รายงานวิจัยฉบับสมบูรณ์

รหัสแก้ไขความผิดพลาดและการประมวลผลสัญญาณสำหรับ
การบันทึกข้อมูลความหนาแน่นสูง

โดย ศาสตราจารย์ ดร.พรชัย ทรัพย์นิธิ

สัญญาเลขที่ RSA568005

รายงานวิจัยฉบับสมบูรณ์

รหัสแก้ไขความผิดพลาดและการประมวลผลสัญญาณสำหรับ
การบันทึกข้อมูลความหนาแน่นสูง

ศาสตราจารย์ ดร.พรชัย ทรัพย์นิธิ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ

Project Code : RSA568005
(รหัสโครงการ)

Project Title : รหัสแก้ไขความผิดพลาดและการประมวลผลสัญญาณสำหรับการบันทึกข้อมูล
ความหนาแน่นสูง

Investigator : ศาสตราจารย์ ดร.พรชัย ทรัพย์นิธิ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

E-mail Address : pornchai.su@kmitl.ac.th

Project Period : 15 กันยายน 25556 ถึง 14 กันยายน 2559 (3 ปี)

เทคโนโลยีการบันทึกข้อมูลความหนาแน่นสูงสำหรับอนาคตมีทางเลือกสองทาง คือการ
บันทึกเชิงแม่เหล็ก อาทิเช่น การบันทึกแม่เหล็กแนวตั้งขั้นสูง, การบันทึกแม่เหล็กความร้อนช่วย,
การบันทึกสื่อแพทเทิรนบิตและ การบันทึกข้อมูลซ้อนทับ ที่ใช้งานในอุปกรณ์ฮาร์ดดิสก์ไดรฟ์
นอกจากนี้ยังมีการบันทึกแบบแฟลชที่ใช้งานในโซลิดสเตทไดรฟ์ การบันทึกเชิงแม่เหล็กได้รับ
ผลกระทบจากปรากฏการณ์หลายอย่าง ได้แก่ การเลื่อนทรานซิชันไม่เป็นเชิงเส้น (nonlinear
transition shift: NLTS) ความขรุขระอุณหภูมิ thermal asperity)   การแทรกสอดระหว่าง
สัญลักษณ์ (intersymbol   interference: ISI)   การแทรกสอดระหว่างแทร็ก (inter-track
interference: ITI) รวมทั้งปัญหาเฉพาะในเทคโนโลยีแต่ละประเภท เช่น ผลกระทบจากความร้อน
ต่อหัวอ่านในการบันทึกแม่เหล็กความร้อนช่วย การเขียนบิตผิดพลาดในการบันทึกสื่อแพทเทิรน
บิต เป็นต้น สำหรับการบันทึกแบบแฟลชนั้นได้รับผลกระทบจากสัญญาณรบกวนหลายชนิด เช่น
การเก็บรักษาบิต (retention noise) การแทรกสอดระหว่างเซลล์ (inter-cell interference) และการ
รบกวนการเขียน-อ่าน (read/write disturb) เป็นต้น สัญญาณรบกวนและความผิดเพี้ยนเหล่านี้ล้วน
มีผลต่อสมรรถนะของระบบการบันทึกข้อมูลเป็นอย่างมาก ในระบบบันทึกเชิงแม่เหล็ก ขนาดแทร็ก
และขนาดบิตที่เล็กลงมีผลทำให้การแทรกสอดระหว่างสัญลักษณ์และการแทรกสอดระหว่างแทร็ก
มีระดับที่รุนแรงขึ้น ในขณะที่ระบบการบันทึกแบบแฟลช การบันทึกหลายระดับต่อเซลล์ (multi
levels per cell) ทำให้การแทรกสอดระหว่างเซลล์มีค่ามากขึ้น ดังนั้นระดับเอสเอนอาร์ที่ต่ำลงจะมี
ผลต่อสมรรถนะของระบบโดยตรง ทำให้มีความจำเป็นที่จะต้องนำเทคนิคการประมวลผลข้อมูลเข้า

มาใช้งาน นอกจากนี้ระดับอัตราบิตผิดพลาดที่ต่ำมากในระดับน้อยกว่า 10^{-12} ทำให้ต้องนำรหัส
แก้ไขความผิดพลาด (error correction code: ECC) เข้ามาใช้งาน

ในโครงการวิจัยนี้ มีวัตถุประสงค์เพื่อพัฒนาวิธีการการประมวลผลสัญญาณแบบใหม่ และ
รหัสแก้ไขความผิดพลาดสำหรับการบันทึกข้อมูล โดยจะทำการออกแบบดีเท็กเตอร์ที่อิงกับกราฟ
และเทรลลิสสำหรับช่องสัญญาณแบบสองมิติ ข้อดีของการดีเท็กต์แบบกราฟคือใช้จำนวนรอบในการ
ทำงานที่ต่ำลง และมีความซับซ้อนแบบเชิงเส้นอีกด้วย สำหรับรหัสแก้ไขความผิดพลาดจะเน้นไปที่
รหัสพาริตี้เช็กความหนาแน่นต่ำหรือรหัสแอลดีพีซีแบบไบนารีและนอนไบนารี ที่ออกแบบด้วย
อัลกอริธึม progressive-edge growth (PEG) และมีการถอดรหัสด้วยการจัดเรียงลำดับการอัพเดท
ค่าที่แตกต่างจากวิธีดั้งเดิม โดยจะมีการจำลองสมรรถนะของระบบบันทึกเชิงแม่เหล็กและแบบ
แฟลช ท้ายที่สุดจะมีการวิจัยการออกแบบการบันทึกข้อมูลซ้อนทับ

คำหลัก : วงจรดีเท็กต์วิเทอร์บิ, แบบจำลองช่องสัญญาณสองมิติ, การบันทึกแม่เหล็กความร้อน
ช่วย, การบันทึกสื่อแพทเทิรนบิต, การบันทึกข้อมูลซ้อนทับ, หน่วยความจำโซลิดเสตท, การ
ประมวลผลวนกลับ, การอิควอไลซ์เทอร์โบ

Abstract

Project Code : RSA568005

Project Title : รหัสแก้ไขความผิดพลาดและการประมวลผลสัญญาณสำหรับการบันทึกข้อมูล
ความหนาแน่นสูง

Investigator : ศาสตราจารย์ ดร.พรชัย ทรัพย์นิธิ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

E-mail Address : pornchai.su@kmitl.ac.th

Project Period : 15 กันยายน 25556 ถึง 14 กันยายน 2559 (3 ปี)

The candidates for the next generation of disk drive (HDD) technologies include advanced perpendicular magnetic recording (PMR), heat-assisted magnetic recording (HAMR), bit patterned media (BPM) recording, shingled magnetic writing (SMR) as well as flash-based recording. The first four technologies are based on magnetic recording, while the other relies on the solid-state technology. The magnetic recording is faced with noise and distortion which include media noise, nonlinear transition shift (NLTS), thermal asperity, thermal response, intersymbol interference (ISI), inter-track interference (ITI) as well as system-specific issues such as heating gradients in HAMR, write errors in BPM as well as multiple shingle track pitch in SMR, all of which degrade the system reliability. For the flash-based memory, there are several types of noise sources including retention process, inter-cell interference (ICI), background pattern noise, and read/program disturb. Such noise sources reduce the storage reliability of magnetic and flash memory significantly. The continuous bit cost reduction of both types of devices mainly relies on aggressive technology scaling, which, however, further deteriorates the storage reliability. In magnetic recording, the bit size and track pitch are continually reduced causing the ISI, ITI issues, while for flash-based recording, multi levels per cell lead to increased ICI issues. In addition, the signal-to-noise ratio (SNR) is further decreased resulting in degraded bit error rate (BER). Signal processing is required to tackle the ISI, ITI and ICI issues. The typical storage reliability requirement is that non-recoverable BER must be

below 10^{-15}. Such stringent BER requirement makes error correction code (ECC) techniques mandatory to guarantee storage reliability. Different ECC techniques are required in different types of magnetic and flash memory. Turbo equalization between the detectors and ECC decoders can be used to further increase the system performance.

In this project, we develop the novel advanced signal processing and coding schemes for memory storage. Specifically, given a two-dimensional channel matrix, representing the ISI and ITI, we design the myriad detector types such as trellis-based and graph-based detectors. Benefits of such a structure include the SNR gains after fewer iterations, linear complexity and significant mitigation reduction of inter-track interference (ITI) and other distortions on the write side such as nonlinear transition shift (NLTS) and on the read side such as thermal asperity (TA), thermal response (TR) and dual peak-to-peak readback amplitude. For the ECC, binary and non-binary low-density parity-check (LDPC) codes based on progressive-edge growth (PEG) algorithm and the product code structure will be beneficial. For the decoders, the layered-scheduling and mixed-scheduling will be designed and proposed for magnetic-based and flash-based memory storage.

Keywords : Viterbi detector, two-dimensional channel model, heat-assisted magnetic recording (HAMR), bit patterned media (BPM), shingled magnetic writing (SMR), solid-state memory, iterative processing, turbo equalization

Executive Summary

Data Storage technology is an integral part of the digital economy, particularly, with the projection of ever increasing amount of data stored at each instants. From personal storage to cloud-based memory, both magnetic and flash recording technologies need to satisfy the insatiable storage demand of the public and the enterprise. Physical layer technology such as advanced signal processing and coding schemes are essential to the high-density magnetic recording beyond terabits per square and the multi-leveled cells (MLC) in solid-state memory. As Thailand is one of the largest manufacturers and a powerhouse in memory storage particularly in magnetic storage, high-quality research is important to raise the human capability as well as producing technology path from within the country. This research project was initiated to address these requirements in developing novel signal processing techniques appropriate for multi-tracking system and powerful error correction codes for both magnetic and flash memory. At the conclusion of this research project, all the research results were published in 3 high-quality journals and presented at 3 international conferences. Three doctoral and one master students graduated in this area of research. In addition, the researcher was invited to host a special session in data storage technology at the APSIPA 2014 conference and invited to be a technical committee member at the prestigious ICC 2014 conference in the data storage track.

Content

List of Tables

List of Figures

1. Introduction

1.1 Introduction to the research problem and its significance

The candidates for the next generation of disk drive (HDD) technologies include advanced perpendicular magnetic recording (PMR), heat-assisted magnetic recording (HAMR), bit patterned media (BPM) recording, shingled magnetic writing (SMR) as well as flash-based recording. The first four technologies are based on magnetic recording, while the other relies on the solid-state technology. The magnetic recording is faced with noise and distortion which include media noise, nonlinear transition shift (NLTS), thermal asperity, thermal response, inter-symbol interference (ISI), inter-track interference (ITI) as well as system-specific issues such as heating gradients in HAMR, write errors in BPM as well as multiple shingle track pitch in SMR, all of which degrade the system reliability. For the flash-based memory, there are several types of noise sources including retention process, inter-cell interference (ICI), background pattern noise, and read/program disturb. Such noise sources reduce the storage reliability of magnetic and flash memory significantly. The continuous bit cost reduction of both types of devices mainly relies on aggressive technology scaling, which, however, further deteriorates the storage reliability. In magnetic recording, the bit size and track pitch are continually reduced causing the ISI, ITI issues, while for flash-based recording, multi levels per cell lead to increased ICI issues. In addition, the signal-to-noise ratio (SNR) is further decreased resulting in degraded bit error rate (BER). Signal processing is required to tackle the ISI, ITI and ICI issues.

Current channel detectors such as Viterbi detector, often known as partial-response maximum likelihood (PRML) detector, are designed to mitigate the ISI rather than ITI. For high-density magnetic recording, particularly, for BPM and TDMR, the ITI needs to be considered. For 2-D interference, the implementation of a 2-D Viterbi detector is impractical due to excessive complexity with exponential complexity as a function of channel lengths. Efforts have been made on reduced-complexity approaches. One such approach is "modified Viterbi detector" which handles ITI by using multiple parallel branches entering and leaving a single state in the trellis. Alternatively, a graph-based structure, which has previously been used as an LDPC decoder, can be designed as a channel detector to mitigate ISI and ITI. The complexity is linear as a function of channel lengths. The structure itself allows SNR gain from running the detector with more iterations

even when the LDPC code is not employed.

The typical storage reliability requirement is that non-recoverable BER must not exceed 10^{-15}. Such stringent BER requirement makes error correction code (ECC) techniques mandatory to guarantee storage reliability. Different ECC techniques are required in different types of magnetic and flash memory. Turbo equalization between the detectors and ECC decoders can be used to further increase the system performance.

In this project, we intend to develop the novel advanced signal processing and coding schemes for memory storage. Specifically, given a two-dimensional channel matrix, representing the ISI and ITI, we design the myriad detector types such as trellis-based and graph-based detectors. Benefits of such a structure include the SNR gains after fewer iterations, linear complexity and significant mitigation reduction of inter-track interference (ITI) and other distortions on the write side such as nonlinear transition shift (NLTS) and on the read side such as thermal asperity (TA), thermal response (TR) and dual peak-to-peak readback amplitude. For the ECC, binary and non-binary low-density parity-check (LDPC) codes based on progressive-edge growth (PEG) algorithm and the product code structure will be beneficial. For the decoders, the layered-scheduling and mixed-scheduling will be designed and proposed for magnetic-based and flash-based memory storage.

1.2 Objectives

1. Design detectors for one-dimensional recording channels in advanced perpendicular magnetic recording (PMR) and heat-assisted magnetic recording (HAMR).

2. Design detectors for two-dimensional recording channels in bit patterned media (BPM), shingled recording (SMR).

3. Develop the LDPC decoder with mixed-schedulings.

4. Design binary and non-binary LDPC codes based on progressive-edge growth (PEG) algorithm and the product codes.

5. Investigate the flash-based recording channels and design soft-decisioned detector.

6. Perform the experiments and make a trade-off study in shingled magnetic recording (SMR).

2. Theory

2.1 Candidates for future magnetic recording system technology

At present, the areal density achieved for commercial purpose is estimated at around 800-900 Gbits/in^2. To design and plan for the next-generation high-density magnetic storage systems, the areal density at the level of over 1 Tb/in^2 (terabits per square inch) has been experimented and prototyped and it is envisioned that the level of 10 Tb/in^2 will be achieved. To increase the areal density, for magnetic recording, two parameters which need to be reduced are bit per inch (BPI) and track per inch (TPI), while for flash-based recording system; the number of bits per cell needs to be increased. Increasing these parameters lead to a decrease of signal-to-noise ratio (SNR), hence, the performance will degrade. For magnetic recording, at present, four candidate technologies are under consideration for the near future. They include advanced perpendicular recording, heat-assisted magnetic recording (HAMR) [1-2], bit patterned media (BPM) [3-5] and, recently proposed, shingled writing and two-dimensional magnetic recording (TDMR) [6].

HAMR achieves the balance by allowing high anisotropy media to be written by heating the media during the writing process (e.g., by laser light) to temporarily lower the anisotropy, hence, lower BPI is obtained [1]. HAMR employs a continuous domain writing approach. The image of HAMR is shown in Fig. 2.1.

On the other hand, BPM writes data on single domain magnetic island (a few grains); hence, discrete writing is used [2]. As shown in Fig. 2.2, in conventional recording, the information is stored at the transition between magnetic directions, resulting in a continuous region. The aerial density is increased by reducing the bit cell; hence, fewer magnetic particles are contained, causing lower response and the instability of the magnetic field. This effect is often known as "superparamagnetic effect". In PMS, however, each bit is recorded on an ordered array of highly uniform islands on a track of the magnetic film and each recorded bit is represented by a single-domain magnetic island and each island includes a single magnetic grain to achieve high density and stability [2].
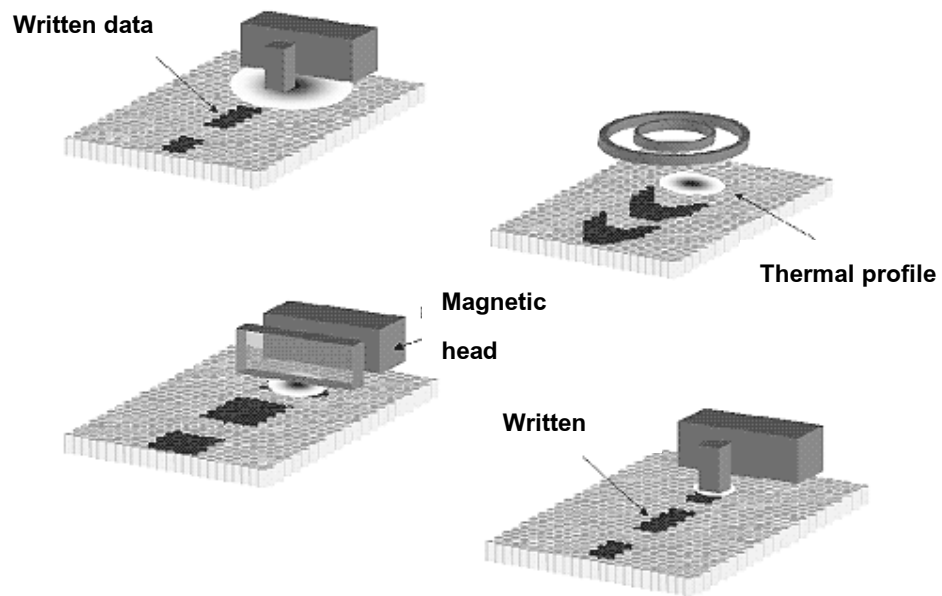
Fig. 2.1. Heat-assisted magnetic recording (HAMR).
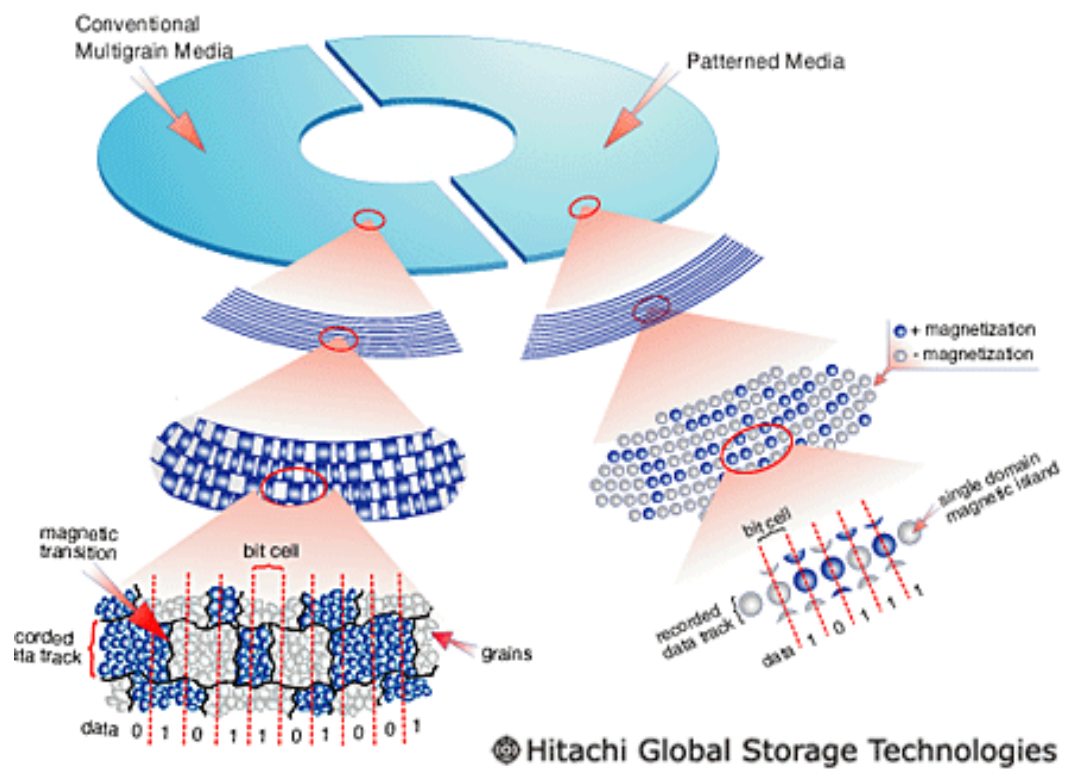
Fig. 2.2. Comparison between conventional continuous domain recording and patterned media recording system.

Moreover, the non-magnetic boundaries between two adjacent bit islands can reduce the transition noise problem, which is one of the serious noises in the conventional continues magnetic media.

