . \tag{1.11}$$

One way of generalizing the 1-D detection schemes to the 2-D case is to treat each column as a symbol consisting of *M* binary bits, then the 2-D data block can be considered as a 1-D 2^M-ary data sequence, as shown in Fig. 1.7.

Now the full trellis has $\left(2^M\right)^{I_x}$ states with 2^M transitions per state, so the total number of transitions for one symbol is $2^M \times \left(2^M\right)^{I_x} = \left(2^M\right)^{I_x+1}$, therefore the complexity of optimal detecting one data block is in the order of $N2^{M(I_x+1)}$, which means the complexity of optimal 2-D detection is not only exponential to the ISI memory, but also to the data width [14]. In addition, it is been proved in [15] that the optimal MLSD in 2-D detection is NP-complete. Due to all of these reasons, in the real world we have to use suboptimal 2-D detection algorithms.

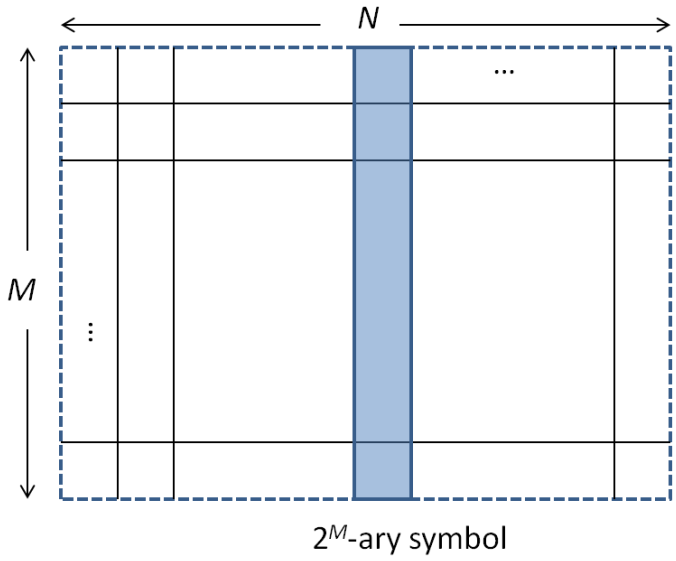

Fig 1.7. Construction of 2-D trellis [14].

1.5. Overview of the Dissertation

The goal of this dissertation is to provide practical solutions to the challenges caused by TDMR systems. Since we have a fairly good Voronoi channel model in place, we focus on the problems of channel detection, in short words, dealing with 2-D ISI and media noise under manageable computational complexity.

The rest of the dissertation is organized as follows: Chapter 2 is a review of the conventional 1-D detection techniques, including fundamental channel partial response equalization process and channel detection algorithms, such as the Viterbi algorithm (VA), the Bahl-Coke-Jelinek-Raviv (BCJR) algorithm and soft-output VA (SOVA). In Chapter 3 we focus on channel detectors dealing with ITI such as multi-track detectors (MTD) and ITI cancellers, and then we compare their performance results on a squeezed PMR channel and the TDMR channel model.

Chapter 4 gives the solution of dealing with colored media noise by using pattern dependent noise predictive (PDNP) detectors. Finally we give a conclusion of the dissertation and some remarks on future research work in Chapter 5.

Chapter 2 1-D Magnetic Recording Channel Detection

Before discussing the complicated 2-D detection algorithms, in this chapter we revisit the fundamentals of channel detection in a magnetic recording system, and then we take some detailed reviews of the current technology working on 1-D magnetic recording channels.

2.1. Magnetic Recording System

The digital magnetic recording system is designed to provide reliable data storage and recovery with very little error tolerance in an HDD. First of all the binary written data {0,1} are encoded to NRZ rectangular waveforms of {-1, +1}, indicating two opposite directions of the magnetic units. The data stream is recorded on the magnetic media through the write head. The magnetic recording media can be generally considered as a communication channel, although the channel response and characteristics are different from each other, like in LMR, PMR or the future BMPR, HAMR, TDMR, etc., Then the magneto-resistance (MR) read head flying above the rotating magnetic platter picks up the readback signal as the channel output, which will be filtered and sampled, followed by the equalizer and detectors to recover the original data bits. Fig 2.1 shows how those different components work together in a typical magnetic recording system.

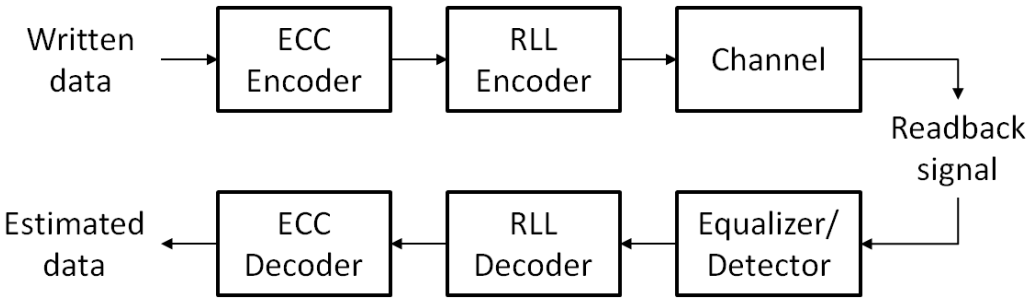

Fig 2.1. A block diagram of a magnetic recording system.

In order to provide some extra error control, the system usually includes run-length limited (RLL) and error correcting codes (ECC) encoders and decoders. Run-length limited (RLL) codes, or also called modulation codes, control the minimum and maximum run-length of the data transitions, because overly frequent transitions will surely increase the transition noise, while the absence of transitions will make the timing recovery more difficult.

ECC encoding/decoding is a very important area of the research in magnetic recording. Basically ECC alleviates the random errors by adding redundancy to the information bits. In general ECC can be categorized into convolutional codes and block codes. Some well-known block codes used in magnetic recording system include low-density parity check (LDPC) codes and Reed-Solomon (RS) codes. For more details about these coding techniques, please refer to [16], [17].

We are focused on the signal processing aspect of the magnetic recording system, to be specific, the equalization and detection of the signal coming out of the channel. Since the readback signal is usually contaminated by ISI, electronic

thermal noise and media noise, channel equalization and detection are essential processes of reducing the ISI and making reliable decisions on the written data.

2.2. Detection Algorithms for ISI Channel with AWGN

Once the magnetic recording media is modeled as an ISI channel and the noise is only AWGN, the data can be estimated by optimal channel detectors. The optimal detection on an ISI channel with AWGN can be accomplished by MLSD or MAP detectors. MLSD, which minimizes the sequence error, is usually implemented by the VA. On the other hand, MAP detection, which minimizes the symbol error, can be done by the BCJR algorithm. SOVA derived from the VA is also used as a suboptimal alternative MAP detector. We will review those algorithms as the foundation of advanced detection algorithms in the following subsections.

2.2.1. Viterbi algorithm

The VA [18] has been applied in decoding convolutional codes in wireless communications and channel detection for many years. As linear codes and ISI channels can be both represented by a trellis, which is an illustration of a finite state machine, the VA finds the noiseless sequence that has the smallest distance to the received sequence through the trellis, where the distance here refers to Euclidian distance in channel detection and Hamming distance in decoding, in this way, the VA is an optimal detector in minimizing the sequence error.

Let us look at an example of a discrete magnetic recording channel system. $\mathbf{a} = [a_1, a_2, \cdots, a_N]^T$ is the input data sequence with length *N*, $\mathbf{h} = [h_0, h_1, \cdots, h_I]^T$ is the channel ISI impulse response with memory *I*, the received signal $\mathbf{r} = [r_1, r_2, \cdots, r_N]^T$ can be represented as:

$$r_k = \sum_{i=0}^{I} h_i a_{k-i} + n_k, \quad k = 1, 2 \cdots N, \tag{2.1}$$

where $\mathbf{n} = [n_1, n_2, \cdots, n_N]^T$ is AWGN with zero mean and variance of σ^2.

To simplify the following notations, we denote the noiseless channel output as

$$z_k \triangleq \sum_{i=0}^{I} h_i a_{k-i}, \; k = 1, 2 \cdots N, \tag{2.2}$$

The VA finds the most likely sequence $\hat{\mathbf{a}}$ by minimizing the Euclidean distance between the noiseless output u_k and the received signal:

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}} \left\{ p(\mathbf{r} \mid \mathbf{a}) \right\}. \tag{2.3}$$

Appling the chain rule for the Markov model,

$$\begin{aligned} p(\mathbf{r} \mid \mathbf{a}) &= \prod_{k=1}^{N} p(r_k - z_k) \\ &= \prod_{k=1}^{N} p(n_k). \end{aligned} \tag{2.4}$$

And (2.4) is essentially the probability density function of the AWGN. Assuming the noise is i.i.d. Gaussian, we shall get

$$\begin{aligned}\hat{\mathbf{a}} &= \arg\max_{\mathbf{a}} \prod_{k=1}^{N} \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\left(r_k - z_k\right)^2 / 2\sigma^2 \right] \right\} \\ &= \arg\max_{\mathbf{a}} \left\{ \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N \exp\left[-\sum_{k=1}^{N} \left(r_k - z_k\right)^2 \Big/ 2\sigma^2 \right] \right\}.\end{aligned} \tag{2.5}$$

If we take the logarithm of (2.5) and disregard irrelevant items,

$$\begin{aligned}\hat{\mathbf{a}} &= \arg\max_{\mathbf{a}} \sum_{k=1}^{N} \left\{ \ln \frac{1}{\sqrt{2\pi\sigma^2}} - \left(r_k - z_k\right)^2 / 2\sigma^2 \right\} \\ &= \arg\max_{\mathbf{a}} \sum_{k=1}^{N} \left[-\left(r_k - z_k\right)^2 / 2\sigma^2 \right] \\ &= \arg\min \sum_{k=1}^{N} \left(r_k - z_k\right)^2 / 2\sigma^2,\end{aligned} \tag{2.6}$$

where we define the branch metric as

$$\lambda_k = \left(r_k - z_k\right)^2 / 2\sigma^2 . \tag{2.7}$$

Therefore the goal is to find the sequence that minimizes the path metric M_k, which is the sum of the branch metrics till time k. The data sequence with the minimum path metric is called survivor path. Note that the term $1/2\sigma^2$ in (2.7) can be dropped in hard decision VA, but we will keep it for log likelihood calculation in SOVA.

Briefly, the VA runs in the following steps: At time 0, the path metrics need to be initialized as $M_0(S_0 = S_0') = 0$, $M_0(S_0 \neq S_0') = -\infty$ where S_0' is a known state at time 0. At time *k*, the state S_k has two incoming paths from two S_{k-1} for binary inputs. Each path updates the path metric by adding the branch metric $\lambda_k\left(S_{k-1}, S_k\right)$

to the existing path metric $M(S_{k-1})$. Then the smaller path metric of the two paths is chosen to be $M(S_k)$,

$$M(S_k) = \min\left(M(S_{k-1}) + \lambda_k\left(S_{k-1}, S_k\right)\right). \tag{2.8}$$

Updating path metrics is simply noted as the "add-compare-select" process and is repeated as *k* increases. As long as the paths merge at a time *k-D*, where *D* is a large enough delay, the decisions of the ML sequence before time *k-D* can be made.