In PMS, the distances between islands in along-track and cross-track directions can be reduced to achieve the ultra-high aerial density. And as the read head senses the magnetization, the resulting readback signal is corrupted by a two-dimensional (2D) interference that consists of inter-symbol interference (ISI) and inter-track interference (ITI) [3]-[5], which is one of the main limitations for PMS since it degrades the performance of the data recovery channel.

For TDMR, shingled writing is envisioned, whereby the write head writes from a corner only. In addition, written tracks partially overlap the previous ones [7] as shown in Fig. 2.3, hence, higher track density is achieved. For all of these proposed alternatives, as the read head senses the magnetization, the resulting readback signal is corrupted by a two-dimensional (2D) interference that consists of inter-symbol interference (ISI) and inter-track interference (ITI).
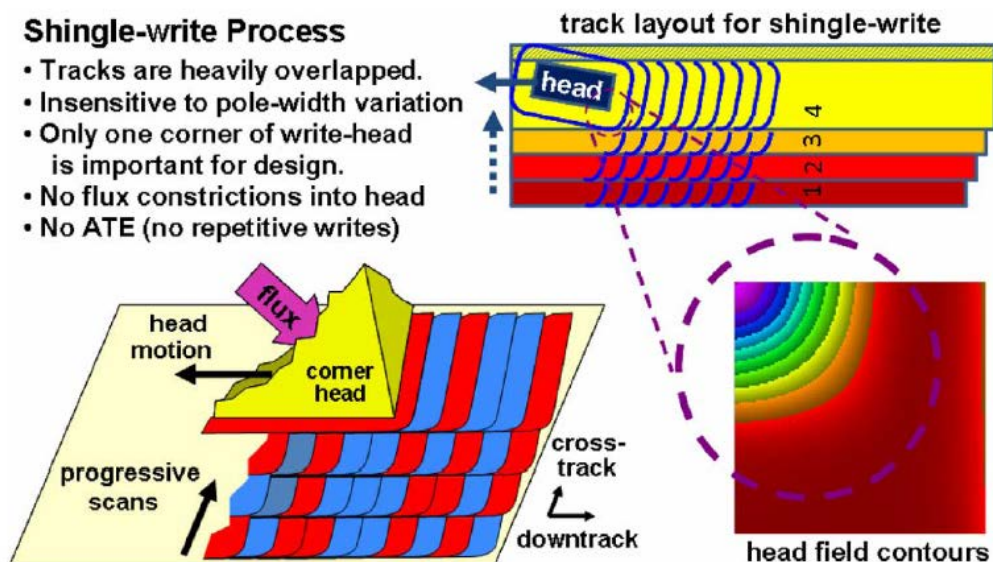
Fig. 2.3. Two-dimensional magnetic recording (TDMR) system [7].

In Fig. 2.4, overall encoding/decoding blocks of the TDMR system are shown. The signal processing and coding are drastically different HAMR and BPM system.
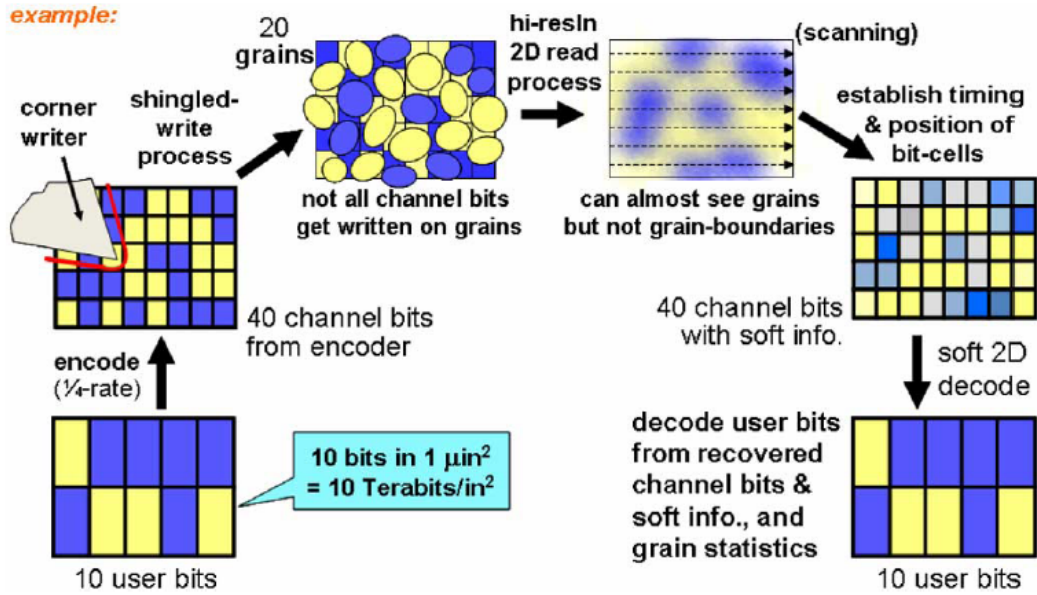
Fig. 2.4. Overall encoding/decoding of TDMR system [7].

The main difference lies in low code rates and ambiguity of the grain boundaries. On the read side, two-dimensional signal processing is required for a 2-D channel model.

2.2 Two-dimensional recording channel models

The 2-D channel model may be obtained from a single read head or multiple read heads. For a single head, a read head detects the signal mainly from the track under it together with some fractions from the above and below tracks [8]. The 2-D readback data can be obtained from a single head scanning a number of tracks, storing in a buffer and releasing the signal to the detector. For multiple read heads, the readback signal from multiple tracks is obtained. The 2-D channel model encompasses the effects from center track and adjacent tracks.
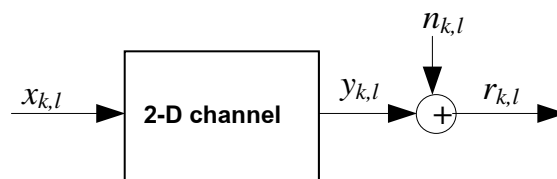
Fig. 2.5. The discrete-time channel model of typical PMS recording system.

6

A typical discrete-time channel model of 2-D recording system is illustrated in Fig. 2.5. In this model, the data bit at time k along the lth track, $x_{k,l}$, is recorded on the media. The obtained readback signal $r_{k,l}$ is expressed as

$$r_{k,l} = y_{k,l} + n_{k,l}, \tag{2.1}$$

$$= \sum_n \sum_m h_{n,m} x_{k-n,l-m} + n_{k,l} \tag{2.2}$$

where $h_{n,m}$'s are the 2-D channel response coefficients and $n_{k,l}$ is assumed to be additive white Gaussian noise (AWGN) [8]. Without the track misregistration (TMR), a discrete-time 3x3 symmetric channel response matrix **H** [8], [12], in the form of

$$\mathbf{H} = \begin{pmatrix} h_{k-1,l-1} & h_{k,l-1} & h_{k+1,l-1} \\ h_{k-1,l} & h_{k,l} & h_{k+1,l} \\ h_{k-1,l+1} & h_{k,l+1} & h_{k+1,l+1} \end{pmatrix} = \begin{pmatrix} \gamma & a & \gamma \\ \eta & b & \eta \\ \gamma & a & \gamma \end{pmatrix} \tag{2.3}$$

where b, a, η and γ are the channel coefficients, is a possible channel model among those in [3], [8].

Consider the 3x3 channel model in Fig. 2.6,
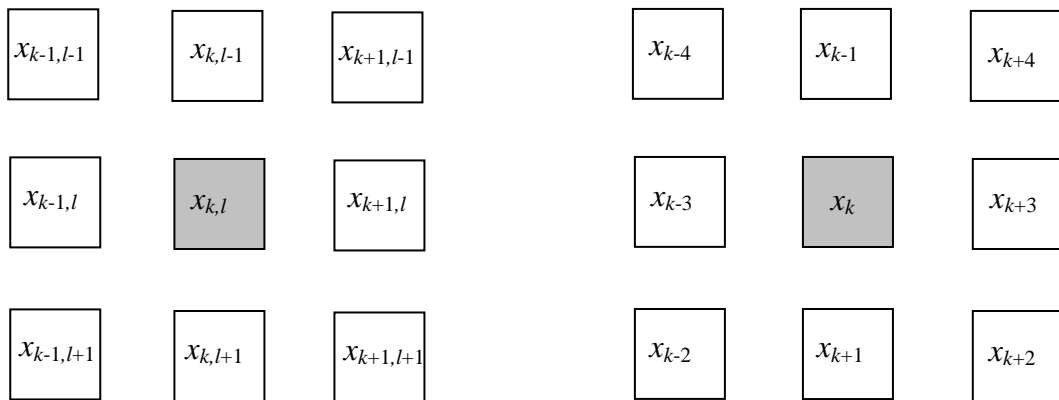
Fig. 2.6. A 3x3 bit configuration.

the readback signal in (2.2) can be rewritten as

$$r_k = bx_k + \eta(x_{k-3} + x_{k+3}) + a(x_{k-1} + x_{k+1}) + \gamma(x_{k-4} + x_{k-2} + x_{k+4} + x_{k+2}) + n_k. \quad (2.4)$$

The first four terms correspond to the noiseless channel output, $y_{k,l}$, in (1). The readback signal is corrupted by the ISI (second term), ITI (third and fourth terms) and noise (last term), hence, the challenge is to recover stored bit $x_{k,l}$ in the environment with such distortions. Based on (2.4), the PMS readback channel can be modeled using the delay blocks as shown in Fig. 2.7
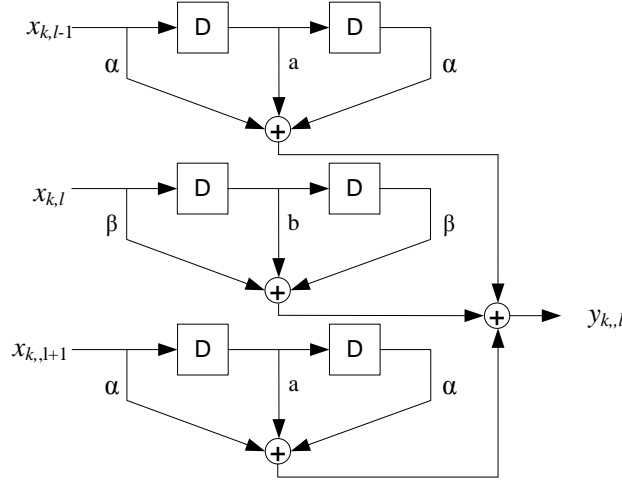
Fig. 2.7. The 2-D channel model with delay blocks.

2.2.1 Partial-response maximum likelihood (PRML) detector

Like a communications channel, the magnetic recording system uses the optimum detection method for the data recovery system. The optimum detection method is based on the maximum-likelihood criterion for detecting sequence of symbols; also known as *maximum-likelihood sequence detection* (MLSD). It minimizes the probability of error for a received bit sequence [13]. An efficient algorithm for implementing MLSD is the Viterbi algorithm (VA). Although the computational efficiency of the VA is up to the implementation of MLSD, its complexity increases exponentially depending on the span of the ISI. In such cases, the complexity of MLSD needs to be reduced in the data recovery system like reducing the ISI by a partial response (PR) equalization method. Fig. 2.8 shows a block diagram of the PRML detector, which consists of the PR equalizer and the optimum MLSD detector, on the receive side.
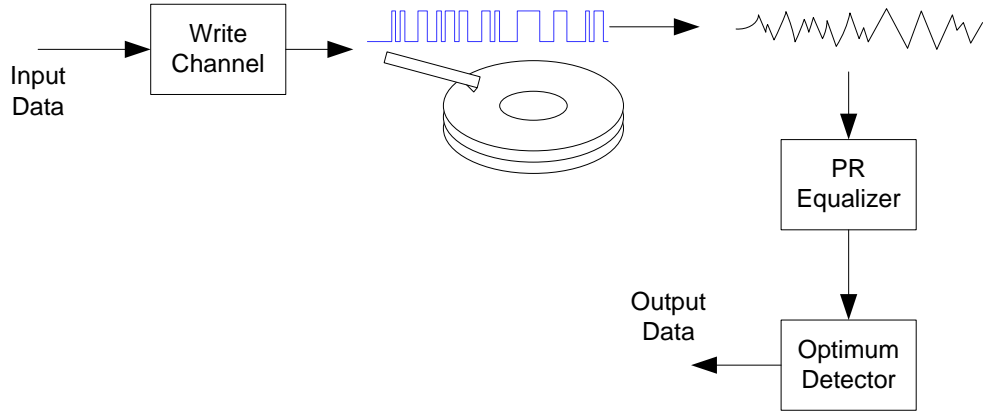
Fig. 2.8. Block diagram of PRML detector.

In this study, we design targets (HD) and equalizers (FD) by using the minimum-mean squared error (MMSE) method to minimize the mean-squared error (MSE) between desired outputs and equalizer outputs [17][19][21]. The target $H(D)$ and its corresponding $F(D)$ can be obtained by minimizing

$$
\begin{aligned}
E\left\{w_k^2\right\} &= E\left[\left(z_k - d_k\right)^2\right] \\
&= E\left[\left\{\left(s_k * f_k\right) - \left(a_k * h_k\right)\right\}^2\right]
\end{aligned}
\tag{2.5}
$$

where w_k is the difference between output of equalizer, z_k and the desired output, d_k of designed target, * is the convolution operator and $E\{.\}$ is the expectation operator, h_k and f_k stand for the coefficients of $H(D)$ and $F(D)$. The MMSE can be expressed as

$$
\varepsilon^2 = E\left\{w_k^2\right\} = \mathbf{F}^T \mathbf{R} \mathbf{F} + \mathbf{H}^T \mathbf{A} \mathbf{H} - 2\mathbf{F}^T \mathbf{P} \mathbf{H}
\tag{2.6}
$$

where the vector $\mathbf{H} = [h_0 \ h_1 \ h_2 \ \dots \ h_{L-1}]^T$ represents the L–tap generalized partial response (GPR) target and $\mathbf{F} = [f_{-K}, \ f_{-K+1}, \dots f_0 \ \dots \ f_K]^T$ represents the K-tap equalizer by where the length of the equalizer is N ($N = 2K+1$). $\mathbf{A}$ is an LxL autocorrelation matrix of a_k, $\mathbf{R}$ is an MxM autocorrelation matrix of sequence s_k, and $\mathbf{P}$ is an MxL cross-correlation matrix sequence of a_k and s_k . During the minimization process, the specified constraint must be used to avoid the trivial solution of $\mathbf{F} = \mathbf{0}$ and $\mathbf{H} = \mathbf{0}$.

Firstly, by minimizing (2.6) subject to a monic constraint, we fix h_0=1 and compute

$$\varepsilon^2 = \mathbf{F}^T \mathbf{R} \mathbf{F} + \mathbf{H}^T \mathbf{A} \mathbf{H} - 2\mathbf{F}^T \mathbf{P} \mathbf{H} - 2\lambda \left(\mathbf{I}^T \mathbf{H} - 1 \right)$$

$$\lambda = \frac{1}{\mathbf{I}^T \left(\mathbf{A} - \mathbf{P}^T \mathbf{R}^{-1} \mathbf{P} \right)^{-1} \mathbf{I}}$$

$$\mathbf{H} = \lambda \left(\mathbf{A} - \mathbf{P}^T \mathbf{R}^{-1} \mathbf{P} \right)^{-1} \mathbf{I}$$

$$\mathbf{F} = \mathbf{R}^{-1} \mathbf{P} \mathbf{H}$$

(2.7)

where λ is the Lagrange multiplier and $\mathbf{I}$ is an L-element column vector which 1st element is 1 and the rest is 0. Secondly, we fix the second target h_1=1 constraint. Column vector $\mathbf{J}$ that 2nd element is 1 and the others are 0. This is identical to monic constraint solution but $\mathbf{I}$ is replaced by $\mathbf{J}$. Thirdly, the energy $\mathbf{H}^T\mathbf{H}$=1 is fixed to minimize (28) called the unit energy constraint.

$$\varepsilon^2 = \mathbf{F}^T \mathbf{R} \mathbf{F} + \mathbf{H}^T \mathbf{A} \mathbf{H} - 2\mathbf{F}^T \mathbf{P} \mathbf{H} - 2\lambda \left(\mathbf{H}^T \mathbf{H} - 1 \right)$$

(2.8)

After differentiating and setting the result to 0, the final constraint
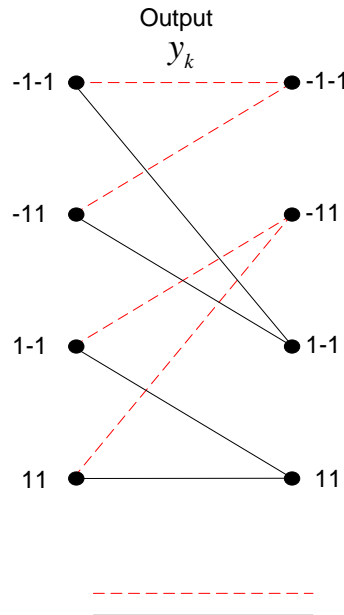
Fig. 2.9. The trellis structure of the Viterbi detector for the channel with two ISI.

The VA processes based on the trellis structure. If the PR equalizer reduces the length of the ISI from two neighboring signal (the number of delay block) to m, the number of states in the trellis structure is 2^m. For example, if there are the two ISI effects in the received signal, the number of states in the trellis will be 4 as shown in Fig. 2.9.

The detector uses Euclidean distance as the branch metrics to find the maximum likelihood sequence of symbol when the received signal is corrupted by additive white Gaussian noise. The branch metric l_k at time k is

$$l_k = (r_k - y_k)^2 \tag{2.9}$$

where r_k is the noisy received signal and y_k is the noiseless ideal channel output from the trellis. Then the path metric corresponding to each path in the trellis is calculated by adding all branch metrics along the path. The VA detector finds a path in the trellis with the minimum path metric as the maximum likelihood path or survivor path and generates the estimate input data by tracing along this path. Viterbi detector is optimal when the entering noise is AWGN, however, in magnetic recording channel, both media noise and colored noise exist; hence, a noise-predictive version or NPML is proposed [14]. The noise-whitening filter is inserted in front of the Viterbi detector; the trellis complexity is increased as a result.

2.2.2 Modified Viterbi detectors for inter-track interference (ITI) mitigation

Current channel detectors such as PRML and NPML detectors are designed to mitigate the ISI rather than ITI. However, for high-density magnetic recording beyond 1 Tb/in^2, the ITI needs to be considered as well. For 2-D interference, the implementation of a 2-D Viterbi detector is impractical due to excessive complexity, hence, efforts have been made on the reduced-complexity approaches. One such approach is a *modified Viterbi detector* [8] which handles the ITI by using multiple branches from a single state in the trellis. For the channel model in (2.3), assuming that possible values of $x_{k,l}$ are chosen from $\{\pm 1\}$, the noiseless channel output $y_{k,l}$ from the detected island in (2.4) can be expressed as

$$y_{k,l} = bx_{k,l} + h(x_{k-1,l} + x_{k+1,l}) + v_{k,l} + w_{k,l} \tag{2.10}$$

with the first ITI term $v_{k,l} \in \{\pm 2a, 0\}$ and the second ITI term $w_{k,l} \in \{\pm 4\gamma, \pm 2\gamma, 0\}$. The ISI portions, the first two terms in (8), result in a trellis with 4 states. Assuming that all corner ITI coefficients, γ, are zero, there are only 3 possible channel outputs. As a result, 3 parallel branches are obtained as shown in Fig. 10. If the values of $\pm 2\gamma$ are neglected, only 9 possible outputs are generated, resulting in 9 parallel branches, otherwise, 15 possible outputs are generated, resulting in 15 parallel branches.

Alternatively, the high complexity of 2-D detector can be avoided by reducing the ISI and mitigating the ITI by using a 2-D equalization method in [10], [12]. In [10], a 2-D generalized partial response (GPR) equalizer using the minimum mean squared error (MMSE) criterion was proposed to mitigate ITI effect in the readback signal before a 1D Viterbi detector. Moreover, in [9], Keskinoz proposed the iterative decision feedback detection (IDFD) method as an alternative method to the 2-D equalization with 1-D detector. For turbo equalization, the soft-output information is required for the decoder of the inner correction code. Two common algorithms are BCJR algorithm [15] and soft-output Viterbi algorithm (SOVA) [16]. In [17], we propose a partial ITI mitigation method (PIMM) based on multi-track processing with the use of an iterative detection by splitting the codeword into three tracks. The resulting trellis has only a single outgoing branch between connected states since all ITI effects are ignored. After the first iteration, hard estimates from the outer code are used compute partial ITI information.
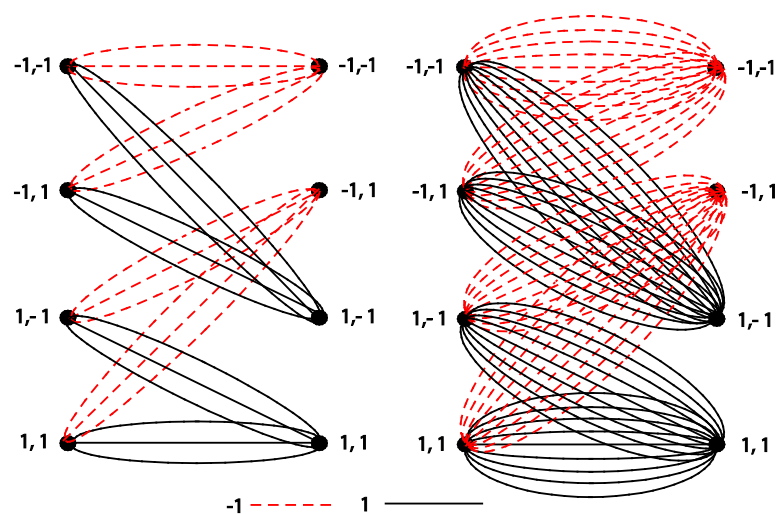
Fig. 2.10. Trellis diagrams for modified Viterbi detectors with 1 branch, 3 branches and 9 branches two connected states.

2.2.3 Graph-Based Detector

A graph-based detector can be used to handle 2-D interference channels. It is a 2-D detector which processes the readback sequences from multiple tracks, then produces the bit estimates of multiple tracks. A factor graph consists of two categories of nodes, the factor nodes which represent the readback signal $r_{j,k}$ and the bit nodes which represent the input bits $x_{j,k}$. During the detection process, an edge between the factor node and the bit nodes carries the reliability information in terms of the *a posteriori* log-likelihood ratio (LLR) based on the received signal. Figure 2.11 shows the factor graph for the 2-D channel in (2) partially by focusing on the edges from $r_{j,k}$ to the bit nodes (solid lines) and from $x_{j,k}$ to the factor nodes (dashed lines) only. Therefore, a full factor graph will have to include other factor nodes and bit nodes. The factor node $r_{j,k}$ connected to nine bits nodes passes the reliability information to $x_{j-1,k-1}$, $x_{j-1,k}$, $x_{j-1,k+1}$, $x_{j,k-1}$, $x_{j,k}$, $x_{j,k+1}$, $x_{j+1,k-1}$, $x_{j+1,k}$, $x_{j+1,k+1}$, then each of these bit nodes updates and sends back the message information to the connected factor nodes. Therefore, the three readback sequences are detected by the graph detector to produce the bit decisions of all three tracks $\{x_{j,k-1}, x_{j,k}, x_{j,k+1}\}$.
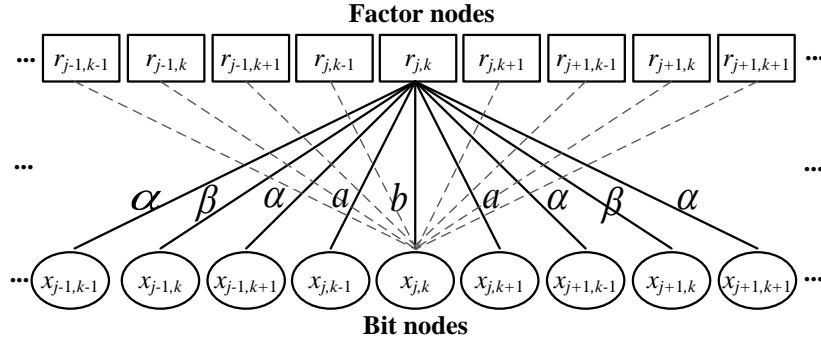
Fig. 2.11. The factor graph of the 2-D channel in (3).

A graph-based detection employs a belief propagation algorithm [47] to compute the LLR information at the factor nodes and bit nodes. For simplicity, $r_{j,k}$ and $x_{j,k}$ with the two indices (j,k) are first expressed as the variables r_n and x_p, each with a single index. The LLR information is calculated by the sum-product update rule based on the incoming messages from other connected bit nodes and the received signal. Using the max-log approximation $\log(e^a + e^b) \approx \max(a,b)$, the update message $L(\mu_{n \to p})$ from the factor node n to the bit node p is

13

$$L(\mu_{n \to p}) = \max_{B^n; x_p = +1} \left\{ -\frac{(r_n - y_n)^2}{2\sigma^2} + \sum_{x_l \in B^n_{\square x_p}, x_l = +1} L(\mu_{l \to n}) \right\}$$

$$- \max_{B^n; x_p = -1} \left\{ -\frac{(r_n - y_n)^2}{2\sigma^2} + \sum_{x_l \in B^n_{\square x_p}, x_l = -1} L(\mu_{l \to n}) \right\} \tag{2.11}$$

where y_n is the noiseless channel output, σ^2 is the variance of an AWGN sequence. After receiving the readback signal r_n, each factor node calculates the LLR messages for every connected bit nodes using (4). For the first term of the numerator, the squared *Euclidean distance* between r_n and y_n is calculated for all the possible values of y_n based on (2), for example, corresponding to all the combination of bit nodes, B^n, where x_p is set to +1. For each combination, the summation term is computed with the messages from all the bit nodes x_l ,except x_p, $B^n_{\square x_p}$, which has the value of +1 only in the combination set. Finally, the maximum value of the first term is obtained. Similarly, the denominator in (5) is calculated for the set of bit nodes when x_p is fixed to -1. The LLR information then obtained from the difference of the two maximum values.

After receiving the messages from the factor nodes, all the bit nodes, in turn, generate the LLR message $L(\mu_{p \to n})$ from the bit node p to the factor node n from

$$L(\mu_{p \to n}) = \sum_{l \in C^p_{\square n}} L(\mu_{l \to p}) \tag{2.12}$$

where $C^p_{\square n}$ represents the set of factor nodes connected to the bit node x_p excluding the factor node r_n. For the first iteration, the messages $L(\mu_{k \to n})$ are initialized to zero since all the input bits are assumed to be equi-probable. From the second iteration onward, the factor nodes update the messages using the incoming new messages $L(\mu_{l \to n})$. The updating and exchanging message process will continue until the final iteration is reached. After the specified number of iterations, the LLR information of the estimated output data $\hat{x}_n$ at bit node n can be computed from

$$L(\hat{x}_n) = \sum L(\mu_{l \to n}). \tag{2.13}$$

The hard estimates of the input bits can be obtained from

$$\hat{x}_n = \begin{cases} +1, \text{ if } L(\hat{x}_n) \geq 0 \\ -1, \text{ otherwise} \end{cases} . \tag{2.14}$$

14

Next, we propose two methods to improve the performance of the 2-D graph detector. The channel and target matrices are based on (2).

2.2.4. Mutual Information

The entropy value of a random variable X is measure of uncertainty of the random variable, the amount of information required on the average to describe a random variable. The mutual information between the equally likely X and the LLR for symmetric and consistent *L*-values by [27]. In graph detector, after the specified number of iterations, the LLR values of each bit can be obtained from (6). For the proposed methods (M1) and (M2), the two resultant LLR values for each bit must be averaged first. We can then measure the mutual information *I*(*L*;*X*) which is defined as

$$I(L; X) = 1 - \int_{-\infty}^{+\infty} p(L \mid x = +1) \log_2(1 + e^{-L}) dL \qquad (2.15)$$

or

$$I(L; X) = 1 - E\left\{\log_2(1 + e^{-L})\right\} \qquad (2.16)$$

where *E*{.} is the expectation operator.

2.3 Background on Low-density parity-check (LDPC) codes

2.3.1 Progressive-edge growth (PEG) algorithm

Low-parity density-check (LDPC) codes were invented in the early 1960's by Robert Gallager [22] but was forgotten. In the 1990's, LDPC codes were again introduced by MacKay and Neal [23]. An LDPC code is a linear block code which shows a good performance approaching the Shannon's limit [23-24] when decoded with the soft decision belief-propagation (BA) or the sum-product algorithm (SPA). LDPC codes can be constructed by random methods. One type of the random LDPC codes is a progressive edge-growth (PEG) LDPC code [25]. This method constructs a parity-check matrix or, equivalently, a Tanner graph, on an edge-by-edge basis giving a large girth and shows an impressive decoding performance through iterative decoding. Unfortunately, the random

LDPC codes do not render efficient encoding structure. Therefore, quasi-cyclic (QC) LDPC codes are preferable since they can be efficiently encoded using simple feedback-shift registers [26]. A disadvantage of the QC-LDPC codes is the limitation of flexible block sizes and code rates. Consequently, in [6], the QC-LDPC codes based on the PEG algorithm or PEG-QC-LDPC codes have been introduced to improve the girth property and the memory requirement. These methods construct the parity-check matrices based on the PEG algorithm [25] with established circulant sub-matrices instead of edges. These codes establish the flexible codes and offer comparable decoding performance to the conventional QC-LDPC codes. Recently, some methods to improve the random LDPC codes based on PEG algorithm have been presented [28-29]. An extended PEG algorithm [29] offers the high-rate random LDPC codes with large girth by considering the local girth on the variable nodes before generating the parity-check matrices, resulting in a high decoding performance. However, it does not provide the maximized girth property. Motivated by this fact, we propose the girth-maximized PEG algorithm and use it to generate the QC-LDPC codes with maximized girth property. This method generates each element of the parity-check matrix based on the PEG algorithm and compute the girth property for every node. Then we select the nodes with the maximum local girth to create the circulant sub-matrix and insert it to the parity-check matrix.

An LDPC code is a linear block code defined by an $m \times n$ sparse parity-check matrix **H** which can be represented by the Tanner graph [30]. A Tanner graph consists of the set of nodes S and edges E, where S combines the variable nodes and check nodes, $S = S_c \cup S_v$, where $S_c = \{c_0, c_1, c_2, ..., c_{m-1}\}$ is the set of check nodes and $V_v = \{v_0, v_1, v_2, ..., v_{n-1}\}$ is the set of variable nodes. E comprises the edges between the variable nodes and the check nodes, with $h_{ji} \neq 0$, where h_{ji} is the element in parity-check matrix at the jth row and the ith column. The set of edges connected to the variable nodes v_i is defined as $E = E_{v_0} \cup E_{v_1} \cup ... \cup E_{v_{n-1}}$. The degree of variable nodes v_i is denoted by d_v and the degree of check nodes c_i is denoted by d_c. For a given variable node v_i, a tree diagram [25] that described a tree spreading from variable node v_i in the Tanner graph can be drawn as shown in Fig. 2.12 where $N_{v_i}^l$ is the set of check nodes reached by a tree spreading from v_i at a *depth-l*. The complementary of $N_{v_i}^l$ is $\overline{N_{v_i}^l}$, where $\overline{N_{v_i}^l} = S_c \setminus N_{v_i}^l$. For each variable node v_i, the local girth g_{v_i} is defined by the length of the shortest cycle passing through that variable node, or equivalently $g_{v_i} = 2(l+2)$.
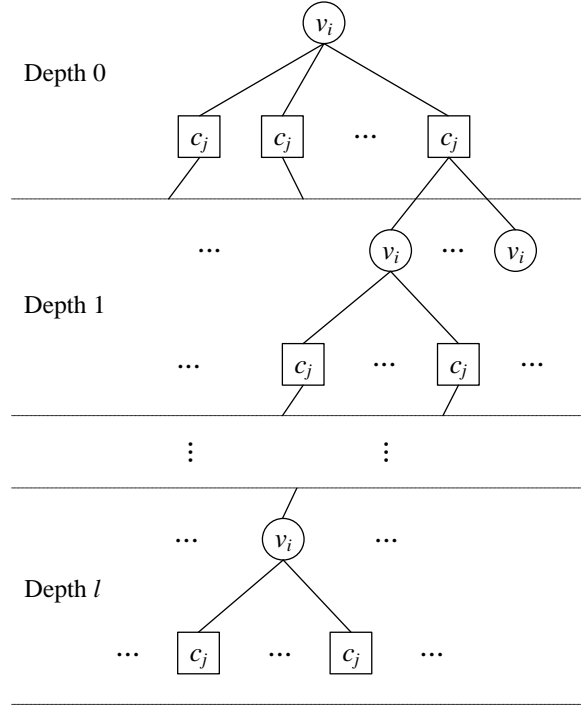
Fig. 2.12. A tree diagram [25] expanding from the variable node v_i with a depth-l.

A. Construction of QC-LDPC codes based on PEG algorithm

The QC-LDPC codes consists of a $c{\times}t$ array of circulant matrixes. A circulant matrix is a square matrix with dimension $p{\times}p$ in which each row is the cyclic shift of the previous row, and the first row is the cyclic shift of the last row. The parity-check matrixes for QC-LDPC follow the format

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{0,0} & \mathbf{H}_{0,1} & \cdots & \mathbf{H}_{0,t-1} \\ \mathbf{H}_{1,0} & \mathbf{H}_{1,1} & \cdots & \mathbf{H}_{1,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{c-1,0} & \mathbf{H}_{c-1,1} & \cdots & \mathbf{H}_{c-1,t-1} \end{bmatrix} \tag{2.17}$$

where $\mathbf{H}_{i,j}$ is circulant matrix with $0 < i < c$ and $0 < j < t$.

A. PEG-QC-LDPC codes

The conventional QC-LDPC codes have a limitation of flexible block sizes and code rates. Consequently, the QC-LDPC codes based on the progressive edge-growth (PEG) algorithm have been introduced [27]. The set of variable nodes S_v are divided into small

groups with each group having p variable nodes. An edge connected to the first variable node in a group is determined using the PEG algorithm. Then, the edges of other variable nodes in the same group can be automatically determined according to the circulant constraint. Figure 2.13 show an example of a PEG-QC-LDPC code with the circulant size of 4×4. The check nodes connected to the variable node v_4 lead to the circulant constraint of other variable nodes in group.
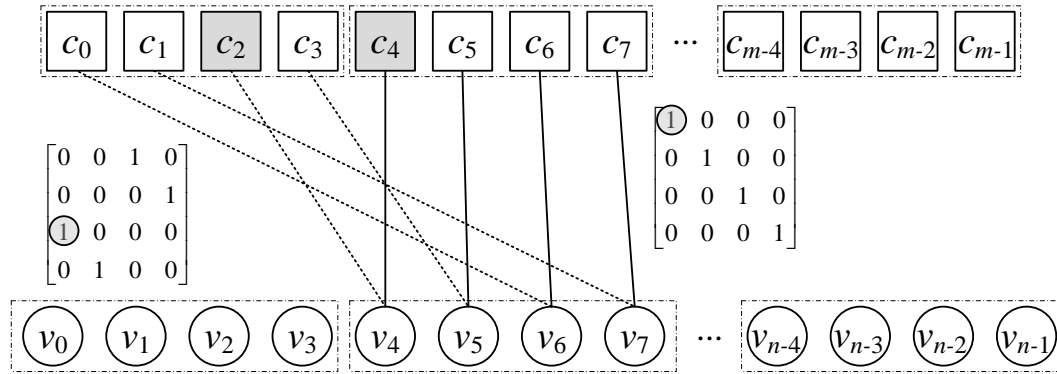
Fig. 2.13. An example of a PEG-QC-LDPC code with the circulant size of 4×4.

2.3.2 Algorithms for LDPC decoder

Low-density parity-check (LDPC) codes [22] is shown to exhibit the performance approaching the Shannon limit in an additive white Gaussian noise (AWGN) channel [23]. These codes have been widely employed in digital communication systems such as wireless LAN [30], and magnetic recording systems [31] among others.

An LDPC code, a kind of linear block codes, is based on a sparse parity-check matrix. To decode the LDPC codes, the belief-propagation (BP) algorithm is a well-known soft iterative decoding algorithm [23]. To improve the convergence speed of LDPC decoding, serial schedulings [32-37] and informed dynamic schedulings [34-37] have recently been proposed. For the serial schedulings, the check nodes are updated according to the numbered sequence of check nodes, while the updated check nodes can be performed in the increasing or decreasing orders in every iterative decoding proceeds. In [33], the authors further improve the performance of the serial schedulings by applying the opposite check node update directions for the even and odd iterations. For the informed dynamic schedulings, the updated check nodes are indicated by the residual (the difference of information from check node to variable node in the current iteration from the previous iteration). In [30], the serial scheduling

(also known as serial belief propagation) is shown to have the same decoding complexity per iteration as flooding scheduling. However, the informed dynamic scheduling which requires the residual generation and ordering of the residual has a very high complexity [37].

Recently, in [38], the serial belief propagation such as layered belief propagation (LBP) [28-30] and shuffled belief propagation (SBP) [31-33] have been proposed in an ultra-high density magnetic recording system called the bit patterned media (BPM) system. The results show that both the LBP and SBP give the fast convergence speed and they significantly improve the bit error rate performance under various media noise levels. However, the performance of LDPC decoding in the written-in error channel [39-41] that is an important challenge in the ultra-high density magnetic recording system has not been studied. In [42], it is also shown that LBP for nonbinary LDPC codes over GF(q) provides the better performance than the conventional LDPC decoding in conventional magnetic recording systems.

The belief-propagation algorithm is based on the exchange of information or reliability (also called the log-likelihood ratio, LLR) between the nodes of a Tanner graph [17] that represents the relation between codewords and parity-check matrices. A graph is composed of the variable nodes v_i that represent the ith code bit and check nodes c_j that represent the jth parity-check equation. In the graph, each variable node is connected to d_v check nodes and each check node is connected to d_c variable nodes. An example of a bipartite graph is shown in Fig. 2.14.
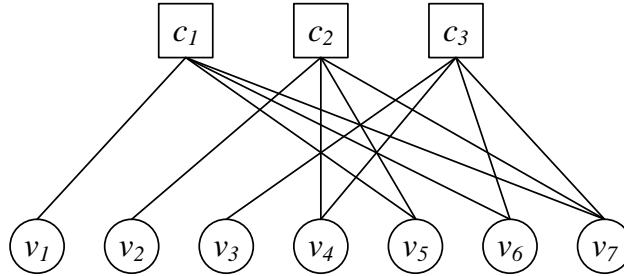
Fig. 2.14. A Tanner graph.

Belief-propagation algorithm relies on the propagation of the reliability from check nodes to variable nodes $L(r_{ji})$ as well as the reliability from variable nodes to check nodes $L(q_{ij})$. The reliabilities can be computed from

$$L(r_{ji}) = 2\tanh^{-1}\left(\prod_{i' \in N(c_j)\setminus i} \tanh\left(\frac{1}{2}L(q_{i'j})\right)\right), \qquad (2.18)$$

19

and

$$L(q_{ij}) = L(v_i) + \sum_{j' \in N(V_i) \setminus j} L(r_{j'i}) \; , \qquad\qquad (2.19)$$

where $i' \in N(c_j) \setminus i$ denotes the variable nodes connected to the check node c_j except the variable node v_i, and $j' \in N(v_i) \setminus j$ denotes the check nodes connected to the variable node v_i except check node c_j. The log-likelihood ratio (LLR) of the variable node v_i is $L(v_i) = \log(P(x_i=0|y_i)/P(x_i=0|y_i))$, where y_i is received signal corresponding to the transmitted codeword x_i. The reliability of a variable node is given by

$$L(Q_i) = L(v_i) + \sum_{j \in N(V_i)} L(r_{ji}) \qquad\qquad (2.20)$$

2.3.3 Layered belief propagation

Layered belief-propagation is illustrated in [7] and exhibits a faster convergence than the BP algorithm. LBP relies on different sequences of check node updates from the BP algorithm. In particular, for the BP algorithm, each check node propagates information $L(r_{ji})$ from the first until the last check nodes before the propagation of reliability $L(q_{ij})$ from variable nodes can occur. In the LBP, after a check node c_j passes the updates to connected variable nodes, each of these variable nodes will need to send the updates to its connected check nodes, except check node c_j , before the next check node can start the propagation so that it is updated using the most recent information. Sequence of check-node updates indicated in LBP is also known as horizontal shuffle scheduling [8-9]. Furthermore, the sequence of check-node propagation in LBP has the same decoding complexity per iteration as the traditional decoding [9-10]. Figure 2.15 shows a graphical explanation of the LBP strategy based on Fig.1. The LBP strategy will schedule to first propagate the information from the same check node c_j as shown in Fig. 2.15(a). Afterward, the information is propagated from the variable nodes connected to the check node c_j as shown Fig. 2.15(b). An iteration of LBP strategy will be counted after all the check nodes completely propagate the information $L(r_{ji})$. In the previous works, most authors consider the left-to-right direction of check node updates. In general, for LBP, the updated check nodes can start from the leftmost check node of the graph or start from the

rightmost one. In Section IV, we will investigate the difference in error locations from both methods.