2.2.2. SOVA

SOVA [20] is a soft-output detection algorithm based on the VA. In addition to the hard decision Viterbi detector as MLSD, SOVA also provides reliability information of each symbol by including probability of choosing the wrong path. Therefore SOVA is also a suboptimal MAP symbol detector. Intuitively, when the difference of the path metrics between survivor path and the loser path are quite large, it is more likely we choose the correct survivor path, on the contrary, when they are close together, it is relatively easier to make the wrong decision.

If we denote the probability of choosing the right or wrong path at time k as P_r and P_w, then the log likelihood ratio (LLR) of choosing the right path over wrong path can be represented by the difference between the survivor and the loser path metrics, as referred in (2.7), (2.8),

$$\log\frac{P_r}{P_w} = \log P_r - \log P_w = \left|M_S - M_L\right| = \Delta_k . \tag{2.9}$$

Correspondingly,

$$P_w = \frac{1}{1+e^{\Delta_k}} \ . \tag{2.10}$$

Then let $\hat{P}_k$ denote the probabilities each symbol on the survivor path is wrong, then $\hat{P}_k$ is updated based on the rules in [20]:

$$\hat{P}_k = \hat{P}_k(1-P_w) + (1-\hat{P}_k)P_w \, . \tag{2.11}$$

The LLR for each symbol can be computed:

$$LLR(a_k) = \log\frac{1-\hat{P}_k}{\hat{P}_k} \, . \tag{2.12}$$

2.2.3. BCJR algorithm

The BCJR algorithm [21] is a maximum *a posteriori* (MAP) detection algorithm developed by Bahl *et al*. that minimizes the symbol error. It estimates the *a posteriori* probability (APP) of each symbol or bit in the binary case, based on the knowledge of the trellis, the observation of the received signal, and the *a priori* probability of each bit.

A. Algorithm for finding the APP

Let us still use the example of (2.1). The APP $P(a_k \,|\, \mathbf{r})$ can be expressed by the Bayes rule:

$$P(a_k \,|\, \mathbf{r}) = \frac{P(a_k, \mathbf{r})}{P(\mathbf{r})}. \tag{2.13}$$

Since a_k is determined by the transition from S_{k-1} to S_k, we can calculate the joint probability $P(S_{k-1}, S_k, \mathbf{r})$ instead, which can be decomposed as the following three items:

$$\begin{aligned} P(S_{k-1}, S_k, \mathbf{r}) &= P(S_{k-1}, r_1^{k-1}) P(S_k, r_k \mid S_{k-1}) P(r_{k+1}^N \mid S_k) \\ &= \alpha_{k-1}(S_{k-1}) \gamma_k(S_{k-1}, S_k) \beta_k(S_k), \end{aligned} \tag{2.14}$$

where $\gamma_k(S_{k-1}, S_k) \sim P(S_k, r_k \mid S_{k-1})$ denotes the branch transition probability,

$$\begin{aligned} P\left(S_k, r_k \mid S_{k-1}\right) &= \frac{P(S_k, r_k, S_{k-1})}{P(S_{k-1})} \\ &= P(S_k \mid S_{k-1}) P(r_k \mid S_k, S_{k-1}), \end{aligned} \tag{2.15}$$

where $P(S_k \mid S_{k-1}) \sim P(a_k)$ is the a priori probability of a_k, and $P(r_k \mid S_k, S_{k-1})$ can be evaluated from (2.14) for AWGN channel, then the branch metric

$$\gamma_k(S_{k-1}, S_k) = P(a_k) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\left(r_k - z_k\right)^2 \Big/ 2\sigma^2 \right]. \tag{2.16}$$

The other two items, $\alpha_k(S_k)$ and $\beta_k(S_k)$ are defined as the forward and backward transition probability, which can be computed recursively:

$$\alpha_k(S_k) = P(S_k, r_1^k) = \sum_{S_{k-1}} \alpha_{k-1}(S_{k-1}) \gamma_k(S_{k-1}, S_k), \tag{2.17}$$

$$\beta_k(S_k) = P(r_{k+1}^N \mid S_k) = \sum_{S_{k+1}} \beta_{k+1}(S_{k+1}) \gamma_{k+1}(S_k, S_{k+1}). \tag{2.18}$$

The algorithm begins with the forward recursion with the initial conditions: $\alpha_0(S_0 = S_0') = 1$, $\alpha_0(S_0 \neq S_0') = 0$ where S_0' is a known state at time 0. $\alpha_k(S_k)$ is updated with $\gamma_k(S_{k-1}, S_k)$ as time *k* goes from 1 up to *N*. When the forward

recursion is over, the backward recursion starts with $\beta_N(S_N = S'_N) = 1$, $\beta_N(S_N \neq S'_N) = 0$, and $\beta_k(S_k)$ is updated as *k* goes from *N* down to 1.

B. APP in the log domain

In practice, some of the above calculations are operated in the log domain in order to convert the multiplications into summations. The BCJR algorithm operating in the log domain is also called the log-MAP algorithm.

For instance we take (2.17) in the log domain and use the notation $\tilde{\alpha}_k(S_k) \sim \ln \alpha_k(S_k)$ as well as for other terms,

$$\begin{aligned} \ln \alpha_k(S_k) &= \ln \sum_{S_{k-1}} \exp\left[\ln \alpha_{k-1}(S_{k-1}) + \ln \gamma_k(S_{k-1}, S_k)\right] \\ \tilde{\alpha}_k(S_k) &= \ln \sum_{S_{k-1}} \exp\left[\tilde{\alpha}_{k-1}(S_{k-1}) + \tilde{\gamma}_k(S_{k-1}, S_k)\right]. \end{aligned} \tag{2.19}$$

Similarly,

$$\tilde{\beta}_k(S_k) = \ln \sum_{S_{k-1}} \exp\left[\tilde{\beta}_{k+1}(S_{k+1}) + \tilde{\gamma}_k(S_k, S_{k+1})\right]. \tag{2.20}$$

From (2.16),

$$\tilde{\gamma}_k(S_{k-1}, S_k) = \ln P(a_k) - \ln \sigma\sqrt{2\pi} - \frac{(r_k - z_k)^2}{2\sigma^2}. \tag{2.21}$$

Then the LLR can be calculated as below,

$$
\begin{aligned}
LLR(a_k) &= \ln \frac{P(a_k = +1/\mathbf{r})}{P(a_k = -1/\mathbf{r})} \\
&= \ln \frac{\sum_{+1} P(S_{k-1}, S_k, \mathbf{r})}{\sum_{-1} P(S_{k-1}, S_k, \mathbf{r})} \\
&= \ln \frac{\sum_{+1} \alpha_{k-1}(S_{k-1}) \gamma_k(S_{k-1}, S_k) \beta_k(S_k)}{\sum_{-1} \alpha_{k-1}(S_{k-1}) \gamma_k(S_{k-1}, S_k) \beta_k(S_k)} \\
&= \ln \sum_{+1} \exp\left[\tilde{\alpha}_{k-1}(S_{k-1}) + \tilde{\gamma}_k(S_{k-1}, S_k) + \tilde{\beta}_k(S_k) \right] \\
&\quad - \ln \sum_{-1} \exp\left[\tilde{\alpha}_{k-1}(S_{k-1}) + \tilde{\gamma}_k(S_{k-1}, S_k) + \tilde{\beta}_k(S_k) \right],
\end{aligned}
\tag{2.22}
$$

where the +1 and -1 denotes the state transition is triggered by $a_k = +1$ and $a_k = -1$ respectively.

BCJR returns the LLR as the soft information of each symbol for iterative decoding or detection. On the other hand, the hard decision of each bit can be made by observing the polarity of the LLR,

$$
\hat{a}_k = \begin{cases} +1, LLR(a_k) \geq 0 \\ -1, LLR(a_k), < 0 \end{cases} \tag{2.23}
$$

2.3. Partial Response Channel Equalization

The optimal detection is based on the trellis whose size is exponential on the length of the ISI. Naturally the channel ISI response is fairly long in a high density magnetic recording system, therefore reducing the ISI is very important to the detection performance. Linear zero-forcing equalization (ZFE) provides an equalizer with the inverse frequency response of the channel, such that the overall

channel response is forced to be ISI-free, but through filtering it also brings noise enhancement. Linear minimum mean-square error (MMSE) equalization makes a tradeoff between ISI and noise enhancement. These linear equalizers have infinite responses so they are not used in practice.

The alternative approach is equalizing the ISI channel to a partial response (PR) channel. Unlike the ZFE, the PR equalizer is not making the channel ISI-free, but shaping the channel to a finite-length PR target, which is usually much shorter than the original ISI channel. Then detection on the PR channel can be done by MLSD or MAP symbol-by-symbol (SBS) detectors, as mentioned in Section 2.2.

The design of the PR target is essential to the detection, and sometimes it is a tradeoff between performance and complexity. Short targets will surely bring equalization error, which is the difference between the equalized channel output and the PR channel output, but the number of trellis state is exponential on the memory of the PR channel. In the past people used to choose desired PR targets with integer values in low density magnetic recording systems, such as PR4 $(1-D^2)$, EPR4 $(1+D-D^2-D^3)$, E^{2}PR4 $\left(1+2D-2D^3-D^4\right)$, etc. Generalized PR (GPR) targets are preferred in higher density magnetic recording systems, as they approximate the channel more accurately and achieve better performance. The most effective solution for optimizing the GPR target is based on minimizing the mean-square-error (MSE) between the equalizer output and the PR channel output.

The conventional PMR system uses a single track equalizer to optimize the ISI channel into a 1-D GPR target with a 1-D equalizer. Here we provide a brief review

of the optimization of the single track equalizer, where the formulas can be found in [22].

The block diagram of a typical single-track equalizer is shown in Fig 2.2. The notations in the figure are described as follows: $a_k \in \{-1,1\}$ is the channel input data and y_k is the discrete signal from channel after filtering and sampling to the baud rate. $\mathbf{w} = [w_{-M}, \cdots, w_0, \cdots, w_M]^T$ denotes the (2M+1)-tap finite impulse response (FIR) filter, and $\mathbf{t} = [t_0, \cdots, t_I]^T$ denotes the GPR target with length *I*+1.

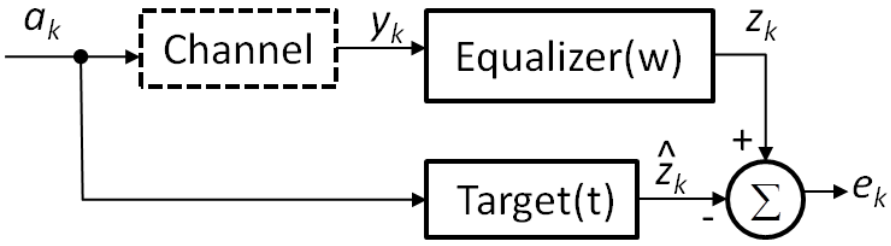

Fig 2.2. Block diagram of a single-track equalizer.