(a)   (b)
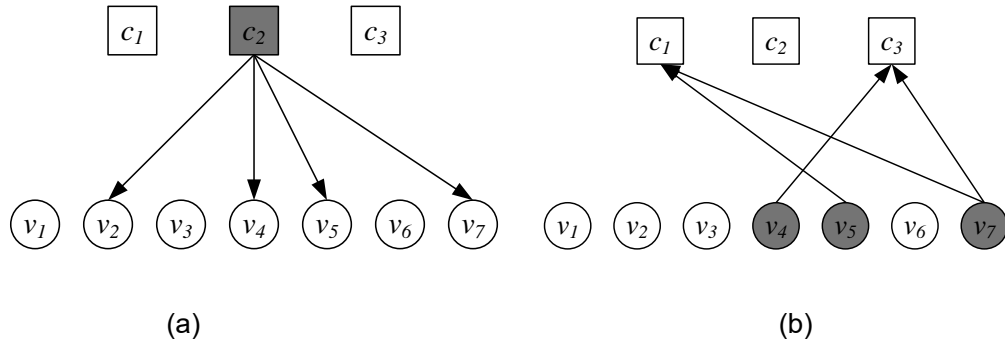
Fig. 2.15. Illustration of procedure for LBP strategy.

2.3.4 Shuffled belief propagation (SBP)

The shuffled belief propagation [11-12] is a type of serial scheduling which converges faster than the traditional BP decoding. The shuffled belief propagation (also called the vertical shuffle scheduling) is used to sequentially update the variable nodes instead of the check nodes in LBP. However, the sequential update of the variable nodes in SBP has a higher complexity than BP and LBP [10]. The graphical explanation of the SBP strategy is shown in Fig. 2.16. For LBP strategy, each check node will propagate all the information $L(r_{ji})$. On the other hand, SBP strategy does not schedule to propagate all the $L(r_{ji})$ from the same check node c_j but will schedule to propagate the $L(r_{ji})$ from the check node connected to the variable node as shown in Fig. 2.16(a). Afterward, the information is propagated from the variable node to all of its connected check nodes as shown Fig. 2.16(b). An iteration of SBP strategy will be counted after each variable node is updated completely.

(a)   (b)
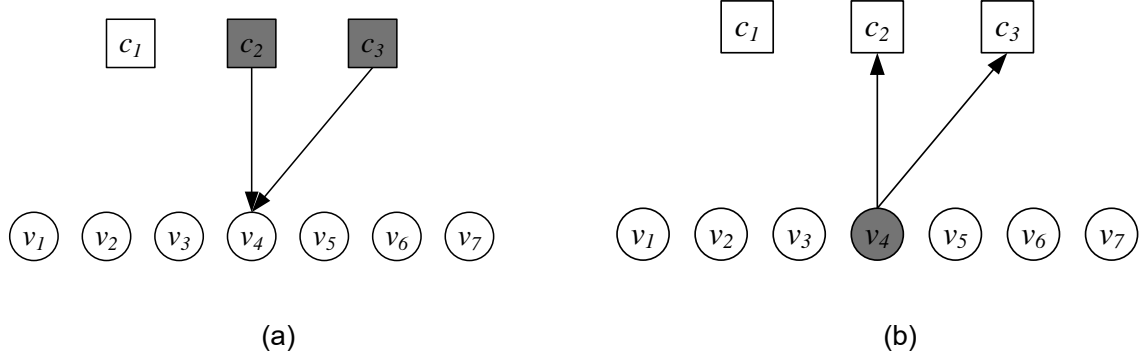
Fig. 2.16. Illustration of procedure for SBP strategy.

2.4 Heat-assisted magnetic recording (HAMR) systems

The Thermal Williams-Comstock (TWC) model can be used to determine the transition center x_0 and transition width a. First, defined he effective field (h) to account for heating behavior, i.e.,

$$h(x) = \frac{H_a(x)}{H_c(T(x))}.$$

(2.21)

The thermal Williams-Comstock slope equation can then be derived from

$$\left.\frac{dM(x)}{dx}\right|_{x_0} = \left.\frac{dM(H)}{dH_a(T(x_0))}\right|_{H_c(T(x_0))} \times \left[\left.\frac{dH_h(x)}{dx}\right|_{x_0} + \left.\frac{dH_d(x)}{dx}\right|_{x_0} - \left.\frac{dH_c(T(x_0))}{dT}\right|_{T(x_0)} \left.\frac{dT}{dx}\right|_{x_0}\right].$$

(2.22)

Since dM/dx and dH_d/dx depend on the parameter a, and it can be solved by using (2.22). To obtain dM/dx, it is assumed that the transition during recording can be described by the arctangent function from which dM/dx can be derived.

$$M(x) = \frac{2M_r(T(x))}{\pi} \tan^{-1} \frac{x - x_0}{a}$$

(2.23)

and dM/dx at x_0 can be solved by differentiating (2.23), i.e.,

$$\frac{dM(x)}{dx}\bigg|_{x_0} = \frac{2M_r(T(x_0))}{\pi a} \tag{2.24}$$

The term $dM(h)/dh$ from (2.22) at x_0 can be solved as slope of hysteresis loop is always positive at either transition is made from positive or negative H_c

$$\frac{dM(h)}{dh}\bigg|_{H_c(T(x_0))} = \left| \frac{1}{H_c(T(x_0))} \times \frac{M_r(T(x_0))}{1-S*} \right| \tag{2.25}$$

The term dH_h/dx of perpendicular recording at x_0 can be solved by using and the symmetry of the longitudinal Karlqvist head field component. The perpendicular head field can be derived by facilitating to turned sideways of longitudinal as

$$H_h = \frac{H_0}{\pi} \left[\tan^{-1}\left(\frac{y+g/2}{x}\right) - \tan^{-1}\left(\frac{y-g/2}{x}\right) \right], \tag{2.26}$$

And

$$\frac{H_h(x)}{dx}\bigg|_{x_0} = -\frac{gH_0}{\pi\left(x^2 + \left(g/2\right)^2\right)}, \tag{2.27}$$

where H_0 is the deep gap field, g is the gap width between pole head and its image ($g = 2d+2t$), t is the medium thickness and the field is evaluated at the center of medium ($y = t/2$). The parameter H_d for perpendicular recording is obtained by convolving the unit step response at the original and the magnetization gradient, i.e.,

$$H_{d,perp} = -\frac{\partial M(x)}{\partial x} * H_y^{step} = -\frac{\partial M(x)}{\partial x} * \frac{1}{\pi}\tan^{-1}\left(\frac{2x}{t}\right) \tag{2.28}$$

and

$$\frac{dH_d(x)}{dx} = -\frac{4}{\pi^2}\int_{-\infty}^{\infty}\frac{M_r(T(x'))a}{a^2 + (x'-x_0)^2}\frac{t}{t^2 + 4(x-x')^2}dx' +$$
$$\frac{4}{\pi^2}\int_{-\infty}^{\infty}\tan^{-1}\left(\frac{x'-x_0}{a}\right)\frac{dM_r(T(x'))}{dT}\bigg|_{T(x')}\frac{dT(x')}{dx'}\frac{t}{t^2 + 4(x-x')^2} \tag{2.29}$$

Here, H_c and M_r are linearly temperature-dependent (T),i.e,

$$H_c = -H_{c0}T(x, y) + H_{c,const}$$ (2.30)

and

$$M_r = -M_{r0}T(x, y) + M_{r,const}$$ (2.31)

where H_{c0} and M_{r0} are the temperature sensitivities of H_c and M_r, $H_{c,const}$ and $M_{r,const}$ are H_c and M_r at 0 Kelvin (K). So dH_c/dT is $-H_{c0}$ and dM_r/dT is $-M_{r0}$. A two-dimensional (2D) Gaussian thermal profile is given as

$$T(x, z) = T_{peak} \exp\left(-\frac{(x-c_0)^2}{2\sigma_t^2}\right)\exp\left(-\frac{z^2}{2\sigma_t^2}\right) + 300 \quad (K)$$ (2.32)

where T_{peak} is the peak temperature in the medium above room temperature in degree Celsius, c_0 is the laser spot position, x is defined to the position in down-track direction and z is defined to the position in cross-track direction. Differentiating (2.29), we have

$$\frac{T(x, z)}{dx} = T_{peak}\left(-\frac{(x-c_0)}{\sigma_t^2}\right)\exp\left(-\frac{(x-c_0)^2}{2\sigma_t^2}\right)\exp\left(-\frac{z^2}{2\sigma_t^2}\right).$$ (2.33)

2.4.1. Transition center (x₀) and transition parameter *a*

The position x_0 is defined as the point where the medium reverses its direction of magnetization and a is measured and related to the width of the magnetization transition. Achieving the narrowest possible transition (smallest a) allows placing recording bits close together and hence results in high linear density. The position x_0 of perpendicular HAMR with large spot and linear relationship for H_c and M_r can be calculated by

$$x_0 = \frac{g + \sqrt{g^2 - a\tan^2\left(\pi\frac{H_c\left(T\left(x_0\right)\right)}{H_0}\right)\left(y^2 - \left(g/2\right)^2\right)}}{2\tan\left(\pi\left(\frac{H_c\left(T\left(x_0\right)\right)}{H_0}\right)\right)} \tag{2.34}$$

and *a* can be also calculated from

$$a = -\frac{\gamma}{2} + \frac{1}{2}\sqrt{\gamma^2 + \frac{4H_c\left(1 - S*\right)t}{\Delta\pi}}\Bigg|_{x_0} \tag{2.35}$$

where

$$\Delta \equiv \frac{H_g g}{\pi\left(x_0^2 + \left(g/2\right)^2\right)} - \frac{dH_c}{dT}\frac{dT}{dx}\Bigg|_{x_0} \tag{2.36}$$

and

$$\gamma \equiv \frac{2M_r}{\Delta\pi} - \frac{t}{2} + \frac{2H_c(1 - S*)}{\Delta\pi} \tag{2.37}$$

For non-large spot HAMR, x_0 can be solved by

$$x0, j+1 = \frac{g + \sqrt{g^2 - 4\tan^2\left(\pi\frac{H_c(T(x_0)) + H_d(x_0, j, a_k)}{H_0}\right)\left(y^2 - \left(g/2\right)^2\right)}}{2\tan^2\left(\pi\frac{H_c(T(x_0)) + H_d(x_0, j, a_k)}{H_0}\right)} \tag{2.38}$$

and *a* can be solved by

$$a_{k+1} = \frac{-\beta + \sqrt{\beta^2 + 4\pi\left(\alpha + \delta\right)H_c\left(1 - S*\right)t}}{2\pi\left(\alpha + \delta\right)} \tag{2.39}$$

where

$$\alpha = \frac{H_g g}{\pi\left(x_{0,j+1}^2 + (g/2)^2\right)}, \delta = \frac{dH_d\left(a_k, x_{0,j+1}\right)}{dx} - \frac{dHc}{dx},$$ (2.40)

and

$$\beta = -\frac{\pi\alpha t}{2} + 2M_r - \frac{\pi t}{2}\left(\frac{dH_d\left(a_k, x_{0,j+1}\right)}{dx} - \frac{dH_c}{dx}\right) + 2H_c\left(1 - S*\right)$$ (2.41)

In this work, the medium is assumed to be $Fe_{25}Ni_{30}Pt_{45}$. For all cases, the laser is assumed to be at the center of the track in the cross-track direction but varied in the down-track direction. The coercivity H_c and remanent magnetization M_r have linear relationship with the temperature. The temperature induced by the laser is assumed to be Gaussian in both dimensions with the peak temperature of 330 °C and track width 20 nm.

2.4.2. The Generation of the HAMR readback signal

The microtrack model is used for 2D process of heating and magnetization of a medium to solve the problem of determining the transition characteristics in HAMR by dividing recorded track into N individual sub-tracks of equal width with different transition parameters. The TWC equation is applied to determined x_0 and a in each individual sub-track. The playback voltage (V) of perpendicular recording is given by

$$V(x, a_i, x_{0,i}) = -CM_r t \ln\left(\frac{\left(g_r/2 - (x-x_0)^2\right) + (d+a+t/2)^2}{\left(g_r/2 + (x-x_0)^2\right) + (d+a+t/2)^2}\right)$$ (2.42)

where C is a system specific constant, g_r is the read head gap, and d is the head-medium spacing and t is denote to media thickness. The overall transition response $p(x)$ can be obtained by weighting the transition response $p_i(x)$ of each sub-track to obtain

$$p_i(x) = \exp^{-\left(-\frac{N+1}{2}\Delta z + i\Delta z\right)^2 \Big/ 2\sigma_r^2} \cdot V(x, a_i, x_{0i})$$ (2.43)

and

$$p(x) = \frac{1}{N} \sum_{i=1}^{N} p_i(x) \qquad (2.44)$$

where a_i and x_{0i} are the transition parameter and transition center of the ith microtrack, i can be 1 to N and Δz stands for the width of each sub-track. The bit response is then

$$h(x) = 0.5\{p(x) - p(x - T_x)\} \qquad (2.45)$$

where T_x denotes the along-track bit period.

2.5 Flash recording channels

NAND flash memory array is divided into blocks and each block consists of a number of pages as shown in Fig. 2.17. Each block has a number of bit lines (memory cells) and word lines (pages). Each bit line has 16 to 64 memory cells.
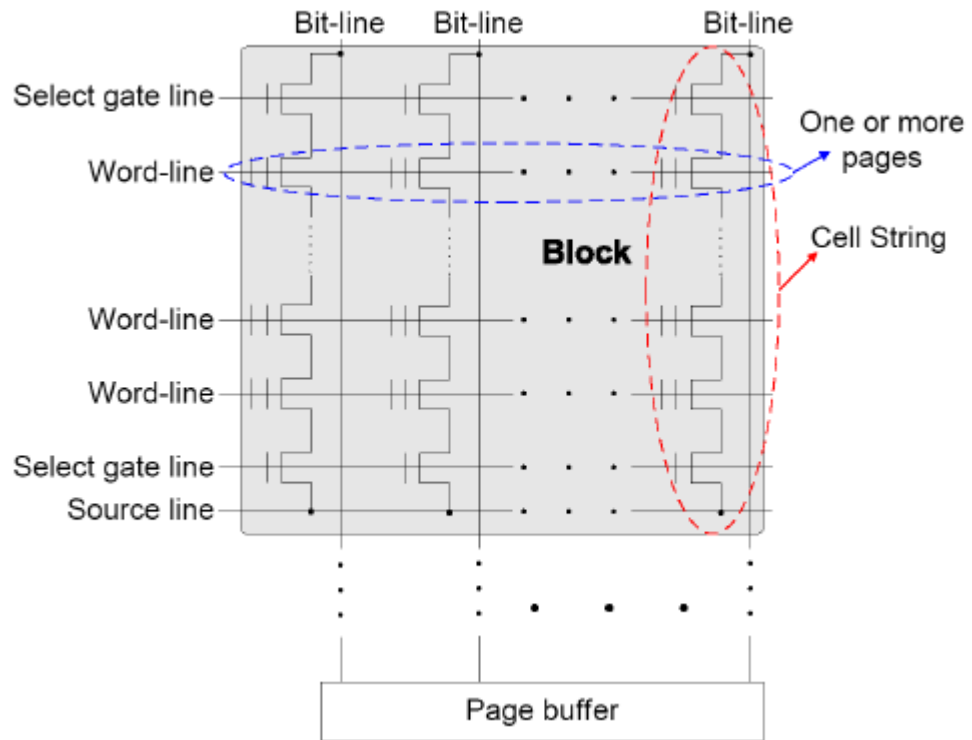
Fig. 2.17. Structure of flash memory blocks.

27

All the memory cells within the same block must be erased at the same time and data are programmed and input in the unit of page. The page size varies from 512 bytes to 8K bytes.

A NAND flash memory channel model is shown in Fig. 2.18.
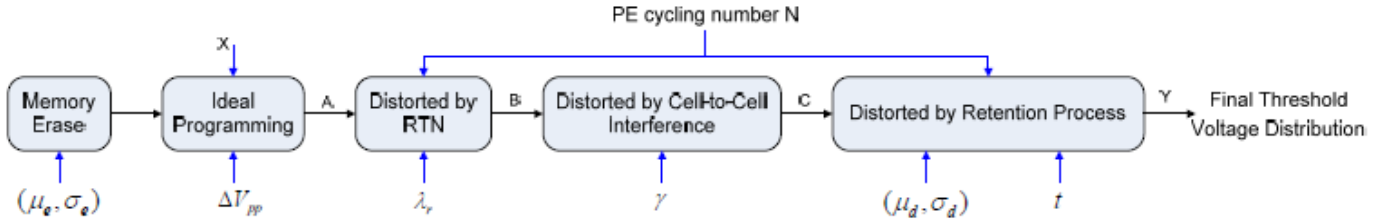
Fig 2.18. Flash memory channel.

A flash memory cell needs to be erased before it can be programmed. In order to erase a cell, all the charges are removed from the floating gate and the threshold voltage is set to the lowest possible values. The distribution of the threshold voltage of erased state can be modeled by a Gaussian distribution $N\left(\mu_e, \sigma_e^2\right)$, i.e.,

$$p_e(x) = \frac{1}{\sqrt{2\pi\sigma_e^2}} e^{-\frac{(x-\mu_e)^2}{2\sigma_e^2}}$$
(2.46)

where μ_e and σ_e are the mean and standard deviation of the erased state.

Each programmed state, with program and verify approach, has the threshold voltage of uniform distribution over $\left[V_p, V_p + \Delta V_{pp}\right]$. If we denote V_p and $V_p + \Delta V_{pp}$ of the k-th programmed state as $V_l^{(k)}$ and $V_r^{(k)}$, then the ideal threshold voltage distribution of the k-th programmed state can be modeled by

$$p_p^{(k)}(x) = \begin{cases} \dfrac{1}{\Delta V_{pp}}, & V_l^{(k)} \leq x \leq V_r^{(k)} \\ 0, & \text{otherwise} \end{cases}$$
(2.47)

In practice, the ideal threshold voltage distribution is affected by the program/erase (PE) cycling which causes (1) the threshold voltage shift and fluctuation known as random telegraph noise (RTN) and (2) in the memory cell causing data retention issue. The probability density function of the RTN-induced threshold voltage is

$$p_r(x) = \frac{1}{2\lambda_r} e^{-\frac{|x|}{\lambda_r}} \tag{2.48}$$

If N is the cycling number, λ_r scales with a power-law fashion. Additionally, the threshold voltage reduction is modeled as a Gaussian distribution $N(\mu_d, \sigma_d^2)$ where μ_d and σ_d^2 scale with N in a power-law fashion and scales with the retention time t in a logarithmic fashion. The cell-to-cell interference in NAND flash memory is caused by parasitic capacitance-coupling effect. The resulting threshold voltage shift of a victim cell can be approximated as

$$F = \sum_k \left(\Delta V_t^{(k)} \cdot \gamma^{(k)} \right) \tag{2.49}$$

where $\Delta V_t^{(k)}$ is the shift in threshold voltage of an interfering cell which is programmed after the victim cell. The coupling ratio is defined as

$$\gamma^{(k)} = \frac{C^{(k)}}{C_{tot}} \tag{2.50}$$

where $C^{(k)}$ is the parasitic capacitance between the interfering cell and the victim cell, and C_{tot} is the total capacitance of the victim cell.

From Fig. 2.18, the threshold voltage distribution after the RTN model is

$$p_{ar}(x) = p_p(x) \otimes p_r(x) \tag{2.51}$$

where $p_p(x)$ is the threshold voltage distribution after the ideal programming, and $p_r(x)$ is from (2.45). For cell-to-cell interference, both the vertical and diagonal coupling ratios are set to be truncated Gaussian distribution, i.e.,

29

$$p_c(x) = \begin{cases} \dfrac{k_c}{\sigma_c\sqrt{2\pi}} e^{-\frac{(x-\mu_c)^2}{2\sigma_c^2}} & ,|x-\mu_c| \leq w_c \\ 0 & ,\text{else} \end{cases} \qquad (2.49)$$

Given that the threshold voltage after the cell-to-cell interference be $p_{ac}(x)$ and the distribution of retention noise distribution be $p_t(x)$, the final threshold voltage distribution $p_f(x)$ is obtained as

$$p_f(x) = p_{ac}(x) \otimes p_t(x) \qquad (2.50)$$

Finally, the retention noise can be modelled as Gaussian distribution where the mean and std depends on the Erase and Programming cycles N, i.e.,

$$\mu_d = K_s(x-x_0)K_d N^{0.5} \ln\left(1+\frac{t}{t_0}\right) \qquad (2.51)$$

and

$$\sigma_d^2 = K_s(x-x_0)K_m N^{0.6} \ln\left(1+\frac{t}{t_0}\right) \qquad (2.52)$$

In [49], proposes the non-uniform fine-grained soft-decision memory sensing and investigates the performance of binary LDPC codes in NAND flash memory channels with 2 bits/cell. In [50], product codes based on the combination of LDPC codes as well as Hamming codes are investigated. For realistic code design, it is essential to understand the error patterns of flash memory, therefore, in [51], the authors spent much efforts characterizing the error patterns in NAND flash memory.

3. Methodology

Channel simulations will be made for various types of storage. For advanced PMR and HAMR, the full system will be simulated, then the designed target equalizers as well as detectors will be made. Since the advanced PMR is affected by nonlinearity on the write side and read side, the system performance will be studied. In addition, the nonlinear compensation methods will be proposed. The write nonlinearity will be based on Volterra's model and the use of pseudo-random sequence is exploited to obtain the dipulse and nonlinearity information, the compensation on the nonlinearity echoes will be emphasized. For the read side, nonlinearity such as sudden amplitude changes will be studied in terms of the probability density function (pdf), then the detector will be designed to detect such events as well as improve the system performance. For HAMR, it is well-known that the intersymbol interference level is higher than that in the advanced PMR channels, and in addition, the shape of dibit pulse differs significantly. Therefore, the study on suitable target-shaping equalizers will be made.

For bit-patterned media and shingled writing, we will focus on the 2-D channel models in order to design the 2-D detectors as well as the 2-D error correction codes for such channels. The designed detectors will be based on trellis structure with the focus on reduced-complexity alternatives. The designed codes are binary as well as nonbinary low-density parity-check (LDPC) codes based on the progressive-edge growth (PEG) algorithm. The decoding schedulings will be the mixed-type scheduling which combines the output from the current layered schedulings with right-to-left and left-to-right direction. These designed codes and decoders will be also be designed specifically for flash memory channels as well to test the improvement in it lifetime.

All the simulations will be on made in MATLAB software. The channels and noise would represent the main characteristics of the recording channels. Finally, some experiments on the applicability and reliability of Shingled magnetic writing (SMR) will be made. Some parameters under study are SNR, reverse overwrite (ROW) ratio, areal density, partial erasure (PE), and others.

For the scope of Research, the channels considered are in the continuous-time domain for advanced PMR and HAMR. For detector design, the bit error rate test is limited for BER level of 10^{-5} only. To simulate the dibit pulse of HAMR, up to 10 microtracks will be used. The types of noise and distortion considered are jitter noise of up to 50 percent, nonlinear models up to two to three orders. The areal density considered is not more than

5 terabits/square inch. For bit-patterned media and shingled writing, we will focus on the 3x3 2-D channel matrix to account for ITI from two adjacent tracks and ISI from two adjacent bits.   For code design, only high-rate LDPC codes of more than 0.8 are considered. When implementing in the system, the BER levels considered is not lower than 10^{-8}. For the experiments in Shingled magnetic writing, we use the write and read heads in the actual manufacturing process; therefore, the sizes will vary significantly.

4. Results and Discussions

4.1. Heat-assisted magnetic recording systems

4.1.1 Simulation of HAMR readback signals

Firstly, we fix M_r magnetization dependencies on temperature to study the behavior of x_0 and a by using various H_c. With magnetization dependencies on temperature (M_r=-1000T+1.8x10^6), we found that at increasing H_c, x_0 is shifted faraway from laser position and a is decreased. Secondly, we vary M_r to study x_0 and a by using H_c = -2900T+2.4x10^6 for evaluation. Fig. 4.1 shows that x_0 has almost the same values and small a is found at low M_r. Small a means narrow transitions, hence, recording bits are packed close together. System parameters are listed in Table 4.1.

TABLE 4.1. SYSTEM PARAMETERS

| Symbol	Parameter	Unit
TW	track width	20 nm
N	number of sub-track	10
c_o	laser position	0 nm
H_0	deep gap field	1x10^6 A/m
T_{peak}	peak temperature	330 C°
σ_t	sigma of temperature profile	16.2 nm
G	gap width between pole head and its image	32 nm
σ_r	sigma of reader sensitivity function	4.23 nm
d	head-medium distance or fly height	6 nm
t	medium thickness	10 nm
T_x	bit period	6 nm
C	a system specific constant	1
H_c	dependence of coercivity on temperature	2900T+2.4x10^6 A/m
M_r	dependence of remanent magnetization on Temperature	600T+1.8x10^6 A/m
AD	Areal density	5 Tb/in^2
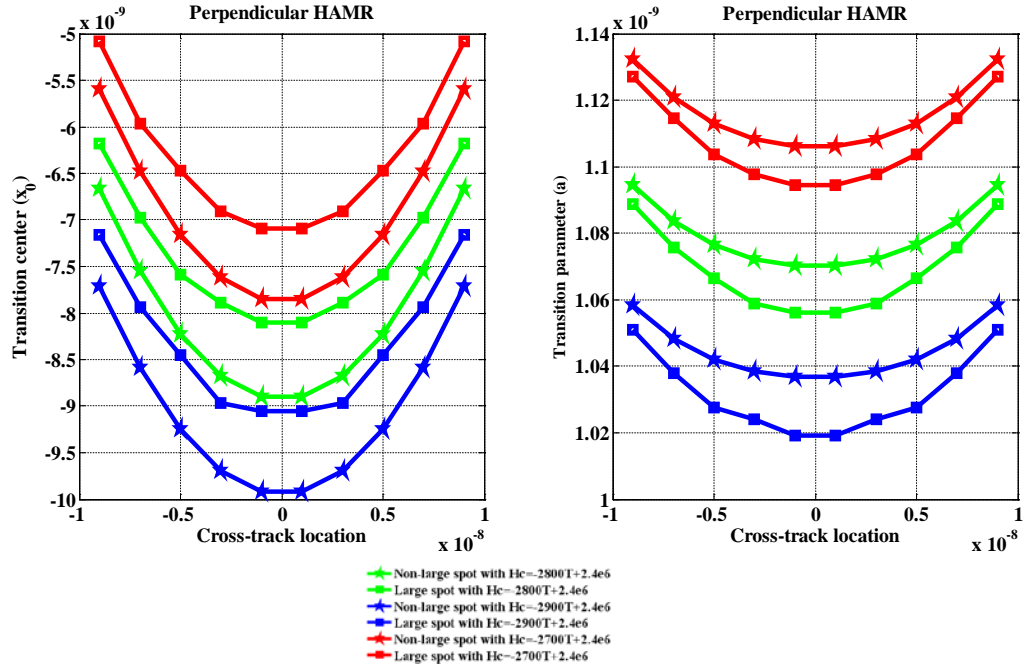
Fig. 4.1. Transition center and parameter with various H_c dependencies on temperatures.
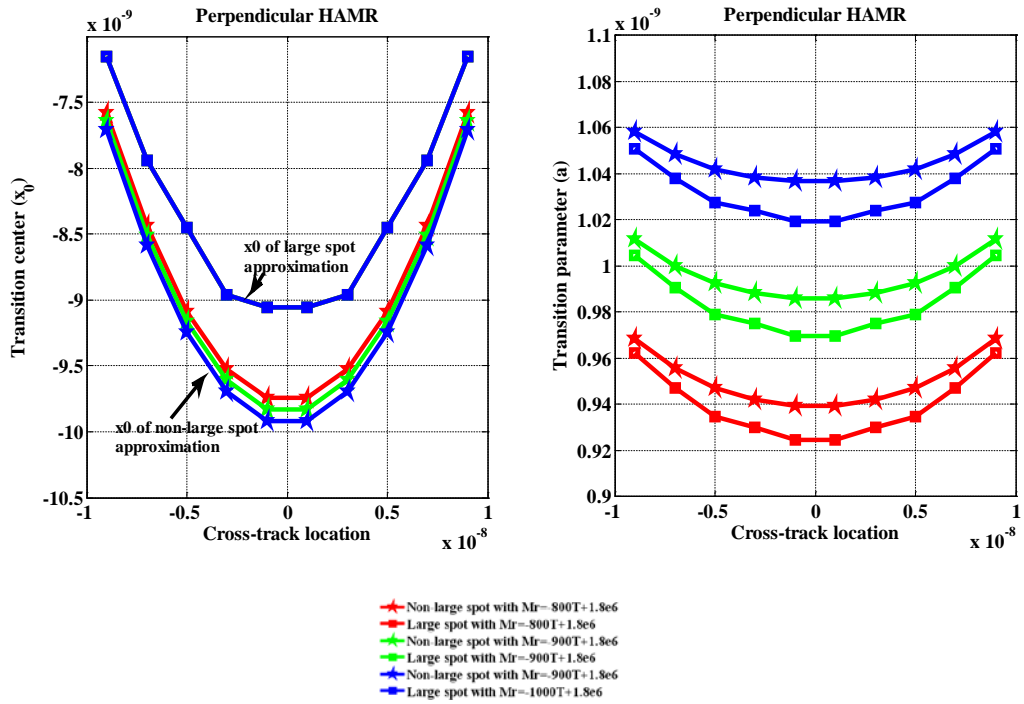

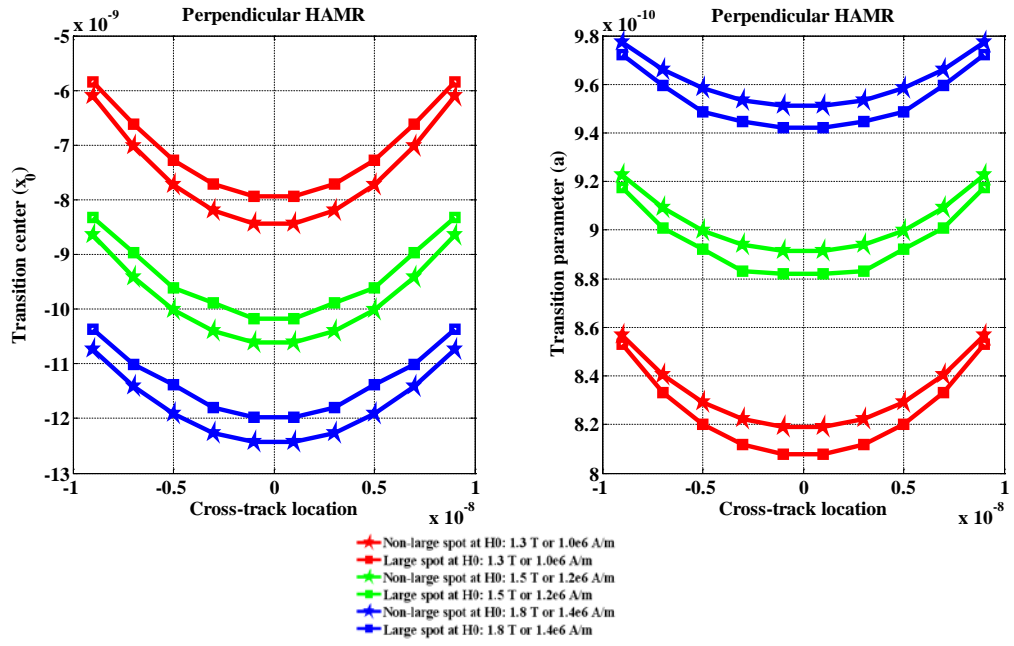**Fig. 4.2.** Transition center and parameter with various M_r dependencies on temperature.

Fig. 4.3. Transition center and parameter with various head fields.
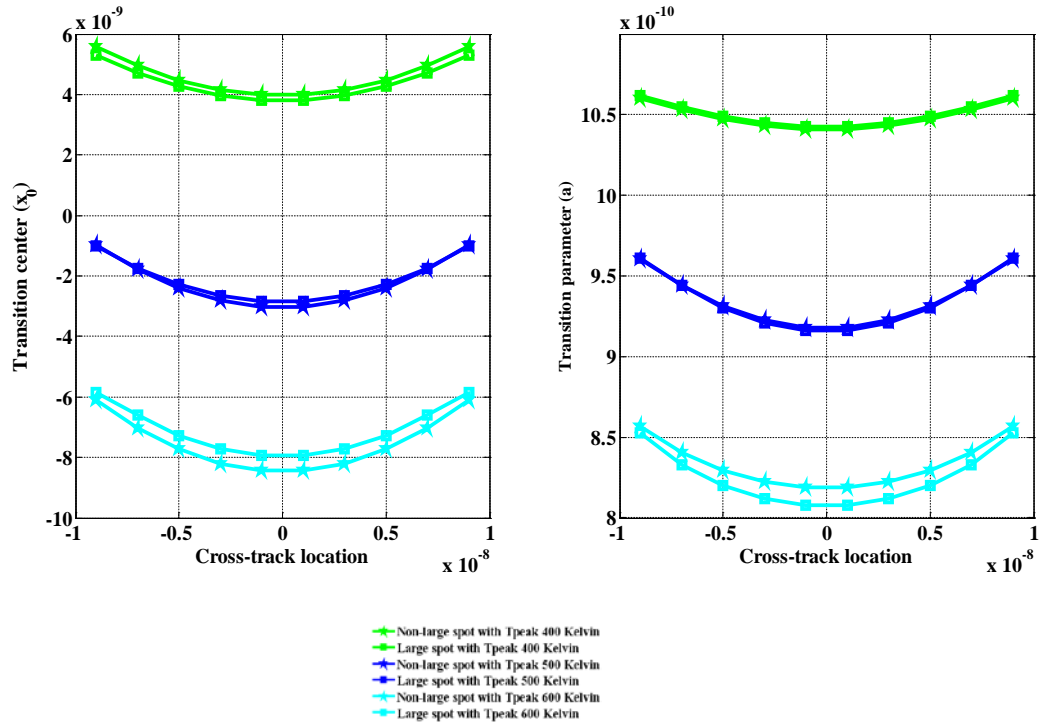
Fig. 4.4. Transition center and parameter with various peak temperatures.

35

Next, we study x_0 and a by focusing on the head field (H_0). The results show that x_0 is shifted far from the laser position and a is wider at high head field as shown in Fig. 4.5. Finally, using high H_c, low M_r and low H_0 to evaluate the system by varying peak temperature (T_{peak}), the results of this experiment show that high T_{peak} will give small a as shown in Fig. 4.6.

As the discussion before, we select high H_c, low M_r, low H_0 and high T_{peak} to evaluate the system. The transition response $p(x)$ and bit response $h(x)$ can be achieved and shown in Fig. 4.7. We get PW_{50} equal to 13.988 nm and hence ND is about 2. The HAMR playback signal is obtained by the convolution between b_k and p_x and corrupted from AWGN. The playback with and without AWGN is shown in Fig. 4.8. Playback signal from LPF is also shown in this Fig. 4.8.
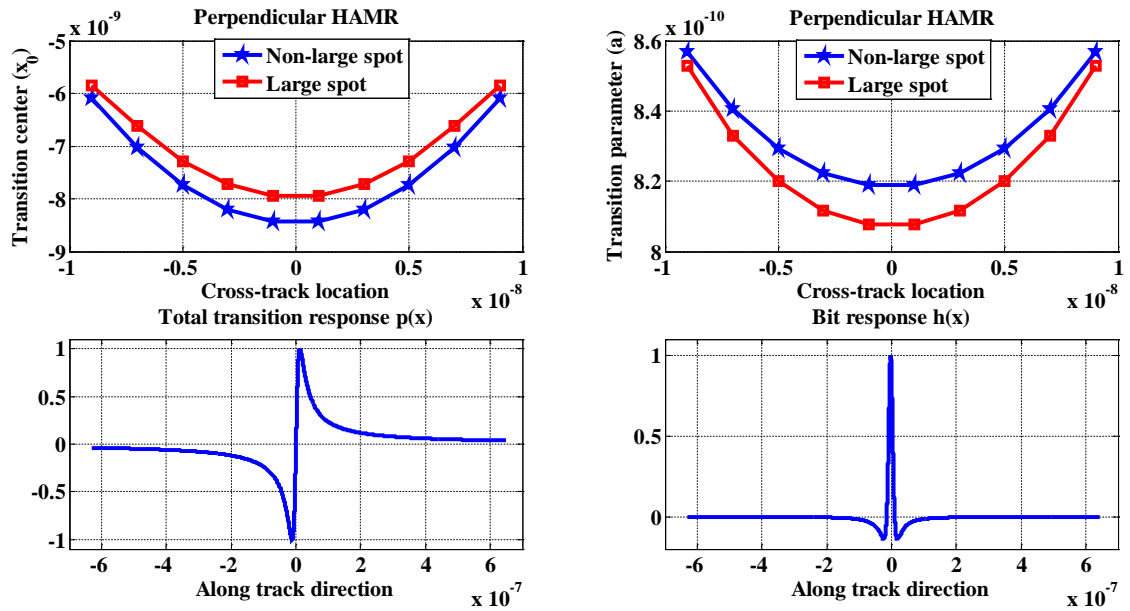
Fig. 4.5. Transition center, parameter, response and bit response.
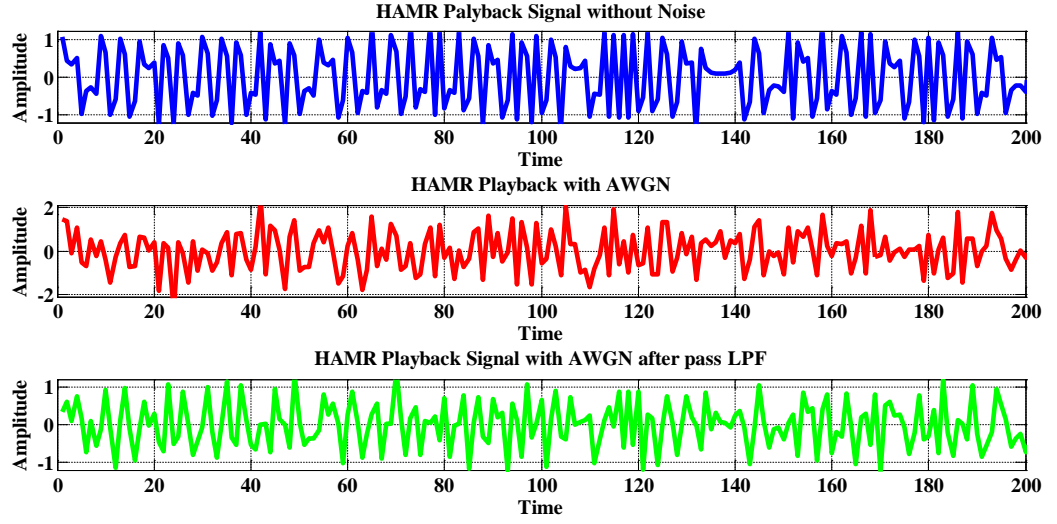
Fig. 4.6. Playback signal with and without AWGN.