The equalization error e_k is the difference between equalizer output z_k and the PR channel output $\hat{z}_k$:

$$e_k = z_k - \hat{z}_k = \sum_{i=-N}^{N} w_i y_{k-i} - \sum_{i=0}^{I} t_i a_{k-i} \,. \tag{2.24}$$

The MSE of e_k can be then expressed in the matrix form,

$$\begin{aligned} E\left\{|e_k|^2\right\} &= E\left\{\|z_k - \hat{z}_k\|^2\right\} \\ &= E\left\{z_k^2\right\} + E\left\{\hat{z}_k^2\right\} - 2E\left\{z_k \hat{z}_k\right\} \\ &= \mathbf{w}^{\mathbf{T}}\mathbf{R}_{\mathbf{y}}\mathbf{w} + \mathbf{t}^{\mathbf{T}}\mathbf{R}_{\mathbf{a}}\mathbf{t} - \mathbf{2}\mathbf{w}^{\mathbf{T}}\mathbf{R}_{\mathbf{y,a}}\mathbf{t}, \end{aligned} \tag{2.25}$$

where $\mathbf{R}_{\mathbf{y}} = E\left\{\mathbf{y}\mathbf{y}^{\mathbf{T}}\right\}$ and $\mathbf{R}_{\mathbf{a}} = E\left\{\mathbf{a}\mathbf{a}^{\mathbf{T}}\right\}$ are autocorrelation matrices of **y** and **a**, respectively, and $\mathbf{R}_{\mathbf{y,a}}$ is the cross-correlation matrix between **y** and **a,** where $\mathbf{R}_{\mathbf{y,a}}(i,j) = E\left\{y_{k-i}a_{k-j}\right\}$, $-N \le i \le N$, $0 \le j \le I$.

Now we need to find **w** and **t** to minimize the expression in (2.25), but that may lead to a trivial solution: $\mathbf{w} = \mathbf{t} = \mathbf{0}$. In order to avoid this situation, a common strategy is setting up a monic constraint of forcing $\mathbf{t}_0$=1, which is done by introducing a term $2\lambda\left(\mathbf{C}^{\mathbf{T}}\mathbf{t} - 1\right) = 0$, where vector $\mathbf{C} = [1, 0, \cdots, 0]^T$ of length $I+1$ and λ is a scalar.

Then (2.25) can be rewritten as

$$E\left\{\left|e_k\right|^2\right\} = \mathbf{w}^{\mathbf{T}}\mathbf{R}_{\mathbf{y}}\mathbf{w} + \mathbf{t}^{\mathbf{T}}\mathbf{R}_{\mathbf{a}}\mathbf{t} - 2\mathbf{w}^{\mathbf{T}}\mathbf{R}_{\mathbf{y,a}}\mathbf{t} - 2\lambda\left(\mathbf{C}^{\mathbf{T}}\mathbf{t} - 1\right), \tag{2.26}$$

To obtain the minimum MSE, we take the derivatives with respect to w_k, t_k and λ, and set them to zero respectively,

$$\frac{\partial E\left\{\left|e_k\right|^2\right\}}{\partial w_k} = 2\mathbf{R}_{\mathbf{y}}\mathbf{w} - 2\mathbf{R}_{\mathbf{y,a}}\mathbf{t} = 0 \tag{2.27}$$

$$\frac{\partial E\left\{\left|e_k\right|^2\right\}}{\partial t_k} = 2\mathbf{R}_{\mathbf{a}} - 2\mathbf{w}^{\mathbf{T}}\mathbf{R}_{\mathbf{y,a}} - 2\lambda\mathbf{C}^{\mathbf{T}} = 0 \tag{2.28}$$

$$\frac{\partial E\left\{\left|e_k\right|^2\right\}}{\partial \lambda} = -2\left(\mathbf{C}^{\mathbf{T}}\mathbf{t} - 1\right) = 0, \tag{2.29}$$

and the following parameters can be derived as:

$$\lambda = \frac{1}{\mathbf{C}^{\mathbf{T}}\left(\mathbf{R}_{\mathbf{a}} - \mathbf{R}_{\mathbf{y,a}}^{\mathbf{T}}\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{R}_{\mathbf{y,a}}\right)^{-1}\mathbf{C}} \tag{2.30}$$

$$\mathbf{t} = \lambda\left(\mathbf{R}_a - \mathbf{R}_{y,a}^T \mathbf{R}_y^{-1} \mathbf{R}_{y,a}\right)^{-1} \mathbf{C} \tag{2.31}$$

$$\mathbf{w} = \mathbf{R}_y^{-1} \mathbf{R}_{y,a} \mathbf{t}\,. \tag{2.32}$$

2.4. Conclusion

In this chapter we revisited the fundamentals of a magnetic recording system, conventional PR equalization and 1-D detection algorithms. In the rest of this dissertation, we choose to use BCJR detector as the optimal SBS detector in order to use the soft decisions and minimize the symbol error rate. As the PR target is relatively short (3 or 4 taps), the complexity of the detector is still manageable.

Here we denote the single-track detector (STD) as the detection scheme of using single-track equalizer followed by a standard BCJR detector, as shown in Fig 2.3.

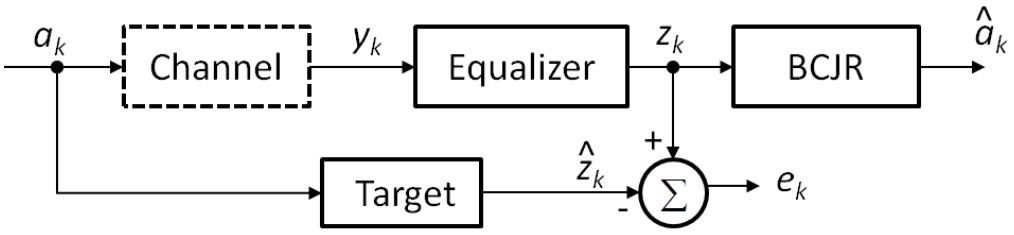

Fig 2.3. Block diagram of the STD.

STD is a simple and efficient way of performing 1-D PMR channel detection, which can also work with ECC decoders for better data recovery. In the next chapters, we treat STD as the benchmark in various computer simulations, and we will introduce more sophisticated 2-D equalization and detection schemes.

Chapter 3 ITI Mitigating Detection on the TDMR Channel Model

Previous research has shown that PR equalization followed by a BCJR detector, which we denote as the STD is a practical and sufficient solution for channel detection on the 1-D PMR channel. When moving from PMR to TDMR, the most important job for TDMR channel detectors is to cope with 2-D ISI, especially the severe ITI on the cross-track direction. An intuitive way is to expand the idea of 1-D equalization into the 2-D case, such as optimizing 2-D GPR targets on the side tracks, and using the side track information to mitigate the effect of ITI.

In the following content, we will discuss several PR equalization techniques such as 2-D equalization and joint-track equalization, and then their applications on TDMR channel detection, such as MTD and ITI cancellation, giving simulations results compared with STD.

3.1. Two-Dimensional Equalization and Optimization

3.1.1. 2-D equalization

2-D equalization has been applied in BPMR channel detection by Nabavi and Kumar [23]. It uses 1-D FIR equalizers on each track, and so the group of 1-D equalizers can be considered as a 2-D equalizer. On the other hand, the GPR target is constrained to 1-D by forcing the side track targets to zero so as to avoid the high complexity of a 2-D detector. This method of using 2-D equalizer and 1-D GPR target is later called “2D1D” equalization.

Note that this technique is still for single track detection. The reason for choosing 2-D FIR equalizer is to get a smaller MSE. But the constraint of 1-D GPR target can be removed, where the GPR targets for side tracks could have non-zero values, and denoted as "2D2D" equalization. This architecture needs a multi-input detector such as the joint-BCJR detector [22].

An example of three-track "2D2D" equalizer is shown in Fig 3.1, where track 0 is the main track, and track -1,+1 are two adjacent tracks. The equalizers on the three tracks are denoted by $\mathbf{w}_{-1} = \left[w_{-1,-N}, \ldots, w_{-1,N} \right]^T$, $\mathbf{w}_0 = \left[w_{0,-N}, \ldots, w_{0,N} \right]^T$ and $\mathbf{w}_1 = \left[w_{1,-N}, \ldots, w_{1,N} \right]^T$, each with length 2N+1. These equalizers can be defined together as a vector $\mathbf{w} = \left[\mathbf{w}_{-1}^{\mathbf{T}}, \mathbf{w}_{\mathbf{0}}^{\mathbf{T}}, \mathbf{w}_{\mathbf{1}}^{\mathbf{T}} \right]^{\mathbf{T}}$ of length 3(2N+1). The GPR target on the center track is $\mathbf{t}_0 = \left[t_{0,0}, \ldots, t_{0,L-1} \right]^T$ with length L, the targets on the side tracks are $\mathbf{t}_{-1} = \left[t_{-1,0}, \ldots, t_{-1,L'-1} \right]^T$ and $\mathbf{t}_1 = \left[t_{1,0}, \ldots, t_{1,L'-1} \right]^T$, both with length L'. In the same say, the GPR targets can be expressed together as $\mathbf{t} = \left[\mathbf{t}_{-1}^{\mathbf{T}}, \mathbf{t}_{\mathbf{0}}^{\mathbf{T}}, \mathbf{t}_{\mathbf{1}}^{\mathbf{T}} \right]^{\mathbf{T}}$ of length L+2L'. Correspondingly, the channel input $\mathbf{a} = \left[\mathbf{a}_{-1}^{\mathbf{T}}, \mathbf{a}_{\mathbf{0}}^{\mathbf{T}}, \mathbf{a}_{\mathbf{1}}^{\mathbf{T}} \right]^{\mathbf{T}}$ of length L+2L' and output $\mathbf{y} = \left[\mathbf{y}_{-1}^{\mathbf{T}}, \mathbf{y}_{\mathbf{0}}^{\mathbf{T}}, \mathbf{y}_{\mathbf{1}}^{\mathbf{T}} \right]^{\mathbf{T}}$ of length 3(2N+1) are also defined in a vector form.

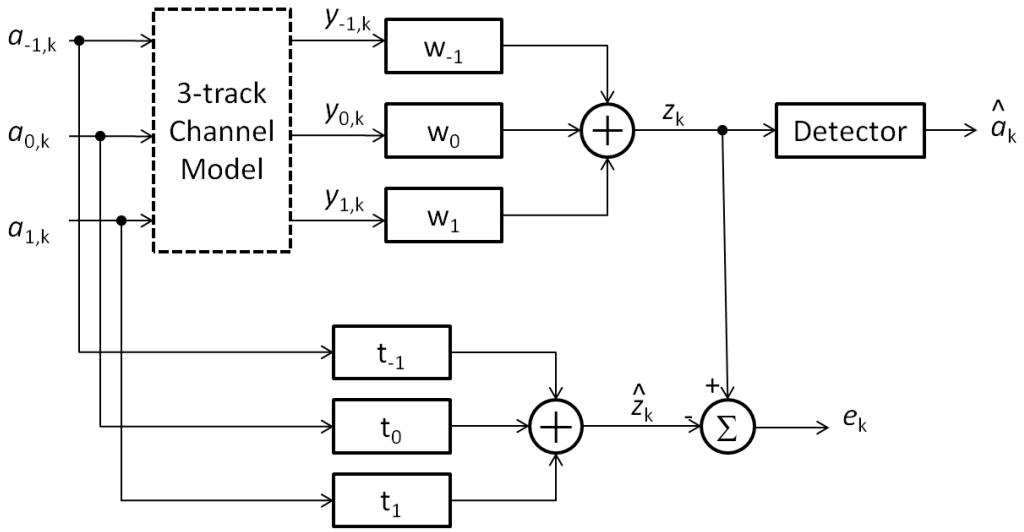

Fig 3.1. 2-D equalization with 2-D GPR targets on a three-track model.