4.1.2. HAMR System

The laser can be positioned either in the direction of the head movement (up-track or *+x*) or opposite to it (down-track or *-x*). However, in this work, the laser is assumed to be at the center of the track in the cross-track direction. The HAMR channels with equalizer design is shown in Fig. 4.7.
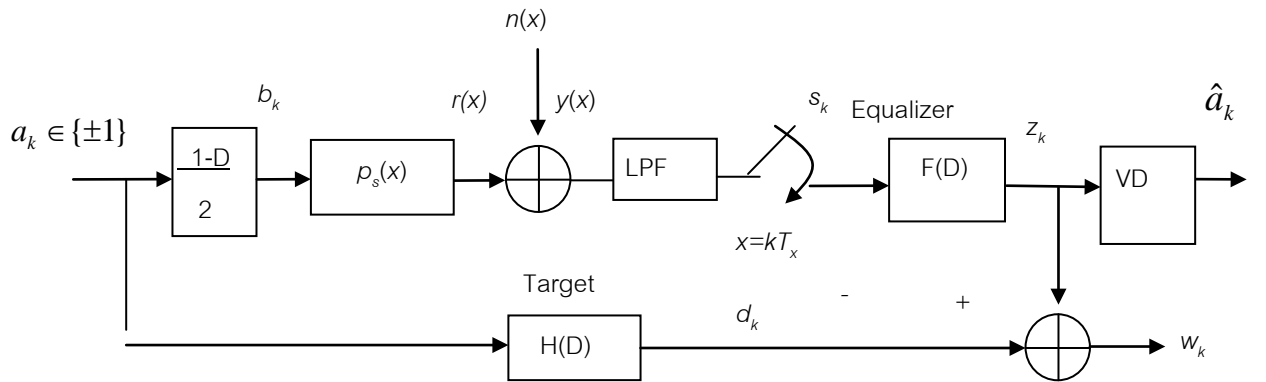
Fig. 4.7. The HAMR system with target-shaping equalization.

where $a_k \in \{\pm 1\}$ is input sequences and filtered by using ideal differentiator (1-*D*)/2. The sequence of the transitions is $b_k \in \{\pm 1, 0\}$, where $b_k \in \{\pm 1\}$ represents the positive and negative transitions, and b_k = 0 means no transition. The playback signal *r(x)* is obtained

by convolution between b_k and transition response p_k and then corrupted by the additive white Gaussian noise (AWGN). In this simulation, SNR is defined as SNR = $10\log_{10}(1/\sigma^2)$, where σ^2 is the variance of AWGN. The playback signal with AWGN $y(x)$ will be passed to low pass filter (LPF), sampled and then put in equalizer (FD) to equalize the signal in order to facilitate the application of Viterbi detector (VD). Finally, the equalized outputs $z(k)$ are detected by VD.

### 4.1.3. Target and Equalizer Design

We use fixed target constraint according to PR form 1-D^2 of the PR-4 target. For MMSE equalizer design, we use the fixed target constraint to minimize the MSE by setting the various number of equalizer taps. Equalizer taps are in the range of 3 to 21 taps. PR1 gives the lowest MMSE from all the targets for this case. The results are shown in Fig. 4.8. The MMSE value at 11-tap equalizer is equal to 0.526. The 11-tap equalizer coefficients are [0.044 -0.036 0.252 0.128 1.320 2.641 1.356 0.404 0.166 0.033 -0.010]

Next, we compare the MSE using four target constraints for evaluation and the range is from 3 to 21 taps as shown in Fig. 4.9. The unit energy constraint has the lowest minimum MSE. Longer target give higher MSE. The 6-tap target with 11-tap equalizer has the MSE equal to 0.058, while the 10-tap target has MSE of 0.056. The $h_1=0$ constraint gives the MSE of 0.085. For $h_0 = 1$ constraint, the minimum MSE of both targets is 0.105, and the fixed-target constraint using PR1 [1 1] has the highest MSE value equal to 0.554 and 0.578.

Since the unit-energy constraint gives the lowest MSE value, we select the unit energy constraint to evaluate and find the MSE with the various number of target taps. Equalizer taps are also in the range of 3 to 21 taps as the previous simulation. For the 11-tap equalizer, the 10-tap target gives the minimum MSE as shown in Fig. 4.10. The MMSE value is 0.055. In Fig. 4.11, we show MMSE of various constraints using 3-tap target with 11-tap equalizer. The PR2 target is worse than the remaining constraints.
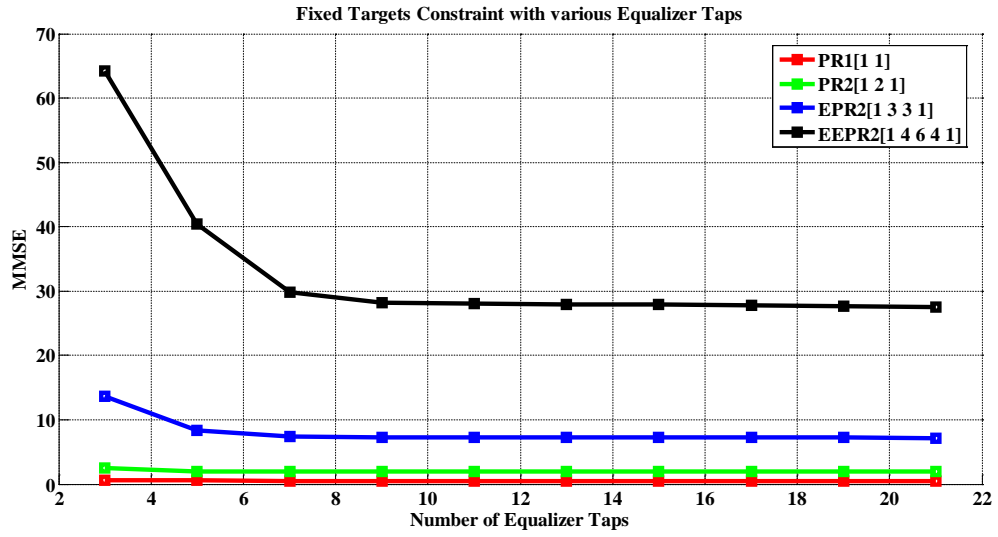
Fig. 4.8. MMSE of fixed-target constraint with various equalizer taps.
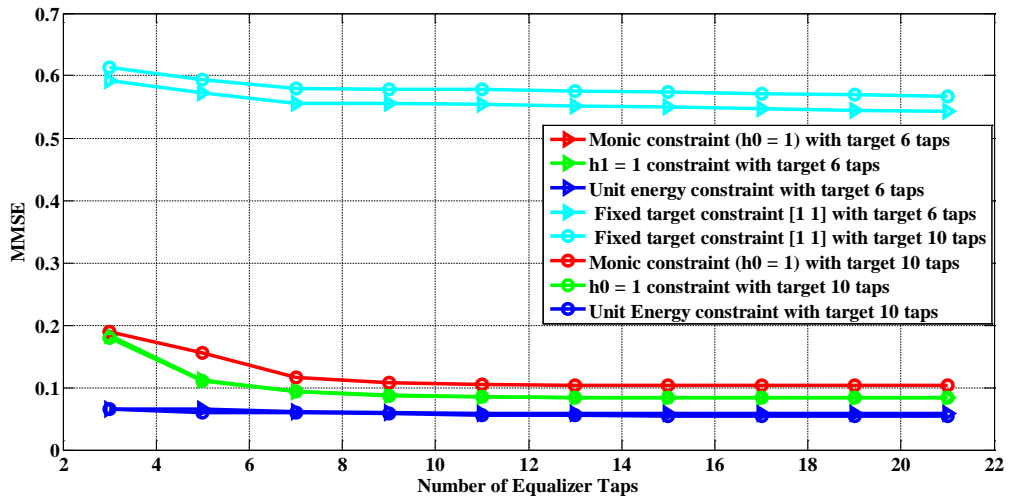
Fig. 4.9. MMSE of various targets constraint using 6 and 10 target taps.
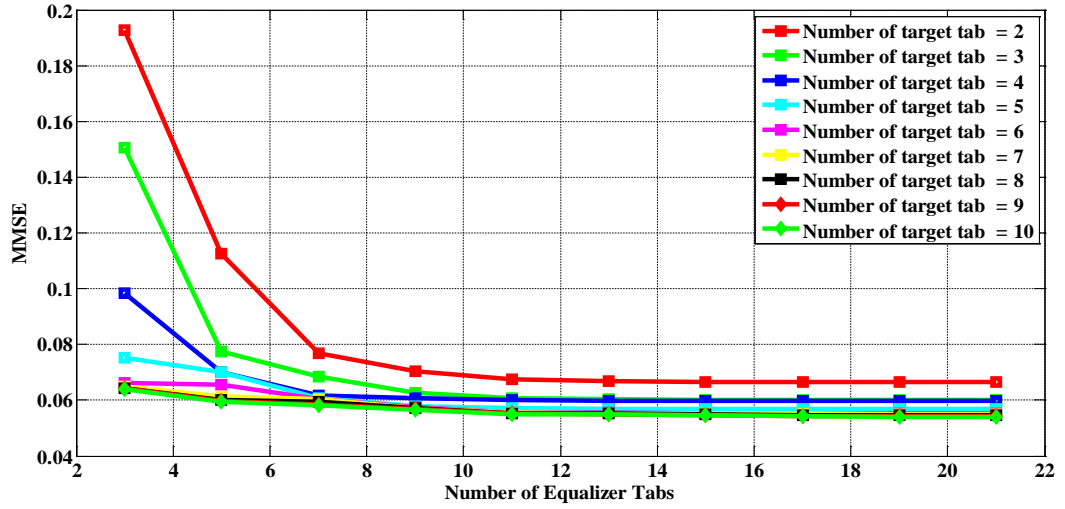
Fig. 4.10. MMSE of unit-energy constraint equalizers.
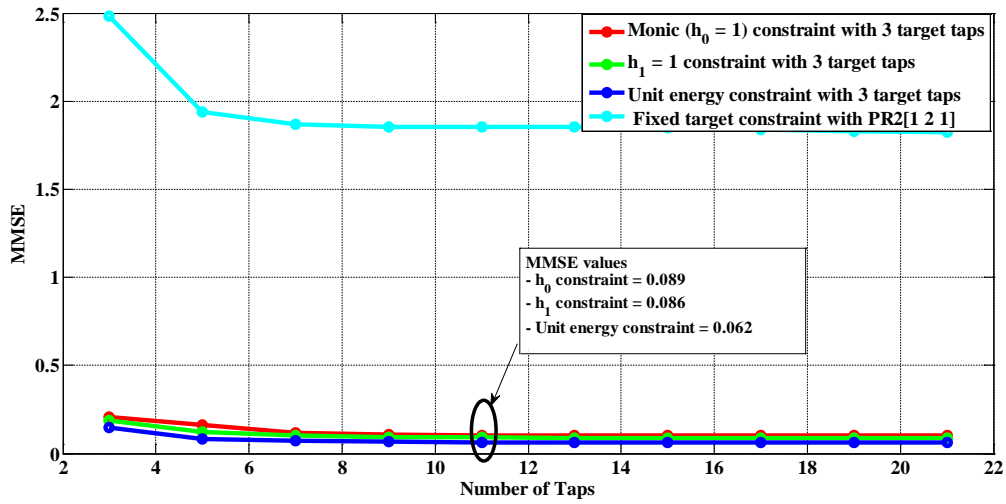
Fig. 4.11. MMSE of various constraints using 3-tap targets with 11-tap equalizer.

As we can see the frequency responses of 4 targets in Fig. 4.12, the frequency response of GPR targets of our system is similar to the frequency response of LMR. Since the frequency response of PR2 target are all dc-full. That is why PR2 target gives high MMSE values and therefore they are not suitable for target design.

40
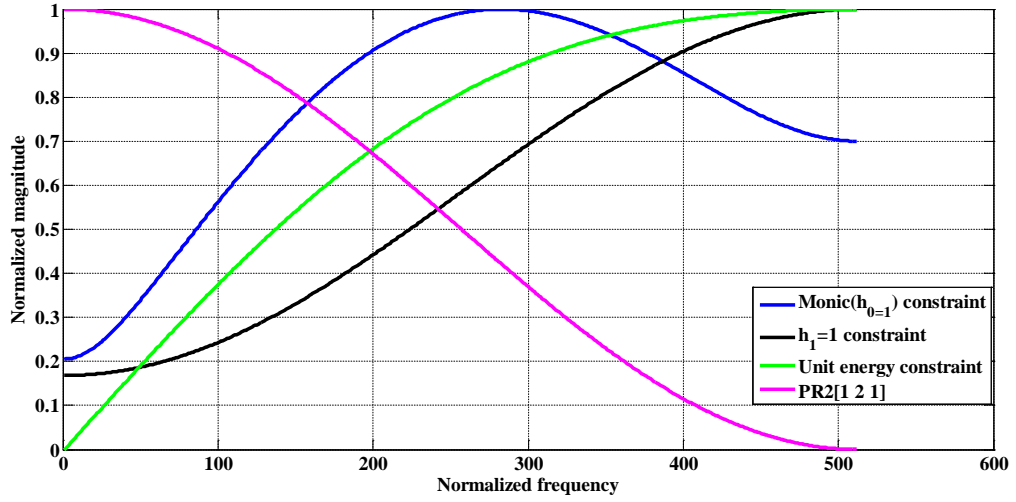
Fig. 4.12. Frequency responses of GPR targets versus fixed-target (PR2).

Fig. 4.13 compares the bit error rates (BER) of the HAMR system using various constraints. At the SNR of 18 dB, the BER of the $h_0=1$ constraint, $h_1=1$ constraint and the unit energy constraint are 1×10^{-4}, 8×10^{-4} and 1.80×10^{-3}, respectively. On the other hand, the BER of the PR2 target constraint is much worse at BER = 1.78×10^{-1}. The BER of the $h_0=1$ constraint gives the best system performance even though the lowest MMSE is from the unit energy constraint as shown in Fig. 4.11. The BER performance of Viterbi detector also depends on the noise correlations and the data patterns. So, the lower MMSE does not necessarily provide the lower BER.
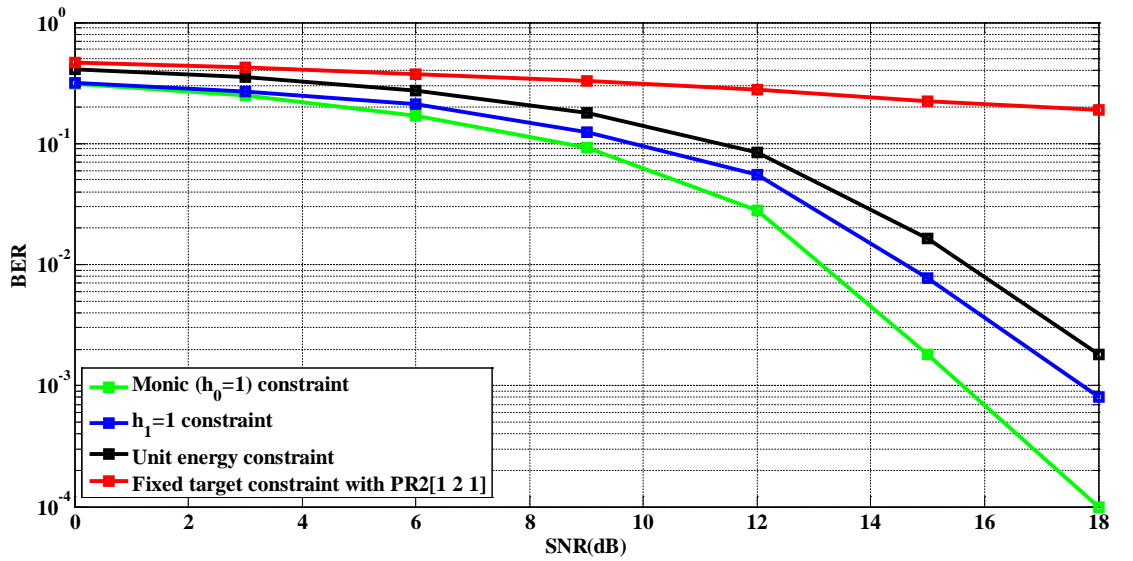
Fig. 4.13. Bit error rates (BER) of the HAMR system.

41

4.2 Detectors for two-dimensional recording channels

4.2.1 Proposed 2-D graph-based detector (M1)

Motivated by [8], where the readback signal from a 2-D channel is equalized to two 1-D targets, vertical and horizontal ones, in this work, the graph-based detectors are used instead of the trellis-based counterpart. In Fig. 4.14, the input sequences $\{x_{j,k-2}, x_{j,k-1}, x_{j,k},$ $x_{j,k+1}, x_{j,k+2}\}$ are written on five tracks at a time and on the receive side, the three tracks of readback sequences $\{r_{j,k-1}, r_{j,k}, r_{j,k+1}\}$ are obtained. In Fig. 4.14, the top (horizontal) 2-D equalizer processes three rows of readback sequences $\{r_{j,k-1}, r_{j,k}, r_{j,k+1}\}$, giving $\{z_{j,k}^{(h)}\}$, then the top equalizer processes the next three rows $\{r_{j,k}, r_{j,k+1}, r_{j,k+2}\}$ to produce $\{z_{j,k+1}^{(h)}\}$ and so on. The bottom (vertical) equalizer processes three columns $\{r_{j-1,k}, r_{j,k}, r_{j+1,k}\}$ at a time, giving $\{z_{j,k}^{(v)}\}$, then the bottom equalizer processes the next three column $\{r_{j,k}, r_{j+1,k}, r_{j+2,k}\}$ to produce $\{z_{j+1,k}^{(v)}\}$ and so on.
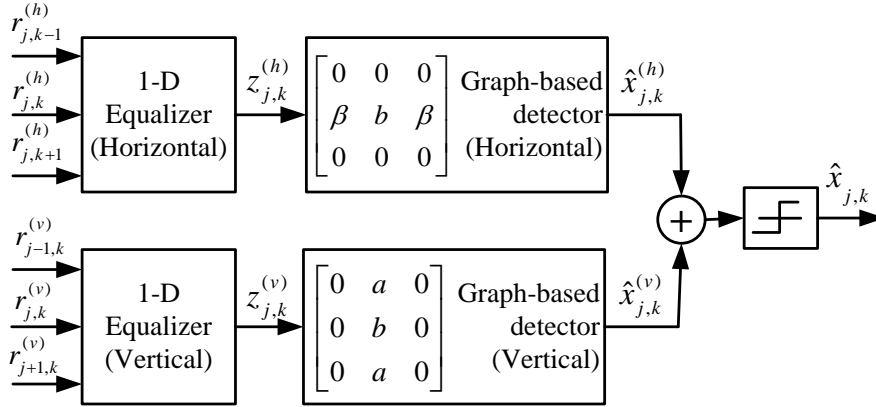
Fig. 4.14. Block diagram of the proposed 2-D equalizers and 2-D graph detector (M1).

(a)

Factor nodes

Bit nodes

(b)
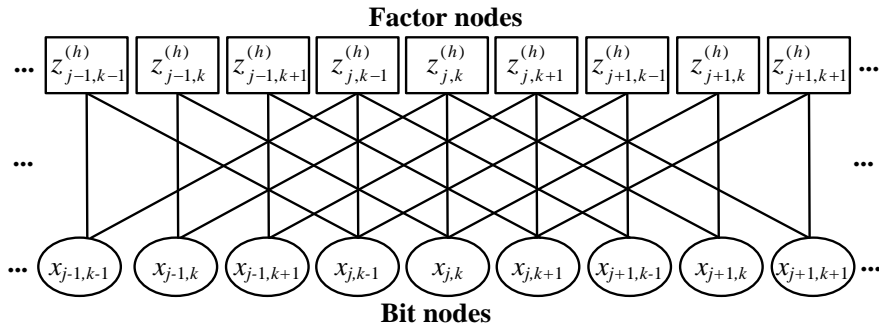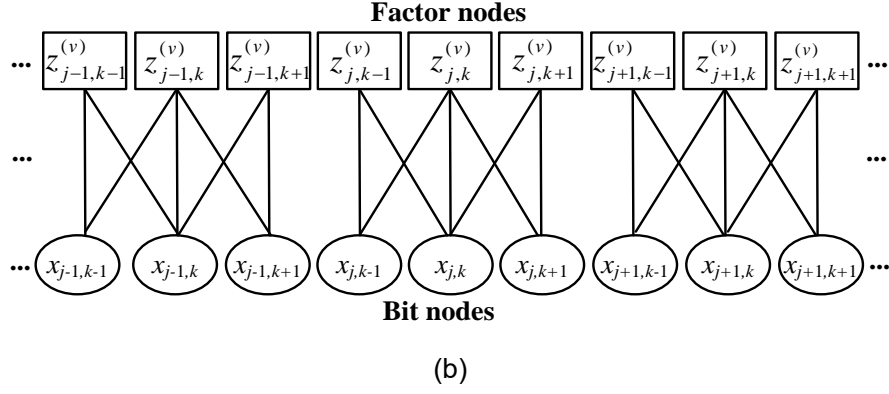
Fig. 4.15. The factor graphs of the (a) horizontal channel and (b) vertical channel.