The equalization error is $e_k = z_k - \hat{z}_k = \mathbf{w^T y - t^T a}$, and the MSE

$$E\left\{\left|e_k\right|^2\right\} = \mathbf{w^T R_y w} + \mathbf{t^T R_a t - 2 w^T R_{y,a} t}\,. \tag{3.1}$$

By using the same method of monic constraint used in (2.26) to minimize the MSE,

$$\lambda = \frac{1}{\mathbf{C^T}\left(\mathbf{R_a - R_{y,a}^T R_y^{-1} R_{y,a}}\right)^{-1}\mathbf{C}}, \tag{3.2}$$

$$\mathbf{t} = \lambda\left(\mathbf{R_a - R_{y,a}^T R_y^{-1} R_{y,a}}\right)^{-1}\mathbf{C}\,, \tag{3.3}$$

$$\mathbf{w} = \mathbf{R_y^{-1} R_{y,a} t}\,, \tag{3.4}$$

where $\mathbf{C} = [0,\ldots,0,1,0,\ldots,0]^T$ of length L+2L' is a similar constraint vector as in the single-track equalizer in order to make the center track target t_0 with the first coefficient equal to one. No wonder that (3.2)-(3.4) resemble (2.30)-(2.32). But the

difference is the matrices in the above equations are made up with vectors from multiple tracks.

3.1.2. Joint-track equalization

Joint-track equalization [24] can be considered as a special case in 2-D equalization. It optimizes the channel using a 1-D equalizer and 2-D GPR target. The 1-D FIR equalizer is just like the one in the STD, but the 2-D GPR target is designed for both center and side tracks. Since the trellis is two-dimensional, it also requires multi-input detectors.

The 2-D GPR target of the joint-track equalization is illustrated in Fig. 3.2.

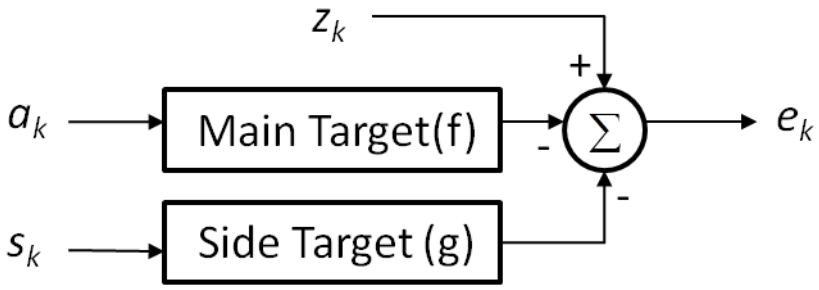

Fig 3.2. 2-D GPR target for joint-track equalization.

Suppose the ITI comes from one side track with the written data s_k, and the GPR target of main track is denoted by $\mathbf{f} = [f_0, \cdots, f_{L-1}]^T$ of length L and the GPR target of the side track is denoted by $\mathbf{g} = [g_0, \cdots, g_{L'-1}]^T$ of length L', where L may not equal L'.

The equalization error and MSE are expressed as

$$\begin{aligned} e_k &= z_k - (a * f)_k - (s * g)_k \\ &= \sum_{i=-N}^{N} w_i y_{k-i} - \sum_{i=0}^{L-1} f_i a_{k-i} - \sum_{i=0}^{L'-1} g_i s_{k-i}, \end{aligned} \tag{3.7}$$

$$\begin{aligned} E\left\{|e_k|^2\right\} = \mathbf{w}^{\mathbf{T}}\mathbf{R}_{\mathbf{y}}\mathbf{w} + \mathbf{f}^{\mathbf{T}}\mathbf{R}_{\mathbf{a}}\mathbf{f} + \mathbf{g}^{\mathbf{T}}\mathbf{R}_{\mathbf{s}}\mathbf{g} \\ +2\mathbf{f}^{\mathbf{T}}\mathbf{R}_{\mathbf{a,s}}\mathbf{g} - 2\mathbf{w}^{\mathbf{T}}\mathbf{R}_{\mathbf{y,a}}\mathbf{f} - 2\mathbf{w}^{\mathbf{T}}\mathbf{R}_{\mathbf{y,s}}\mathbf{g}. \end{aligned} \tag{3.8}$$

To minimize the MSE, we use the same method as in single-track equalization by taking the derivative of MSE with respect to λ, f_k, g_k and w_k and enforcing $f_0 = 1$. The coefficients can be obtained as in [24] and [25]:

$$\lambda = \frac{\mathbf{1}}{\mathbf{C}^{\mathbf{T}}\left(\mathbf{A}_{\mathbf{1}} - \mathbf{B}_{\mathbf{1}}\mathbf{A}_{\mathbf{2}}^{-\mathbf{1}}\mathbf{B}_{\mathbf{2}}\right)^{-\mathbf{1}}\mathbf{C}}, \tag{3.9}$$

$$\mathbf{f} = \lambda\left(\mathbf{A}_{\mathbf{1}} - \mathbf{B}_{\mathbf{1}}\mathbf{A}_{\mathbf{2}}^{-\mathbf{1}}\mathbf{B}_{\mathbf{2}}\right)^{-\mathbf{1}}\mathbf{C}, \tag{3.10}$$

$$\mathbf{g} = -\mathbf{A}_{\mathbf{2}}^{-\mathbf{1}}\mathbf{B}_{\mathbf{2}}\mathbf{f}, \tag{3.11}$$

$$\mathbf{w} = \mathbf{R}_{\mathbf{y}}^{-\mathbf{1}}\left(\mathbf{R}_{\mathbf{y,a}}\mathbf{f} + \mathbf{R}_{\mathbf{y,s}}\mathbf{g}\right), \tag{3.12}$$

$$\text{where } \begin{cases} \mathbf{A}_{\mathbf{1}} = \mathbf{R}_{\mathbf{a}} - \mathbf{R}_{\mathbf{y,a}}^{\mathbf{T}}\mathbf{R}_{\mathbf{y}}^{-\mathbf{1}}\mathbf{R}_{\mathbf{y,a}} \\ \mathbf{A}_{\mathbf{2}} = \mathbf{R}_{\mathbf{s}} - \mathbf{R}_{\mathbf{y,s}}^{\mathbf{T}}\mathbf{R}_{\mathbf{y}}^{-\mathbf{1}}\mathbf{R}_{\mathbf{y,s}} \\ \mathbf{B}_{\mathbf{1}} = \mathbf{R}_{\mathbf{a,s}} - \mathbf{R}_{\mathbf{y,a}}^{\mathbf{T}}\mathbf{R}_{\mathbf{y}}^{-\mathbf{1}}\mathbf{R}_{\mathbf{y,s}} \\ \mathbf{B}_{\mathbf{2}} = \mathbf{R}_{\mathbf{a,s}}^{\mathbf{T}} - \mathbf{R}_{\mathbf{y,s}}^{\mathbf{T}}\mathbf{R}_{\mathbf{y}}^{-\mathbf{1}}\mathbf{R}_{\mathbf{y,a}} \end{cases}. \tag{3.13}$$

Alternatively, [25] has pointed out that the joint-track equalizer is actually a special case of 2D2D equalization when $\mathbf{w}_{-\mathbf{1}}$ and $\mathbf{w}_{\mathbf{1}}$ are set to zero, while $\mathbf{t}_{-\mathbf{1}}$ and $\mathbf{t}_{\mathbf{1}}$

have non-zero values. Then the coefficients can be computed in a general form of (3.2)-(3.4).

3.2. Multi-track Detection on TDMR

Although people use different equalization methods to shape the channel more accurately and get smaller equalization error, they usually do not use all the available information to get more performance gain, because ITI is not a random noise but signal that contains information. Based on the above equalization techniques, Chang and Cruz proposed an MTD method [26], where they explained the effective MSE of a 2D equalizer is actually larger than the MSE in a single track equalizer. But if the side track data is perfectly known, the actual MSE of the 2D equalization can be reduced to the MSE. In reality the side track data is surely unknown, either the hard decision or soft APP of the data on adjacent tracks can be pre-detected, and can be used as *a priori* information for the detection on the main track.

MTD requires the read channel to be equalized to a 2-D GPR target, but the choice of equalizers could be either single-input or multi-input, which we refer to as the joint-track equalizer and a 2-D equalizer. In the following example we only consider the joint-track equalizer because it has lower complexity and similar performance as the 2D2D equalizer. Suppose that the TDMR model has three adjacent tracks, we use the same strategy in [26] to lower the detector complexity with the following assumptions:

a. Two side tracks have the same impulse response.

b. All three-track data are recorded with perfect synchronization.

Based on those assumptions, the binary data on the two side tracks can be combined as a sequence consisting of ternary data {-2, 0, 2}, denoted by d_k. Then the two side tracks can share the same GPR target, and the optimization is simplified to one side track equalization. Since the trellis is expanded to $\{b_k, d_k\}$, this allows us to reduce the overall trellis complexity as well, with each state having $2\times 3 = 6$ outgoing branches instead of $2\times 2\times 2 = 8$ branches, if we were to treat the two side tracks individually.

Finally, the equalized signal is detected using a joint-BCJR detector, which is explained in [24]. In this case, the branch metric is modified by taking the information on the side tracks. The forward and backward transition probability stays the same as (2.17)-(2.18), but the branch transition probability becomes:

$$\gamma_k(S_{k-1}, S_k) = P(b_k, d_k)P(r_k \mid S_{k-1}, S_k, b_k, d_k), \tag{3.14}$$

where $P(b_k, d_k) = P(b_k)P(d_k)$ due to the independence between b_k and d_k. The APP of b_k can be computed by marginalization:

$$P(b_k \mid \mathbf{r}) = \sum_{d_k} P(b_k, d_k \mid \mathbf{r}). \tag{3.15}$$

The block diagram of the MTD we are applying is shown in Fig. 3.3, where the index N refers to the N-th track, $\mathbf{b}_N$ denotes the binary input data block on track N (main track) and $\mathbf{d}_N$ denotes the combined sequence of side-track data $\mathbf{b}_{N-1}$ and $\mathbf{b}_{N+1}$, which are detected by STD separately.

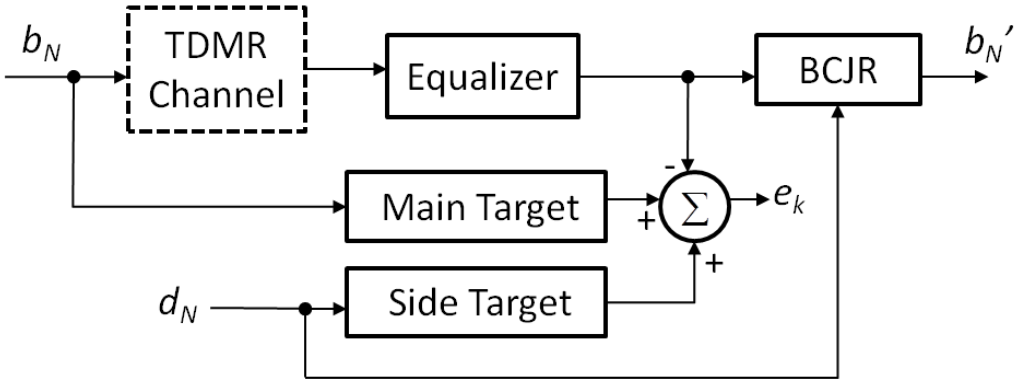

Fig 3.3. Multi-track detector with joint-track equalization.

It has been shown that MTD has significant performance gains against STD on BPMR channel models [26], [27], and it will be interesting to see its performance on the TDMR channels.

3.3. ITI Cancellation

ITI cancellation has been applied to SMR channels [28], where only the ITI from one side track was considered and estimated by reading the tracks in the reverse order of writing, and assuming the availability of a non-shingled track per writing block. Another work [29] has tested both single-sided and double-sided ITI cancellation for different squeeze ratios, which shows double-sided ITI cancellation performs better when we have a high squeeze ratio.

In our channel model, we consider all tracks to be shingled and each track is interfered by two adjacent tracks. Our application of double-sided ITI cancellation is estimating the ITI from both tracks and subtracting them from the readback

signal, but it does not require reading in reverse order, instead, data on the side tracks should be pre-read and detected. Fig. 3.4 illustrates the application of our ITI canceller for a three-track example.

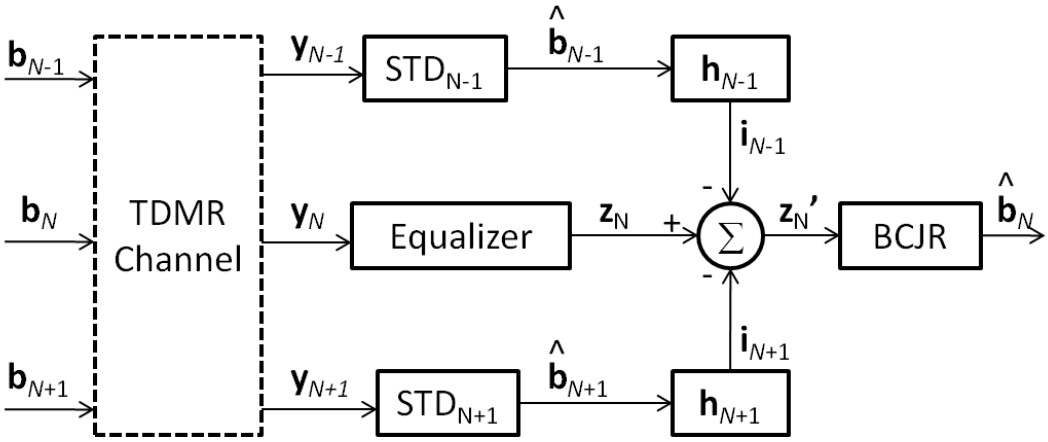

Fig 3.4. ITI canceller for a three-track TDMR channel model

We first pre-compute the ITI response $\mathbf{h}_{N-1}$ by cross-correlating the model input and output signals of the training sequence on track *N*-1. Then the hard decision of the side track data $\hat{\mathbf{b}}_{N-1}$ is detected by STD. The ITI signal $\mathbf{i}_{N-1}$ is estimated by the convolution of $\mathbf{h}_{N-1}$ and $\hat{\mathbf{b}}_{N-1}$:

$$\mathbf{i}_{N-1} = \mathbf{h}_{N-1} * \hat{\mathbf{b}}_{N-1}. \tag{3.16}$$

We repeat the same process on track *N*+1 to find $\mathbf{i}_{N+1}$, and subtract them from the equalizer output $\mathbf{z}_N$:

$$\mathbf{z}'_N = \mathbf{z}_N - \mathbf{i}_{N-1} - \mathbf{i}_{N+1}. \tag{3.17}$$

Finally the ITI cancelled readback signal $\mathbf{z}'_N$ will be processed by the single input

BCJR detector, the structure of ITI canceller is actually simpler and requires less computational complexity than MTD.

Note that ITI signals $\mathbf{i}_{N-1}, \mathbf{i}_{N+1}$ have to be consistent with the signal on the main track. In our example, we subtract the ITI signal after the equalization process, so we use the equalized output on the side tracks to estimate the ITI responses. Alternatively, if we choose to put the equalizer after the ITI subtraction, then the estimation should not include equalization. Actually these two approaches makes very little difference.

3.4. MTD Simulations on a Squeezed PMR Model

Before simulating on the TDMR channel model, we first test the BER results on the PMR channel with ITI. The channel model is a 3-track PMR model with AWGN and first order jitter noise, and the channel is "squeezed" such that the center track signal is attenuated and significant ITI is introduced from the two side tracks.

In order to better quantify the amount of ITI, we introduce a parameter called head-share-ratio, denoted by α for the portion of reading the side track with respect of the whole head width. Suppose the normalized amplitude of the channel response is 1, since the amplitude of the response on adjacent tracks is α, the amplitude of the channel response on the main track is 1-α. If we assume there are two symmetric adjacent tracks, the amplitude of the response on each side track is $\alpha/2$, the channel output is $y = (1-\alpha)y_c + (\alpha/2)y_1 + (\alpha/2)y_2$, as shown in Fig 3.5.

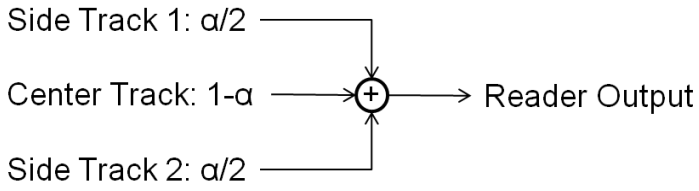

Fig 3.5. Channel output in a squeezed PMR model.

The Monte-carlo simulation on the squeezed PMR channel model is set up on channel density D = 1.1 and head share ratio = 0.3. The BER performance comparison between the traditional STD and MTD is shown in Fig 3.6.

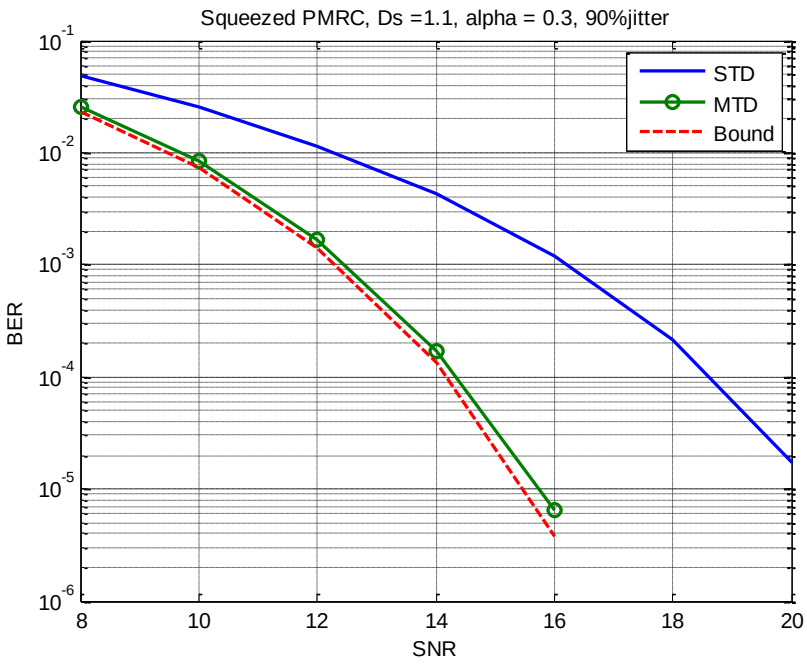

Fig 3.6. Performance results for a squeezed PMR channel with Ds =1.1, α =0.3, 90% jitter noise.

Since the MTD takes advantage of the side track signals, it outperforms the STD substantially. The performance bound in Fig. 3.6 shows the performance of the MTD when the data on the side tracks are perfectly known, and the MTD based on actual detected side track data is getting very close to the bound. It will be interesting to see the performance results on the TDMR channel model as well.

3.5. Simulations on a TDMR Model

For each data configuration, we use the Voronoi-based TDMR model introduced in Chapter 1 to generate two independent sets of data consisting of three adjacent 32-kb sectors. One data set, which contains the input/output data on 3 adjacent tracks without any offset, is used to train the equalizer and optimize the targets for the MTD, as well as estimating the ITI responses. The other set, containing 3 tracks of data, including the readback signals with the read head offset on the cross-track direction, is used to obtain BER performance results.

In our TDMR channel model, once the channel parameters such as channel density and BAR are set up, the channel SNR is fixed. Therefore, instead of testing BER vs. SNR result, we are more interested in the BER vs. read head offset, illustrated by the "bathtub curve". In an SMR system with very narrow tracks, the offset of a read head is likely to happen. Such curve can show how much performance gain of the ITI mitigation detectors can get at a certain offset, or how much tolerance to the read head offset of a certain detector.

The channel model is set at a density of 1Tb/in^2 and the BAR at 2:1, without any error correcting code. The reason for selecting this BAR is to increase the ITI so that we can evaluate the performance of the detectors under more challenging conditions. The amount of ITI is illustrated in Fig 3.7, where the amplitudes are extracted from baud-rate sampled data.

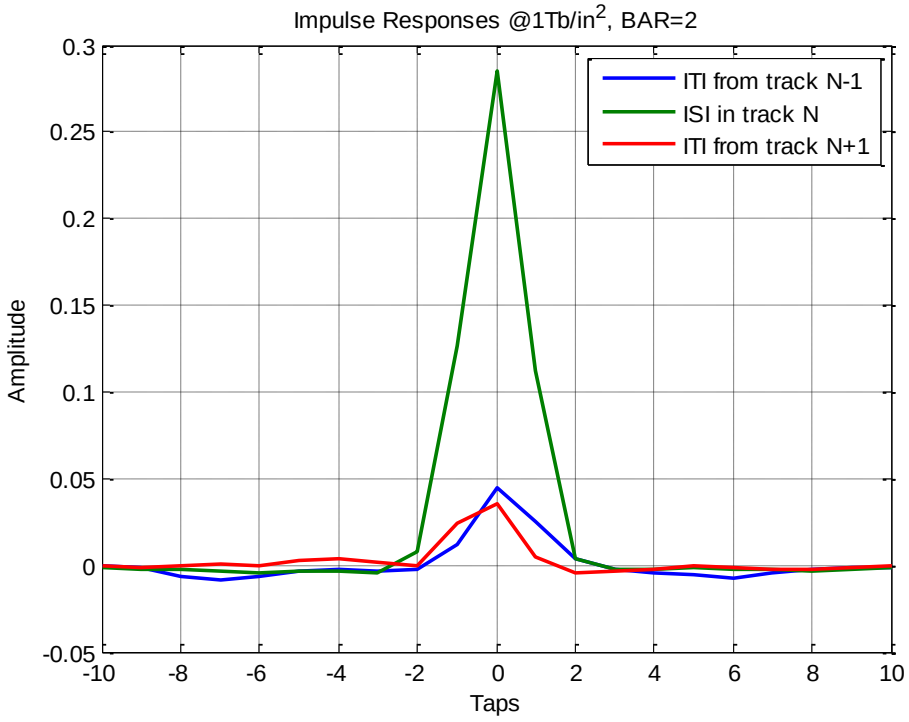

Fig 3.7. ITI responses of the TDMR channel model at 1Tb/in^2, BAR=2.

The plot of "bathtub curves" using STD, MTD, and ITI canceller is shown in Fig. 3.8 where the results are obtained by offsetting the center track reader and keeping the side-track readers fixed.