Since the top equalizer outputs one row at a time, three rows will need to be buffered and then sent to the top (horizontal) graph detector to estimate data of three rows. The bottom (vertical) graph detector, however, estimates data of three columns at a time. The soft information values from these two graph detectors are then added to produce the final bit estimates of three tracks. Actual implementation may vary the buffer size. In Fig. 4.15 (a) and (b), the factor graphs of the 1-D horizontal channel and the 1-D vertical channel, respectively, are shown.

4.2.2. Proposed 2-D graph-based detector (M2)

Since the method (M1) ignores the cornered ITI information but with low complexity due to the 1-D targets, in this method, the 2-D channel is equalized to 2-D targets, with and without cornered ITIs as shown in Fig. 4.16.
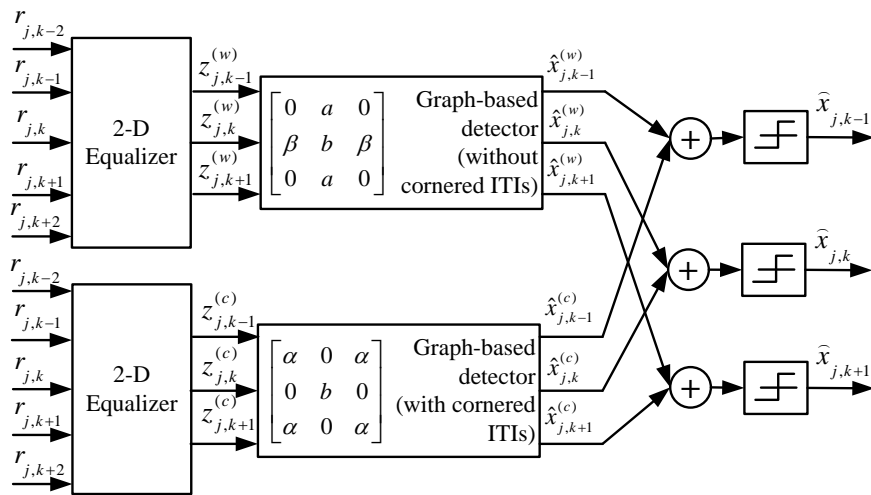
Fig. 4.16. Block diagram for the 2-D graph-based detector (M2).

The factor graphs of the 2-D targets are shown in Fig. 5. Each factor node is connected to 5 bit nodes.

Factor nodes

(a)

Factor nodes

(b)
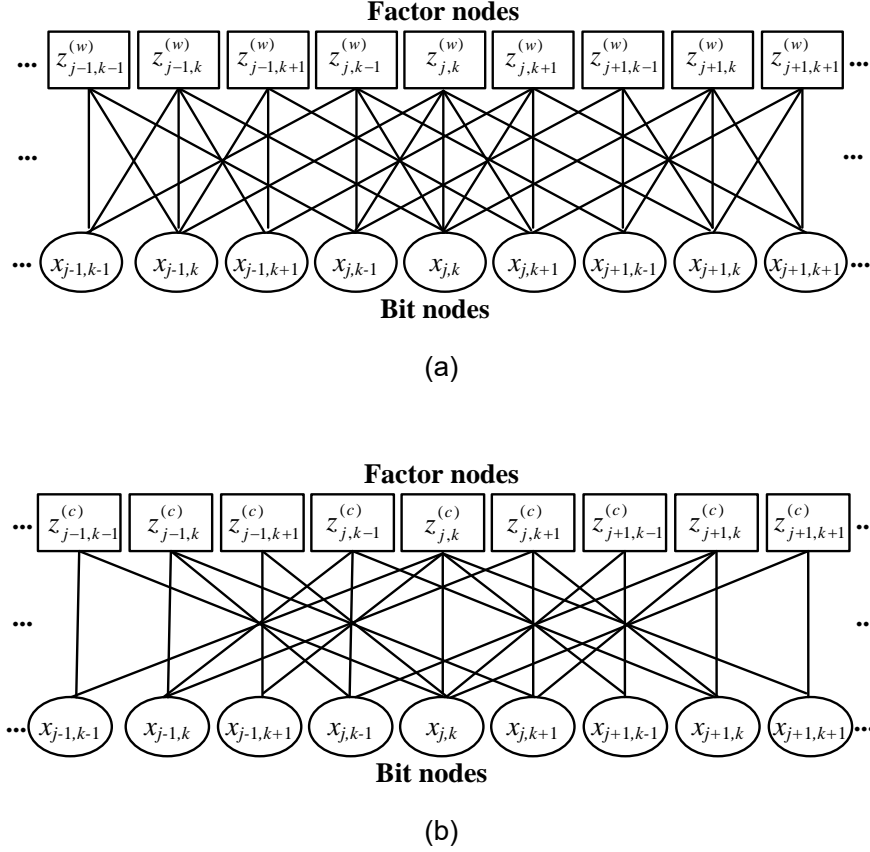
Fig. 4.17. The factor graph of the 2-D graph-based detector (M2) (a) without cornered ITIs and (b) with cornered ITIs.

For simplicity, we measure the complexity of the graph-based detector by the number of edges per factor node. The number of edges per factor node of the full graph-based detector, the method (M1) and the method (M2) is 9, 6 and 10, respectively. The 2-D graph-based detectors (M1) and (M2) can reduce the length-four cycles by about 62% and 31% respectively.

4.2.3 Simulation Results and Discussions

We consider a BPMR system with the areal density of 2 Tbits/in^2, in which both bit period and track pitch are 18 nm., the along track PW50 is 19.4 nm. and the cross track PW50 is 24.8 nm., respectively [1], [3]. The 3×3 matrix channel response is generated by the 2-D Gaussian function. The resultant 3×3 coefficient matrix is

$$\mathbf{H} = \begin{bmatrix} 0.0213 & 0.2321 & 0.0213 \\ 0.0919 & 1.0000 & 0.0919 \\ 0.0213 & 0.2321 & 0.0213 \end{bmatrix}. \qquad (4.1)$$

In the system, each sector includes 4,096 bits, The signal-to-noise ratio (SNR) is defined as $10 \log(V_p/\sigma)$, where $V_p = 1$ is normalized peak value of the readback signal and σ is the standard deviation of AWGN sequence. For the 2-D graph-based detector (M1), the 2-D equalizers for the horizontal and vertical targets have the dimension of 3×7 and 7×3, respectively. In all the figures, the number of iterations for the graph detectors is three.

Fig. 4.18 shows the performance comparison of the three graph-based detectors. The 2-D graph-based detector (M2) performs better than the 2-D graph-based detector (M1) and the 2-D graph-based detector (Full) [18], to achieve the gain of about 0.7 and 1.6 dB, respectively, at BER of 10^{-4}. The full graph detector gives the worst performance mainly due to the large number of length-4 cycles.
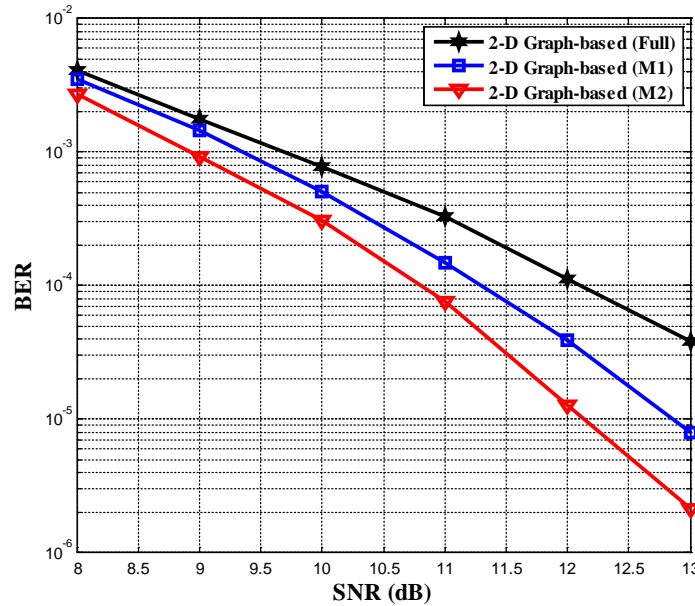
Fig. 4.18. Performance of the 2-D graph-based detector (Full), 2-D graph-based detector (M1) and (M2) of the 2-D channel in (4.1).

In Fig. 4.19, for the fixed ISI levels, we vary the cornered ITI coefficients of the channel matrix in (10) from 2% to 9% and compare the required SNR at BER of 10^{-4} as

the percentage of cornered ITIs coefficients increase. The method (M2) gives the larger SNR gains over the method (M1) and full graph detector as the more cornered ITI levels are present. This is because the method (M2) accounts for the cornered ITIs in the target and reduced the number of length-four cycles.
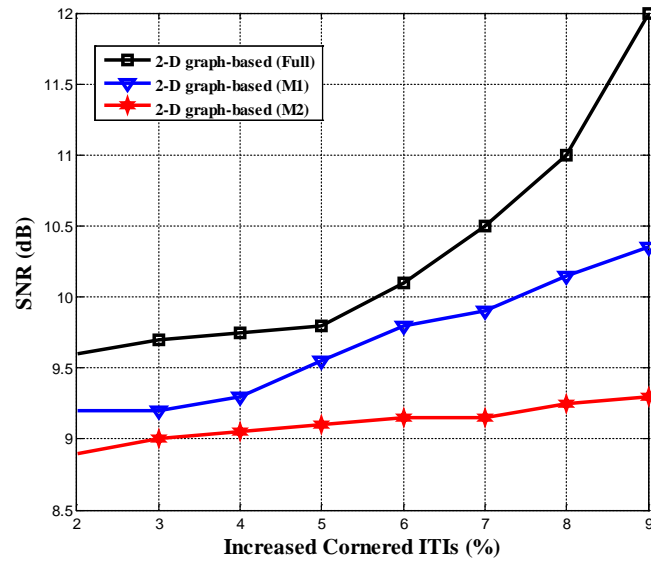
Fig. 4.19. Performance comparison of 2-D graph-based detectors (Full), (M1) and (M2) with increased coefficients of cornered ITIs at the BER of 10^{-4}.
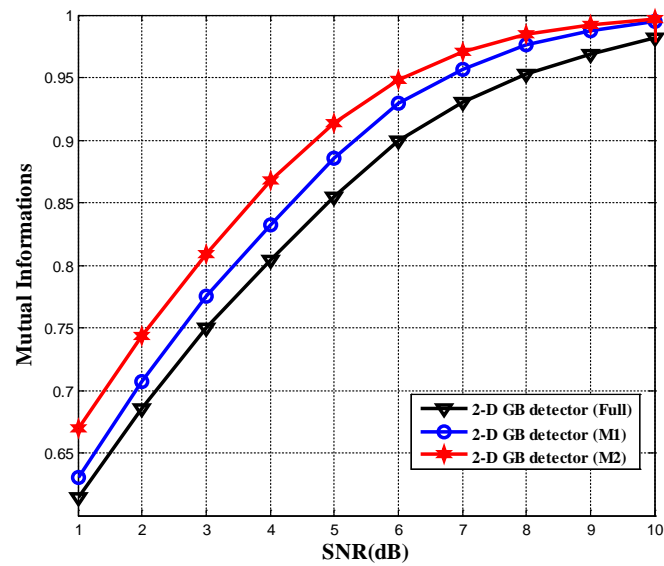
Fig. 4.20. The mutual information of the three 2-D graph-based detectors.
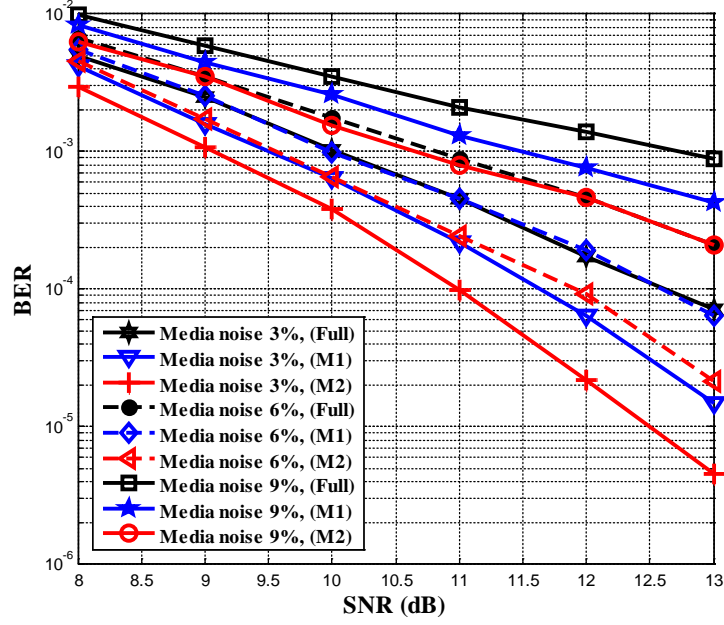
Fig. 4.21. Performance comparison of the three 2-D graph-based detectors in the BPMR system with various media noise levels.

In Fig. 4.20, we the plot the mutual information of the three graph-based detectors at the SNRs from 1 to 10 dB and with 3 iterations. It is evident that the performance of the 2-D graph-based detector (M2) is better than the others. Figure 4.21 shows the performance comparisons of the three graph-based detectors for the BPMR system with media noise due to the location and size fluctuations of the islands [8] at 3%, 6% and 9%. As the media noise levels increase, the performance of the 2-D graph-based detector (M2) still outperforms the 2-D graph-based detector (M1) and the full graph detector (Full).

4.3 LDPC decoder with mixed-schedulings.

4.3.1 Reliability of Variable Nodes in Serial Scheduling

In this section, we investigate the reliability of serial scheduling. We will use the LBP strategy to denote all serial scheduling strategy. In the LBP strategy, the most recent information (from the previous check node update) is used to change the reliability in the last of a series. Hence, LBP has more information than BP when propagating the reliability $L(r_{ji})$ from the last check node. For the LBP with left-to-right direction of check node update

(LBP-LR), the reliability on the left side will affect the reliability on the right side following the direction of the sequential update. On the other hand, For the LBP with right-to-left direction of check node update (LBP-RL), the reliability on the right side will affect the reliability on the left side.

In Fig. 4.22, we show the procedures of updated information for an iteration of LBP. The solid line is the updated information from check nodes to variable nodes and the dashed line is the updated information from variable nodes to check nodes. We use LBP strategy to propagate information on Tanner graphs. Firstly, the LBP strategy will schedule to propagate information from c_1 to v_1, v_5, v_6 and v_7. Afterward, the information is propagated from the variable nodes connected to c_2 and c_3. The information update of 1st check node c_1 is then completed. Secondly, the information from c_2 is propagated to v_2, v_4, v_5 and v_7. The information from these variables are then sent to c_1 and c_3, this is the completion of the 2nd check node update. Finally, after the update of the 3rd check node is finished, the first iteration of LBP-LR as shown in Fig. 4.22(a). The next iterations repeat the fore mentioned processes. The procedure of the check node sequence and information updates of LBP-RL is shown in Fig. 4.21(b).

For each iteration in LBP-LR decoding in Fig. 4.22(a), the node v_5 is updated 3 times twice, while the LBP-RL decoding shown in Fig. 4.22(b). More number of steps indicate that more information is processed before the variable node update. Consequently, the reliability obtained from these two methods are not necessarily the same in Fig. 4.23. In this Fig. 4.23, there are eight iterations and SNR = 4 dB. For the node v_5, or the 5th bit, apposite signs are resulted. Even though these two types exhibit similar bit error rates, the error location may appear at different locations. For the node v_7, even though the number of information updates may be the same for both directions, the total reliability $L(Q_7)$ may differ due to the different update sequence of the check nodes.
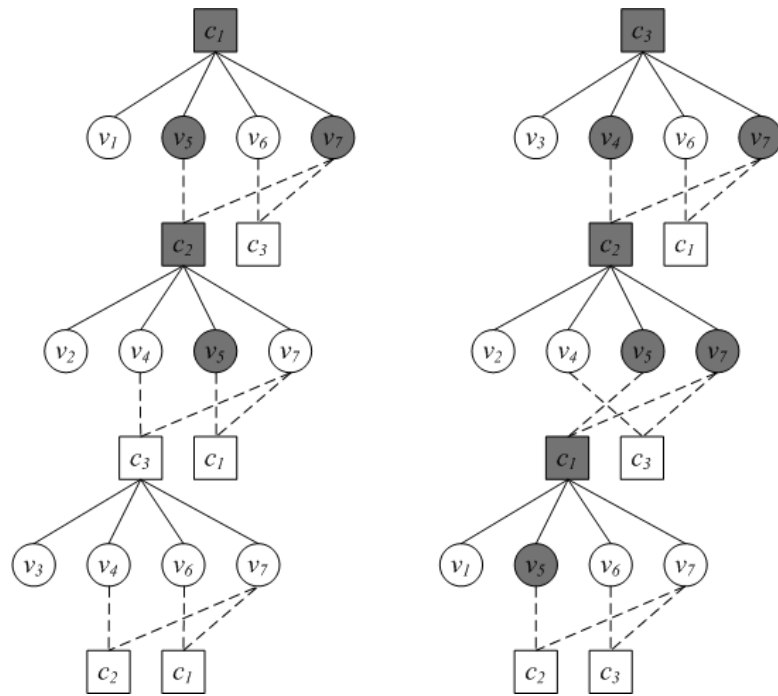
(a) LBP-LR   (b) LBP-RL

Fig. 4.22. Procedures of update information for LBP.
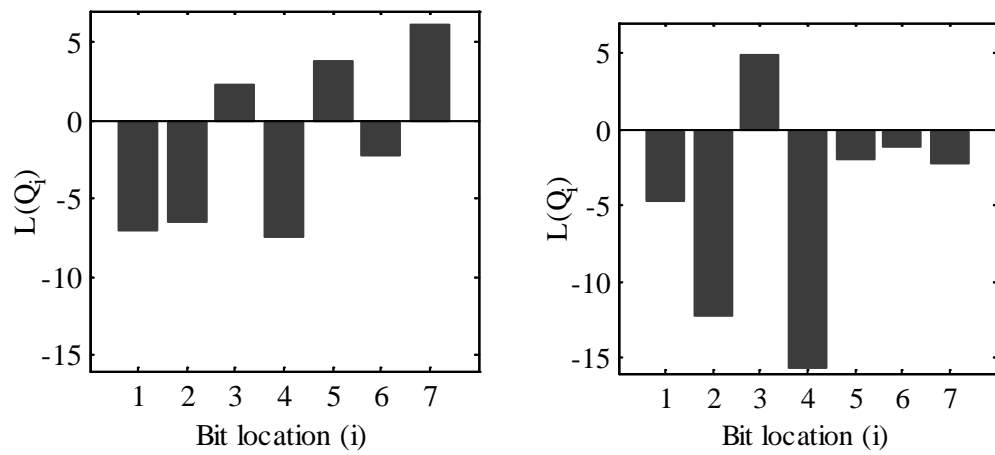
(a) LBP-LR   (b) LBP-RL

Fig. 4.23. The reliability of variable nodes $L(Q_i)$.

4.3.2. Mixed Scheduling for Belief-Propagation Algorithm

In order to obtain more information for LDPC decoding, we propose the mixed scheduling for belief-propagation (MBP). MBP is based on a two-direction scheduling strategy with the bit reliabilities from the left-to-right (LR) direction update and the right-to-left (RL) direction update are combined. Both LBP and SBP algorithms can be used for such purpose. The information from check nodes to variable nodes for the LR direction is denoted by $L_{LR}(r_{ji})$, while that for the RL direction is denoted by $L_{RL}(r_{ji})$. The reliability of LR and RL are generated from

$$L_{LR}(R_i) = \sum_{j \in N(V_i)} L_{LR}(r_{ji}) \, , \qquad (4.1)$$

and

$$L_{RL}(R_i) = \sum_{j \in N(V_i)} L_{RL}(r_{ji}) \, . \qquad (4.2)$$

After the specified number of iterations, the MBP calculates soft information for each bit, that is,

$$L(Q_i) = 2L(v_i) + L_{LR}(R_i) + L_{RL}(R_i) \, . \qquad (4.3)$$

4.3.3 Implementation

The MBP is a two-direction scheduling strategy for LDPC decoding. The LBP is used to sequentially update the check node (SC). We call MBP-C and MBP-V for the MBP based on LBP strategy and SBP strategy, respectively.