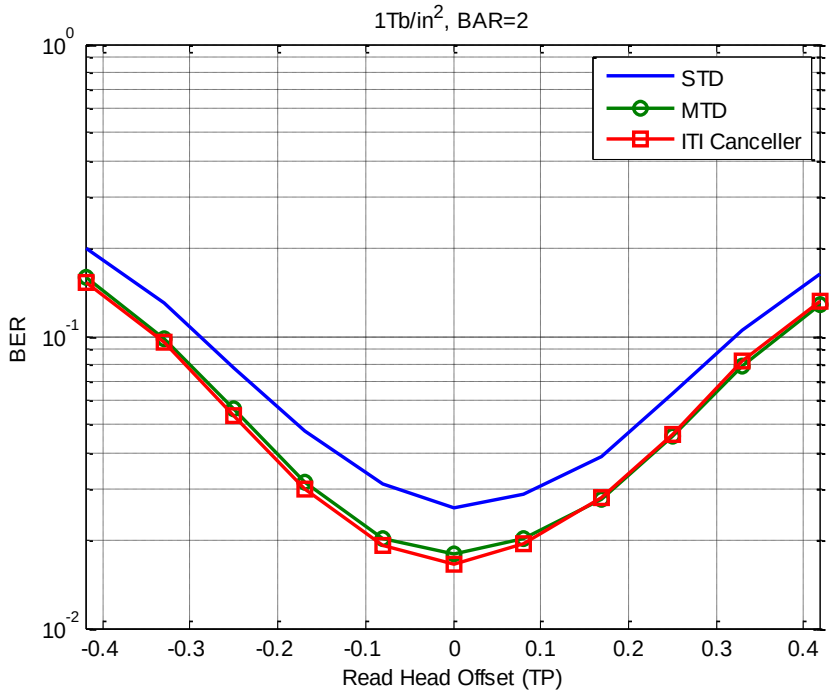

Fig 3.8. Performance results for various read head offsets of the center track

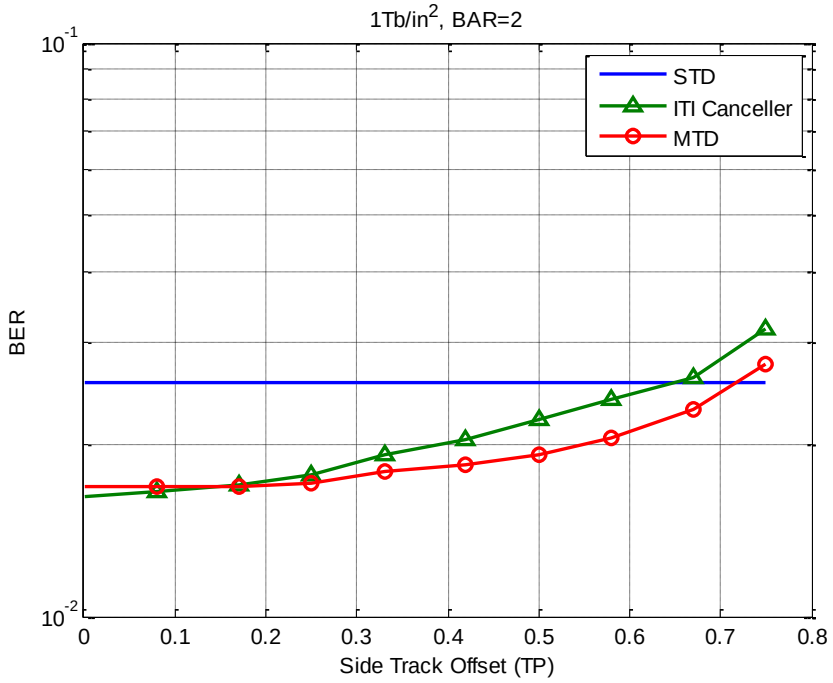

Fig 3.9. Performance results for various read head offsets of the side track

We also test the results with center track reader fixed and moving two side tracks moving toward the center symmetrically, as shown in Fig 3.9.

The performance results showed that both the MTD and the ITI canceller exhibit significant performance gain over the STD, where the gain comes from the pre-detected side track information.

One might notice that the ITI canceller performs slightly better than the MTD at some offsets, because it takes asymmetric side-track responses into consideration, while the MTD assumes that the two side tracks have the same response. However, the MTD still has some advantages compared with the ITI canceller, because the ITI canceller only takes hard decision on the side tracks, while the MTD can take soft information from the side tracks, which might be the reason it has more tolerance to side track offsets.

3.6. Conclusion

In this chapter we focused on solving the problem of 2-D ISI in TDMR channels by using ITI mitigation detectors like MTD and the ITI canceller. Both methods require pre-reading and pre-detection process, which can be done by multiple head or multiple reading by a single head. In order to simulate the real world situation and maintain the overall complexity, we choose STD for detecting the side tracks. The detection it also based on the assumption of perfect synchronization between the tracks. The MTD and ITI canceller take advantage of side track information in different ways. The MTD uses side-track information in the BCJR to improve the

decisions, while the ITI canceller uses side-track data to estimate the ITI, and subtract it from the readback signal to give a cleaner signal for the detector. These methods are suboptimal approaches, but the simulation results showed some significant performance gains in the squeezed PMRC and in Voronoi TDMR channel model at low SNR.

Chapter 4 Noise Predictive Channel Detection

ITI is not the only issue in a 2-D magnetic recording system. Another factor that impairs the performance of the channel detector is the noise, which consists of electronic thermal noise and media noise, where the media noise is dominant at high density systems. In this chapter we focus on the noise predictive detection techniques to suppress the effect of media noise with simulation results on PMRC and TDMR channel models.

4.1. Introduction

Optimal detection on an ISI channel with AWGN can be done by using a Viterbi or BCJR detector. Usually the electronic thermal noise in a magnetic recording channel is modeled as AWGN, which is independent OF the transmitted signals. However, in the PR equalization, the white Gaussian noise goes through the FIR equalizer and becomes correlated, such that those optimal detectors are not optimal in the presence of colored noise. Noise-predictive maximum-likelihood (NPML) techniques are [30] is an effective way of whitening the colored noise, by using the linear combination of past noise samples to predict the current noise sample.

At higher densities, the media noise plays a dominant role in the noise component. In a discrete BPMR channel, the position and shape variation of the discrete islands is the main source of the media noise. In continuous magnetic recording media like in PMR or TDMR, since the media is made up of random

magnetic grains, the border between each bit magnetization is of a zigzag shape, such that there is jitter in the timing of data transitions.

If we recall from Chapter 1, we modeled the media noise in the PMRC model as the first order position jitter noise in (1.4) and (1.5), which is related to the transition sequence. Therefore it depends on particular input data patterns. The idea is intuitive, for example, if the input patter is [-1,-1,…,-1] or [+1,+1,…,+1], there should be no jitter noise, on the other hand if the input pattern is alternate -1s and +1s, there will be much more jitter noise because of the large number of data transitions. In practice the interval of data transitions is controlled by RLL or modulation codes.

Pattern-dependent noise prediction (PDNP) [31], which is an extension of NPML, not only whitens the colored noise, but also takes care of pattern-dependent media noise effectively. Not only does it work on the MLSD, it also works on the SBS soft-output MAP detector. Since the media noise is correlated with the input pattern, which means it contains useful information, by modifying the branch metrics in the Viterbi or BCJR detector, the performance of channel detection can be improved.

4.2. Pattern-Dependent Noise Prediction

In this section we review the process described in [31]. First of all the predicted noise sample $\hat{n}_k$ depends on the linear combination of L previous noise samples, where we denote L as the prediction order or prediction taps.

The noise sample n_k is extracted as the difference between the channel output r_k and the noiseless output z_k.

$$n_k = r_k - z_k, \tag{4.1}$$

The channel output r_k is always available, but the noiseless channel outputs depend on z_k and its past L values for the specific bit pattern $\mathbf{b} = [b_{k-L-I}, \cdots, b_k]$ where I is the PR channel memory.

Therefore the noise vector $\mathbf{n}(\mathbf{b}) = [n_{k-1}(\mathbf{b}), n_{k-2}(\mathbf{b}), \cdots, n_{k-L}(\mathbf{b})]^T$ is a pattern dependent vector, and the predicted noise sample can be represented as an autoregressive process:

$$\begin{aligned} \hat{n}_k(\mathbf{b}) &= q_1(\mathbf{b}) n_{k-1} + q_2(\mathbf{b}) n_{k-2} + \cdots + q_L(\mathbf{b}) n_{k-L} \\ &= \sum_{i=1}^{L} q_i(\mathbf{b}) n_{k-i}, \end{aligned} \tag{4.2}$$

where $\mathbf{q}(\mathbf{b}) = [q_1(\mathbf{b}), q_2(\mathbf{b}), \cdots, q_L(\mathbf{b})]$ is the noise prediction coefficient vector, and we denote the noise correlation vector by $\mathbf{c}(\mathbf{b}) = E\{n_k(\mathbf{b})\mathbf{n}(\mathbf{b})\}$ and the noise correlation matrix by $\mathbf{R}(\mathbf{b}) = E\{\mathbf{n}(\mathbf{b})\mathbf{n}(\mathbf{b})^T\}$,

By applying the Yule-Walker equations to the autoregressive model (4.2), the predictor coefficients and the prediction variance can be obtained for each data pattern:

$$\mathbf{q}(\mathbf{b}) = \mathbf{c}(\mathbf{b})^T \mathbf{R}^{-1}(\mathbf{b}), \tag{4.3}$$

$$\sigma_p{}^2(\mathbf{b}) = \sigma^2 + \mathbf{c}(\mathbf{b})^T \mathbf{R}^{-1}(\mathbf{b}) \mathbf{c}(\mathbf{b}). \tag{4.4}$$

Unlike the AWGN channel, the expected channel output using PDNP will be the noiseless output plus the predicted noise.

$$\hat{r}_k(\mathbf{b}) = z_k(\mathbf{b}) + \hat{n}_k(\mathbf{b}). \tag{4.5}$$

If using a noise predictive Viterbi detector, the branch metric is modified by using (4.5) and replacing σ^2 with $\sigma_p{}^2(\mathbf{b})$

$$\lambda_k = \frac{1}{\sqrt{2\pi\sigma_p(\mathbf{b})^2}} \exp\left[\frac{-\left(r_k - \hat{r}_k(\mathbf{b})\right)^2}{2\sigma_p(\mathbf{b})^2}\right]. \tag{4.6}$$

Similarly, the branch transition probability can be modified for the noise predictive BCJR,

$$\gamma_k(S_{k-1}, S_k) = \frac{1}{\sqrt{2\pi\sigma_p(\mathbf{b})^2}} \exp\left[\frac{-\left(r_k - \hat{r}_k(\mathbf{b})\right)^2}{2\sigma_p(\mathbf{b})^2}\right] P(a_k). \tag{4.7}$$

Note that the size of the trellis in the VA or BCJR is expanded when using PDNP by the number of prediction taps, the total number of states is 2^{I+L}.

In the following experiments we only consider the BCJR as our channel detector. Based on our simulation environment, we need to execute the following steps to obtain the pattern dependent noise prediction.

1. Find the noise sequence **n** using (4.1) based on a random training sequence.
2. Find **c(b)** and **R(b)** by using a sliding window of each data pattern moving on the input sequence, if there is a match, record the noise vector **n(b)**, then the expectations are calculated by averaging the values by the number of total matches on that sequence. Repeat this step for all data patterns.

3. Find **q(b)** and $\sigma_p^{\,2}(\mathbf{b})$ by substituting **c(b)** and **R(b)** into (4.3) and (4.4), stored in a lookup table.
4. Find the predicted noise $\hat{n}_k(\mathbf{b})$ by substituting **q(b)** and $\sigma_p^{\,2}(\mathbf{b})$ into (4.2).
5. In the PDNP-BCJR, the forward probability $\alpha_k(S_k)$ and backward probability $\beta_k(S_k)$ stay the same as (2.17) and (2.18), but the transition probability $\gamma_k(S_{k-1}, S_k)$ needs to be modified as in (4.7).

4.3. Simulation Results

First, let us test the PDNP algorithm on the conventional PMR channel model given in Section 1.1. The system SNR is defined as

$$SNR = \frac{E_b}{N_0 + M_0}, \tag{4.8}$$

where M_0 is the jitter noise power, and N_0 is the power of AWGN. The jitter noise percentage is defined as $\frac{M_0}{M_0 + N_0} \times 100\%$.

We've tested the noise predictive BCJR with two different jitter noise percentages: 50% and 90%, and PMR channel density = 1.3536. The prediction tap set to zero defaults to the ordinary STD. The performance comparison is shown in Figs. 4.1 and 4.2.

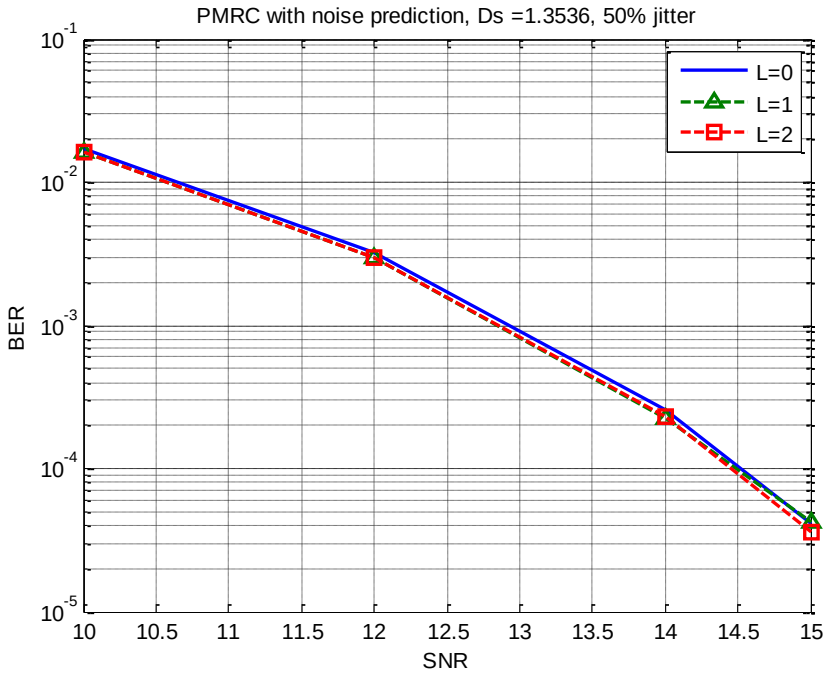

Fig 4.1. BER performance of the PDNP detector on a PMRC with 50% jitter noise.

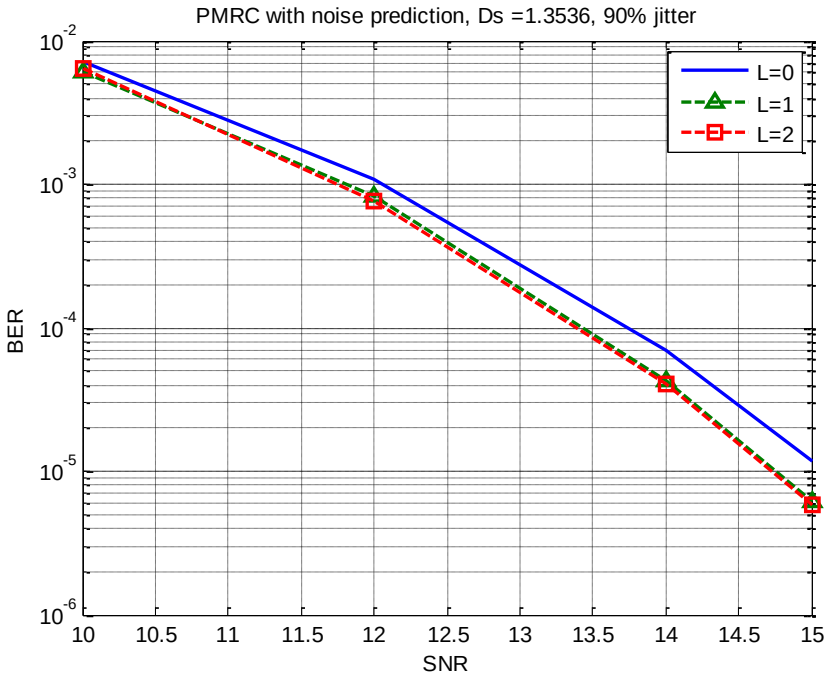

Fig 4.2. BER performance of the PDNP detector on a PMRC with 90% jitter noise.

The results show that the performance gain from the PDNP detector is marginal in 50% jitter noise. But when we increase the jitter noise percentage up to 90%, we achieved more performance gain from the noise predictive detectors of about 0.3dB at BER=10^{-5}, although we get little improvement from increasing the prediction taps.

Next we use the noise predictive detectors on the TDMR model. We took the similar strategy to perform a simulation test of the noise predictive BCJR on the Voronoi channel model with areal density = 1Tb/in^2 and BAR=2, as shown in Table 4.1.

Prediction taps (L)	Bit error rate (BER)
0	2.69×10^{-2}
1	2.67×10^{-2}
2	2.67×10^{-2}
3	2.62×10^{-2}

Table 4.1. Performance results for the PDNP detector on a TDMR channel model at 1Tb/in^2, BAR=2.

Unfortunately we do not get much gain from the noise predictive detector on the TDMR channel model. The reason is that at this density the interference comes mostly from ITI rather than from ISI. An intuitive idea is using a noise predictive BCJR coupled with ITI mitigating techniques that we discussed in Chapter 3. Since the noise prediction also requires trellis expansion, we choose to use ITI

cancellation instead of MTD to keep the low complexity of the overall channel detection. The noise prediction coefficients are based on samples of the signals coming out of the equalizer after ITI cancellation. The results are shown in Table 4.2. Compared with Table 4.1, the BER goes down significantly thanks to the ITI canceller, but the contribution from the noise predictor is not much at this density. The amount of gain obtained by the noise prediction depends on the particular channel density and architecture parameters, such as BAR, etc.

Prediction taps (L)	Bit error rate (BER)
0	1.55×10^{-2}
1	1.49×10^{-2}
2	1.51×10^{-2}
3	1.50×10^{-2}

Table 4.2. Performance results for the PDNP detector on a TDMR channel model at 1Tb/in^2, BAR=2, with ITI cancellation.

4.4. Conclusion

In this chapter we focused on addressing the problem of media noise by using PDNP detectors. The idea is to use past noise samples to predict the current noise sample for different input data patterns, in order to improve the performance of channel detectors in the presence of data dependent noise. Our results showed that for the PMR channel, the noise predictive BCJR provides performance gain in the

presence of large amounts of jitter noise. It can also be combined with ITI cancellation techniques and used on the TDMR channel model for additional gains. But because of the severe 2-D interference environment of TDMR, the PDNP detector may not be a strong contributor to improve the performance of higher-density TDMR channels.

Chapter 5 Conclusions and Future Work

This dissertation focuses on the investigation of signal processing techniques for 1-D and 2-D magnetic recording channels so as to address the challenges of the TDMR channel detection. In this chapter, we summarize the contributions of the dissertation and address some suggestions for future work.

5.1. Concluding Remarks

At the beginning of this dissertation, we took a look at several candidates for the next generation of HDDs, including BPMR, HAMR, MAMR and TDMR. We were interested in TDMR with SMR because it dramatically improves the disk capacity without changing the media. Most of the problems and challenges fall in the detection end, since the reader works on the overlapped tracks which are even narrower than the read head, the channel detector works in a environment of low SNR, two dimensional ISI and colored noise, therefore it requires sophisticated detection techniques to provide reliable data recovery.

To build our research environment, we used a new Voronoi grain based channel model which contains the important features of TDMR with SMR, such as squeezed tracks, tilted bit cells, 2-D ISI, electronic and media noise, etc. Then we conducted an in-depth investigation of channel detection techniques on the TDMR channel. Since the complexity of optimal SBS 2-D detection is exponential on the data width, we had to consider suboptimal solutions. Our approaches were the extension of the conventional 1-D detection techniques, by using joint-track

equalization to optimize the 2-D PR target followed by MTD for joint detection, or using the ITI canceller to estimate and cancel the ITI from side tracks, followed by a standard BCJR detector. We used STD for pre-detecting the side tracks to lower the overall complexity. Then we used PDNP to linearly predict the noise sample, so as to improve the detection performance with colored media noise, especially the data dependent jitter noise. The results showed that our 2-D detectors had significant performance gains over the conventional 1-D detectors, allowing the ECC decoders with of a threshold of BER=$10^{-1.6}$ to work on a 1Tb/in^2 TDMR channel with multiple reads/heads.

5.2. Suggestions for Future Work

To approach the optimal 2-D channel detection is an open problem. Besides the MTD and the ITI canceller that we have considered, there are several 2-D detectors which can be taken into consideration in future research. To our knowledge the iterative row-column soft-decision feedback algorithm (IRCSDFA) [32], Markov-chain Monte Carlo **(**MCMC) method [33] and generalized belief propagation (GBP) [34] are some useful techniques with limited complexity and near-optimal performance. We did not test each one of them on the TDMR model and compare to our MTD and ITI canceller, but that would be a natural recommendation for future work. A summary of these algorithms is included here for completeness.

A. **Iterative row-column soft-decision feedback algorithm (IRCSDFA)**

The IRCSDFA in [32] is an extension of the iterative multi-strip detection [35]. The basic idea of multi-strip detection is to split the 2-D readback signal data block into multiple strips, and use BCJR detectors working on the strips one after another. The IRCSDFA apply the multi-strip detector in both row-by-row and column-by-column directions on the 2-D data block, and the extrinsic information is exchanged between detectors in each direction for a certain number of iterations. We have mentioned in Chapter 1 that the direct implementation of BCJR on an $M \times N$ data block with ISI memory length of $I_M \times I_N$ should have a trellis of size $2^{M \times I_N}$, which is exponential on the data width. In the IRCSDFA the strip width is no less than the ISI width in each direction, therefore the complexity of a regular IRCSDFA is reduced to the order of $2^{I_M \times I_N}$. The complexity is still high for large size 2-D ISI, and a modified version proposed by Zheng *et al.* [36] uses Gaussian approximation (GA) instead of a multi-strip detector to make the strip width reduced to one, such that the overall complexity is dramatically reduced.

B. Markov-Chain Monte Carlo Detection

The MCMC based detection has been applied to the 1-D ISI channel in [37], and to the 2-D ISI channel in [33]. The 2-D MCMC detector uses Gibbs sampler to draw samples from target distribution and calculate the LLR for iterative detection. By using the samples instead of summations and integrations, the MCMC detector in [33] significantly reduced the complexity of 2-D detection from exponential to

polynomial on the ISI size, and showed better performance against the IRCSDFA of [32].

C. Generalized Belief Propagation (GBP)

The generalized belief propagation (GBP) algorithm [38] is a message passing technique based on the standard belief propagation (BP) algorithm. GBP is used as 2-D channel detector in [34], and has also been recently applied as a detector on a TDMR microcell model with 2-D ISI and data dependent noise [37].

The standard BP is commonly used in LDPC decoding, but it cannot be directly applied as an MAP detector since the short cycles within the ISI model will deteriorate the performance, unlike the sparse graph of an LDPC code. GBP avoids this problem by using regions of nodes instead of individual nodes for message passing so the regions and their intersections (also called subregions) form a tree-like graph which has no short cycles. The message passing between the regions can be executed by parent-to-child algorithm, child-to-parent algorithm and two-way algorithm, which are all explained in [38]. It is shown in [10] that the GBP reaches a near-optimal performance on 2-D ISI channel detection, so it is a strong competitor among the 2-D detectors for the next generation of HDDs.

Bibliography

Chapter 1

[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, pp. 917–923, Feb. 2009.

[2] R.Wood, "The feasibility of magnetic recording at 1 terabit per square inch," *IEEE Trans. Magn.*, vol. 36, pp. 36–42, Jan. 2000.

[3] K. S. Chan, R. Radhakrishnan, K. Eason, M. R. Elidrissi, Jim J. Miles, B. Vasic, and A. R. Krishnan, "Channel models and detectors for two-dimensional magnetic recording," *IEEE Trans. Magn.*, vol.46, no.3, pp.804-811, Mar. 2010

[4] H. Richter, A. Dobin, O. Heinonen, K. Gao, R. Veerdonk, R. Lynch, J. Xue, D. Weller, P. Asselin, M. Erden, and R. Brockie, "Recording on bit-patterned media at densities of 1 Tb/in and beyond," *IEEE Trans. Magn.*, vol. 42, no. 10, pp. 2255–2260, Oct. 2006.

[5] R. E. Rottmayer, S. Batra, D. Buechel, W. A. Challener, J. Hohlfeld, Y. Kubota, L. Li, B. Lu, C. Mihalcea, K. Mountfield, K. Pelhos, C. Peng, T. Rausch, M. A. Seigler, D.Weller, and X. Yang, "Heat assisted magnetic recording," *IEEE Trans. Magn.*, vol. 42, no. 10, pp. 2417–2421, Oct. 2007.

[6] J.-G. Zhu, X. Zhu, and Y. Tang, "Microwave assisted magnetic recording," *IEEE Trans. Magn.*, vol. 44, no. 1, pp. 125–131, Jan. 2008.

[7] H. Sawaguchi, Y.Nishida, H. Takano, and H. Aoi, "Performance analysis of modified PRML channels for perpendicular recording systems, " *J.Magnetism Man. Materials*, vol. 235, pp. 265-272, Oct. 2001

[8] K. S. Chan, J. Miles, E. Hwang, B. V. K. Vijayakumar, J. G. Zhu, W. C. Lin, and R. Negi, "TDMR platform simulations and experiments," *IEEE Trans. Magn.*, vol. 45, no. 10, pp.3837-3843, Oct. 2009.

[9] Y. Shiroishi, K. Fukuda, I. Tagawa, H. Iwasaki, S. Takenoiri, H. Tanaka, H. Mutoh, and N. Yoshikawa, "Future options for HDD storage," *IEEE Trans. Magn.*, vol. 45, no. 10, pp. 3816–3822, Oct. 2009.

[10] S. M. Khatami, and B.Vasic, "Generalized Belief Propagation Detector for TDMR Microcell Model," *IEEE Trans. Magn.*, vol. 49, no. 7, pp. 3699-3702, Jul. 2013.

[11] A. Kavcic, H. Xiujie, B. Vasic, W. Ryan, and M. F. Erden, "Channel modeling and capacity bounds for two-dimensional magnetic recording," *IEEE Trans. Magn.*, vol. 46, no. 3, pp. 812–818, Mar. 2010.

[12] R. Todd, E. Jiang, R. L. Galbraith, J. R. Cruz, and R. W. Wood, "Two-dimensional Voronoi-based model and detection for shingled magnetic recording", *IEEE Trans. Magn.*, vol. 48, no. 11, pp. 4594–4597, Nov. 2012.

[13] R. W. Wood and D. T. Wilton, "Readback responses in three dimensions for multilayered recording media configurations," *IEEE Trans. Magn.*, vol. 44, pp. 1874–1890, Jul. 2008.

[14] B. M. Kurkoski, "Towards efficient detection of two-dimensional intersymbol interference channels," *IEICE Trans. Fundamentals*, vol. E91-A, no. 10, Oct. 2008.

[15] E. Ordentlich and R.M. Roth, "On the computational complexity of 2D maximum-likelihood sequence detection," Tech. Rep. HPL-2006-69, HP Laboratories Palo Alto, July 2006.

Chapter 2

[16] W. Chang and J. R. Cruz, "Performance and decoding complexity of non-binary LDPC codes for magnetic recording," *IEEE Trans. Magn.*, vol. 44, no. 1, pp. 211–216, Jan. 2008.

[17] W. Chang and J. R. Cruz, "Optimal channel detection for non-binary coded partial response channels," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 1892–1895, Jul. 2009.

[18] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, no. 3, pp. 363–378, May 1972.

[19] H.-L. Lou, "Implementing the Viterbi algorithm," *IEEE Signal Processing Mag.*, pp. 42–52, Sept. 1995.

[20] J. Hagenauer and P. Hoeher, "A viterbi algorithm with soft-decision outputs and its applications," in *Proc. IEEE Globecom*, pp. 1683—1686, IEEE, 1989.

[21] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.

[22] J. Moon and W. Zeng, “Equalization for maximum likelihood detector,” *IEEE Trans. Magn.*, vol. 31, no. 2, pp. 1083–1088, Mar. 1995.

Chapter 3

[23] S. Nabavi and B. V. K. V. Kumar, “Two-dimensional generalized partial response equalizer for bit-patterned media,” in *Proc. IEEE Int.Conf. Commun.*, Glasgow, Scotland, 2007, pp. 6249–6254.

[24] W. Tan and J. R. Cruz, “Evaluation of detection algorithms for perpendicular recording channels with intertrack interference,” *J. Magn. Magn. Mater.*, vol. 287, pp. 397–404, 2005.

[25] W. Tan and J. R. Cruz, “Signal processing for perpendicular recording channels with intertrack interference,” *IEEE Trans. Magn.*, vol. 41, no.2, pp. 730–735, Feb. 2005.

[26] W. Chang and J. R. Cruz, “Inter-track interference mitigation for bit-patterned magnetic recording,” *IEEE Trans. Magn*., vol. 46, no. 11, pp. 3899-3908, Nov. 2010.

[27] W. Chang and J. R. Cruz, “Intertrack interference mitigation on staggered bit-patterned media,” *IEEE Trans. Magn*., vol.47, no.10, pp.2551-2554, Oct. 2011.

[28] K. Ozaki, Y. Okamoto, Y. Nakamura, H. Osawa, and H. Muraoka, “ITI canceller for reading shingle-recorded tracks,” in *Digests of the Ninth Perpendicular Magnetic Recording Conf.*, Sendai, Japan, 2010, pp.158- 159.

[29] Haratsch, E.F.; Mathew, G.; Jongseung Park; Ming Jin; Worrell, K.J.; Yuan Xing Lee; , "Intertrack interference cancellation for shingled magnetic recording," *Magnetics, IEEE Transactions on* , vol.47, no.10, pp.3698-3703, Oct. 2011

Chapter 4

[30] J. D. Coker, E. Eleftheriou, R. L. Galbraith, and W. Hirt, "Noise-predictive maximum likelihood (NPML) detection," *IEEE Trans. Magn.*, vol.34, pp. 110–117, Jan. 1998.

[31] J. Moon and J. Park, "Pattern-dependent noise prediction in signal-dependent noise," *IEEE J. Select. Areas Commun.*, vol. 19, no. 4, pp. 730–743, Apr. 2001.

Chapter 5

[32] T. Cheng, B. J. Belzer, and K. Sivakumar, "Row-column soft-decision feedback algorithm for two-dimensional intersymbol interference," *IEEE Signal Process. Lett.*, vol. 14, no. 7, pp. 433–436, Jul. 2007.

[33] M. Shaghaghi, K. Cai, Y. L. Guan, and Z. L. Qin, "Markov chain Monte Carlo based detection for two-dimensional intersymbol interference channels," *IEEE Trans. Magn.*, vol. 47, no. 2, pp. 471–478, Feb. 2011.

[34] O. Shental, N. Shental, S. Shamai (Shitz), I. Kanter, A. J. Weiss, and Y. Weiss, "Discrete-input two-dimensional Gaussian channels with memory: Estimation and information rates via graphical models and statistical mechanics," *IEEE Trans. Inform. Theory*, vol. 54, no. 2, pp. 1500–1513, Apr. 2008.

[35] M. Marrow and J. K. Wolf, "Iterative detection of 2-dimensional ISI channels," in *Proc. IEEE Information Theory Workshop*, Paris, France, Mar./Apr. 2003, pp. 131–134.

[36] Zheng, J.; Ma, X.; Guan, Yong Liang; Cai, K.; Chan, Kheong Sann, "Low-complexity iterative row-column soft decision feedback algorithm for 2D inter-symbol interference channel detection with Gaussian approximation," *IEEE Trans. Magn.*, vol.PP, no.99, pp.1

[37] R. Peng, R.-R. Chen, and B. Farhang-Boroujeny, "Markov chain Monte Carlo detectors for channels with intersymbol interference," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2206–2217, Apr. 2010.

[38] J. S. Yedidia, W. T. Freeman, and Y.Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2282–2312, Jul. 2005.

Appendix A - List of Acronyms

1-D	One-Dimensional
2-D	Two-Dimensional
AWGN	Additive White Gaussian Noise
APP	*a posteriori* Probability
BAR	Bit Aspect Ratio
BCJR	Bahl-Cocke-Jelinek-Raviv
BER	Bit-Error Rate
BP	Belief-Propagation
BPMR	Bit-Patterned Magnetic Recording
EAMR	Energy Assisted Magnetic Recording
ECC	Error Correcting Code
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FWHM	Full-Width-Half-Magnitude
GBP	Generalized Belief Propagation
GPR	Generalized Partial Response
HAMR	Heat-Assisted Magnetic Recording
HDD	Hard-Disk Drive
IRCSDFA	Iterative Row-Column Soft-Decision Feedback Algorithm
IFFT	Inverse Fast Fourier Transform
ISI	Inter-Symbol Interference

ITI	Inter-Track Interference
LDPC	Low Density Parity Check
LLR	Log Likelihood Ratio
LMR	Longitudinal Magnetic Recording
MAMR	Microwave-Assisted Magnetic Recording
MAP	Maximum *a posteriori*
MCMC	Markov-Chain Monte Carlo
ML	Maximum Likelihood
MLSD	Maximum Likelihood Sequence Detector
MMSE	Minimum Mean-Squared Error
MSE	Mean-Squared Error
MR	Magneto-Resistance
MTD	Multi-Track Detector
NPML	Noise-Predictive Maximum-Likelihood
NRZ	Non-Return-to-Zero
PDNP	Pattern-Dependent Noise Prediction
PMR	Perpendicular Magnetic Recording
PMRC	Perpendicular Magnetic Recording Channel
PR	Partial Response
PRML	Partial Response Maximum Likelihood
RLL	Run-Length Limited
RMS	Root-Mean-Square

RS	Reed-Solomon
SBS	Symbol-By-Symbol
SMR	Shingled Magnetic Recording
SNR	Signal-to-Noise Ratio
SOVA	Soft-Output Viterbi Algorithm
SSD	Solid-State Drive
STD	Single-Track Detector
SUL	Soft Under Layer
TDMR	Two-Dimensional Magnetic Recording