For efficient implementations, the forward-backward algorithm [10], [18] is the most popular method used to compute the information from check nodes to variable nodes. The hyperbolic tangent core operation is required for forward and backward process. The number of core operations for check updates are typically used for the comparison of hardware complexity purpose [10], [18]. BP and LBP decoding require $3(d_c-2)$ core

operations to compute the information from a check node and SBP requires $d_c(d_c-2)$ core operations for the same task. For MBP algorithm, since the decoding is based on the serial scheduling of belief-propagation, it requires a similar hardware complexity to LBP and SBP, but MBP exhibits more latency in computing the updated check nodes than both LBP and SBP. MBP can exhibit a similar latency to LBP and SBP by implementing both directions simultaneously, but more core operations are required. MBP-C method requires $6(d_c-2)$ core operations, while MBP-V requires $2d_c(d_c-2)$ core operations.

4.3.4 Simulation Results

This section presents the performance of LBP, SBP and the proposed MBP in the additive white Gaussian noise (AWGN) channel. The performance is plotted in terms of the bit error rate (BER) versus signal-to-noise ratio (SNR) in terms of E_b/N_0, where E_b is the bit energy and N_0 is a one-sided power spectral density of AWGN. All the simulations use the QC-LDPC codes. In Fig. 4.24, the performances of the different types of the decoding schemes are shown. The block size is 4064, the code rate considered is 0.88, both of which approximate the code rate and the sector size used in magnetic recording system [6]. The results reveal that the serial scheduling of belief-propagation has a significantly better performance than traditional LDPC decoding. It can be seen that both LBP-LR and

Fig. 4.24. Performance of BP, LBP, SBP and MBP after 15 iterations of the block size is 4064 and the code rate is 0.88.

SBP-LR perform almost indistinguishably from LBP-RL and SBP-RL respectively. The MBP shows the SNR improvement over other types of serial scheduling. The SNR gain of MBP over LBP and SBP is about 0.15 dB at BER = 10^{-6}.

Figure 4.25 and 4.26 show the performance of BP, LBP, SBP and MBP for shorter block size. We construct the LDPC codes at high code rates and medium code rates of the block size is about 1944 for the IEEE 802.11n standard [30]. These simulations show that MBP slightly outperform the LBP, SBP and BP. The performances of MBP-C and MBP-V exhibit comparable performance, however, the SNR gain of MBP-C over the traditional LBP is about 0.25 dB at BER = 10^{-6} illustrated in Fig. 4.25. For medium code rates, in Fig. 4.26, the performance of MBP is about 0.4 dB better than other types of serial scheduling at BER = 8×10^{-5}.

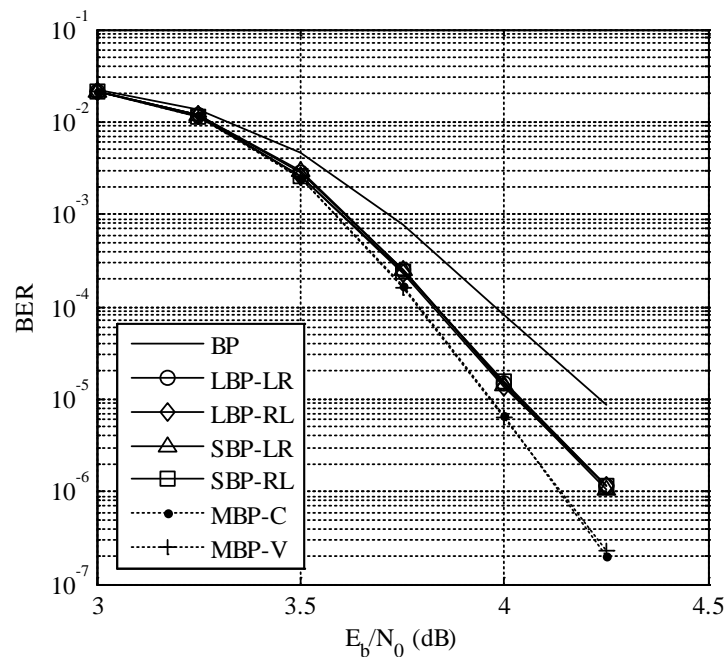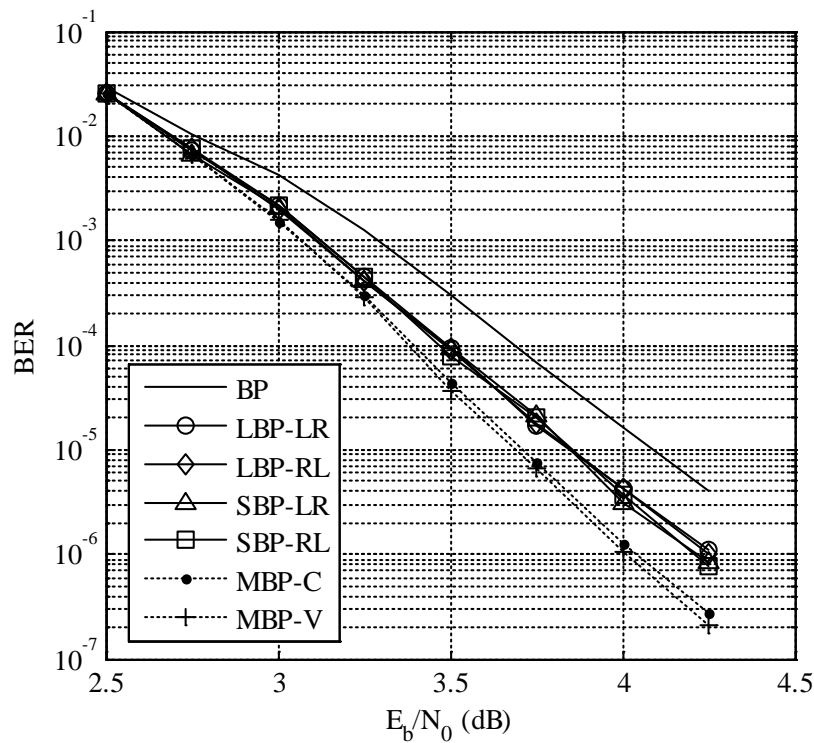**Fig. 4.25.** Performance of BP, LBP, SBP and MBP after 15 iterations of the block size is 1940 and the code rate is 0.8.
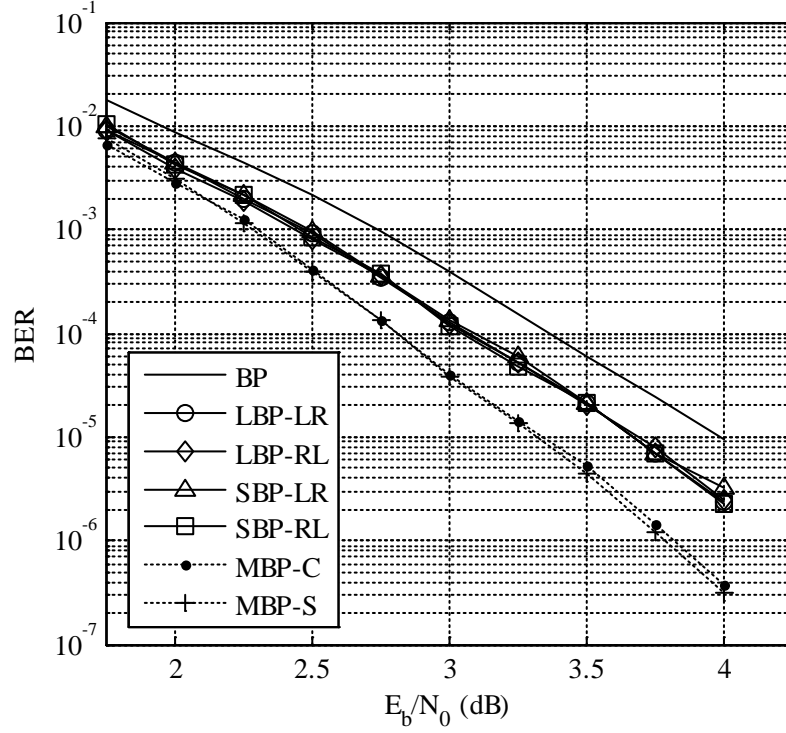
Fig. 4.26. Performance of BP, LBP, SBP and MBP after 15 iterations of the block size is 1930 and the code rate is 0.55.

Figure 4.27 and 4.28 show the comparison between BP, LBP-LR, LBP-RL and MBP at various iterations. We consider the code rates of 0.8 and 0.55, respectively. In Fig. 4.27, for this code rate, the LBP-LR exhibits identical BER until 15 iterations, then shows the lower BER than the LBP-RL afterward. On the other hand, the MBP yields a better performance than both types of LBPs at all iterations. At BER = 9×10^{-5}, MBP requires 15 iterations as opposed to 35 iterations for LBP-LR. For Fig. 4.28, the performance of MBP after 20 iterations is equivalent to that of LBP after 40 iterations

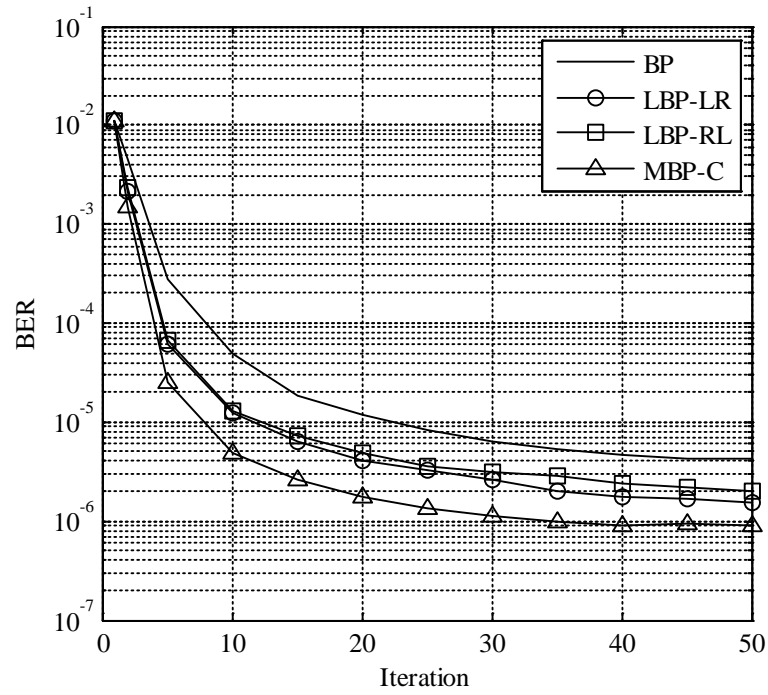**Fig. 4.27.** Performance of BP, LBP and MBP at various iterations for the block size of 1940, code rate of 0.8 and SNR = 3.9 dB.
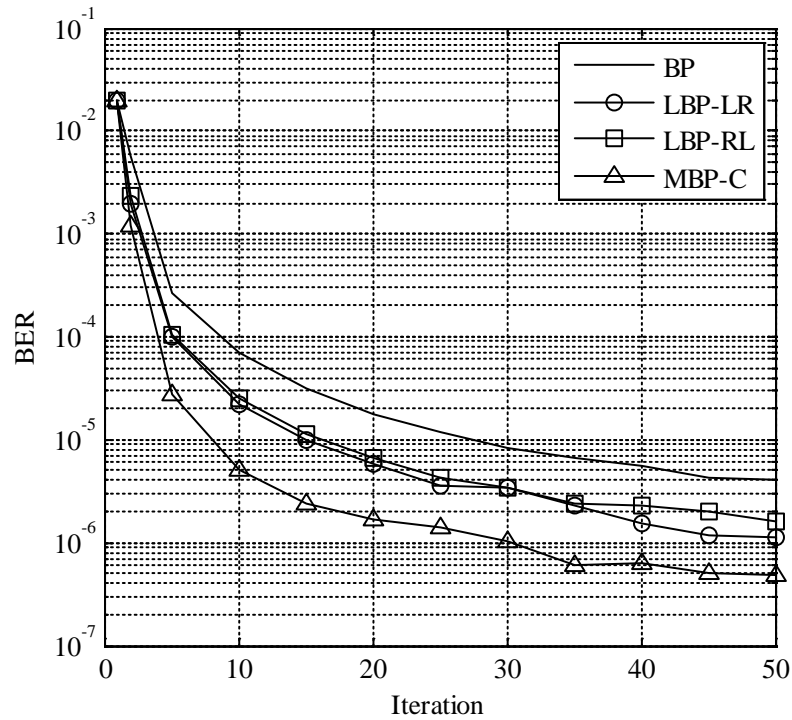

Fig. 4.28. Performance of BP, LBP and MBP at various iterations for the block size of 1930, code rate of 0.55 and SNR = 3.7 dB.

54

4.4 Design of binary and non-binary LDPC codes based on progressive-edge growth (PEG) algorithm and the product codes.

4.4.1. Construction of QC-LDPC Codes Based on PEG Algorithm

Even if the random LDPC codes perform superior error correcting performance [23], the encoding cannot be done using a simple encoder. Alternatively, construction of LDPC codes based on array of circulant matrix or the quasi-cyclic LDPC codes that can be efficiently encoded using feedback-shift register. A QC-LDPC code consists of a *c×t* array of circulant matrixes. A circulant matrix is a square matrix with dimension *p × p* in which each row is the cyclic shift of the previous row, and the first row is the cyclic shift of the last row. The parity-check matrixes for QC-LDPC follow the format

$$
\mathbf{H} = \begin{bmatrix}
\mathbf{H}_{0,0} & \mathbf{H}_{0,1} & \cdots & \mathbf{H}_{0,t-1} \\
\mathbf{H}_{1,0} & \mathbf{H}_{1,1} & \cdots & \mathbf{H}_{1,t-1} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{H}_{c-1,0} & \mathbf{H}_{c-1,1} & \cdots & \mathbf{H}_{c-1,t-1}
\end{bmatrix}
\tag{1}
$$

where $\mathbf{H}_{i,j}$ is circulant matrix with $0 \leq i < c$ and $0 \leq j < t$. The QC-LDPC codes based on the PEG algorithm or PEG-QC-LDPC codes [8] have been introduced to offer flexible parameters of parity check matrices.

For the construction of QC-LDPC codes using PEG algorithm, the set of variable nodes S_v are divided into small groups with each group having *p* variable nodes. An edge that is established from the first variable node in a group is determined using the PEG algorithm. The edges of other variable nodes in the same group can be automatically determined according to a circulant constraint. Figure 4.29 shows an example of the circulant constraint with the matrix size of 4×4. The check nodes connected to the variable node v_2 and v_4 lead to the circulant constraint of other variable nodes in group. Note that the edges connected to variable nodes v_2 and v_4 are determined using the PEG algorithm. The PEG-QC-LDPC algorithm can be briefly summarized in Table 4.2.
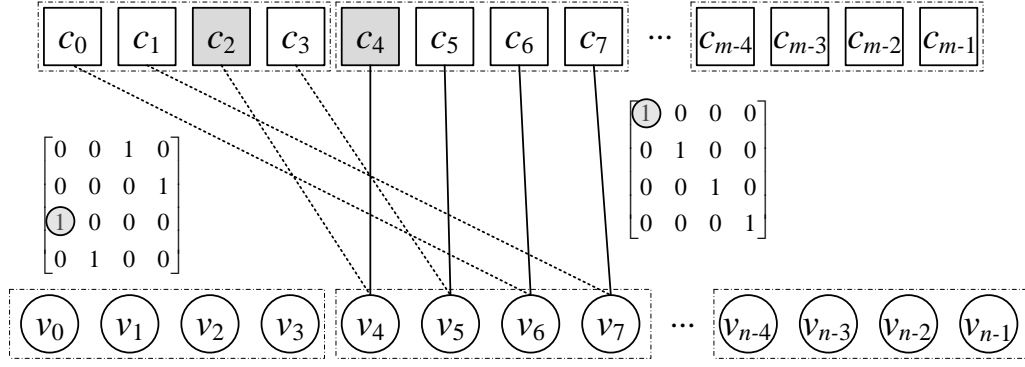
Fig. 4.29. An example of a circulant constraint with matrix size of 4×4.

Table 4.2 PEG-QC Algorithm

---

for $k = 0$ to $t - 1$ **do**

begin

  for $a = 0$ to $d_v - 1$ **do**

  begin

    if $a = 0$

      Establish the first edge incident to variable node v_{kp} by selecting the check node c_j which has the lowest degree under current graph.

    else

      Expand the tree diagram from variable node v_i up to depth l under the current graph such that $N_{v_i}^l \neq \varnothing$ but $N_{v_i}^{l+1} \neq \varnothing$, or the cardinality of $N_{v_i}^l$ stops increasing but is smaller than the number of check node m. Then establish the edge incident to variable node v_{kp} by selecting the check node c_j from the $N_{v_i}^l$ which has the lowest check node degree.

    end

    for $l = 1$ to $p - 1$ **do**

    begin

      Establish the edge incident to variable node v_{kp+1} by using the check node $c_{(j/p)p+\mathrm{mod}(j+l,p)}$ which cyclic shifts l positions of c_j in the circulant $\mathbf{H}_{(j/p),k}$.

    end

  end

end

56

4.4.2. PEG-QC-LDPC codes with maximized girth property

Although the generated PEG-QC-LDPC codes are more flexible than the conventional QC-LDPC codes, the girth is not guaranteed to be the maximum because the multiple choice situation is existence. Motivated by this fact, we propose the girth-maximized method and use it to generate the QC-LDPC codes with maximized girth property.

Clearly, the PEG-QC algorithm established circulant matrices instead of the edges compared with the PEG algorithm. The edge establishing of the first variable node in each group is very important because these establishing will have an effect on the edges of other variable nodes. Especially, in case of all check nodes are reached, the multiple choice situation can occur. This situation will affect the local girth more than the multiple choice situation in PEG algorithm because the tree diagram expanding is used in some edges. Therefore, The QC-LDPC codes that are generated using PEG algorithm do not provide the best possible girth.

To generate the QC-LDPC code with maximized girth property, the set of variable nodes S_v are divided into small groups with each group having p variable nodes. The edges are constructed group by group which the edge incident to first variable node in each group is determined using PEG algorithm. The other edges are constructed according to a circulant constraint. When each group establishes the edges already, the local girth of variable node v_{kp} is calculated. Then, the local girth and the edges incident to variable nodes $v_{(k-1)p}$ and v_{kp} are collected. To provide the maximized local girth, the algorithm will destroy all edges incident to variable node $v_{(k-1)p}$ and v_{kp} and then redo the establishing edges of possible construction. The algorithm will select the edge construction that provide the maximum local girth property and establish the edges incident to variable node $v_{(k-1)p}$ and v_{kp}. The resulting local girth is maximized lead to the girth of codes will be maximized as well. The proposed algorithm is greedy in the sense that the algorithm will evaluate possible choices constructing a graph. However, we can limit this situation by specify the number of iteration. Our proposed algorithm can be shown in Table 4.3.

Table 4.3 PEG-QC-MAX Algorithm

---

for $k = 0$ to $t - 1$ **do**

begin

   while the number of iterations is less than the maximum limit **do**

   begin

      Destroy all of edges incident to variable node $v_{(k-1)p}$ and v_{kp}

      for b = 0 to 1**do**

      begin

         for a = 0 to d_v - 1 **do**

         begin

         if a = 0

            Establish the first edge incident to variable node $v_{(k+b-1)p}$ by selecting the check node c_j which has the lowest degree under current graph.

         else

            Expand the tree diagram from variable node v_i up to depth l under the current graph such that $N_{v_i}^l \neq \varnothing$ but $N_{v_i}^{l+1} \neq \varnothing$, or the cardinality of $N_{v_i}^l$ stops increasing but is smaller than the number of check node m. Then establish the edge incident to variable node $v_{(k+b-1)p}$ by selecting the check node c_j from the $N_{v_i}^l$ which has the lowest check node degree.

         end

         for l = 1 to p - 1 **do**

         begin

            Establish the edges incident to variable node $v_{(k+b-1)p+1}$ by using the check node $c_{(j/p)p+\text{mod}(j+l,p)}$ which cyclic shifts l positions of c_j in the circulant $\mathbf{H}_{(j/p),(k+b-1)}$.

         end

      end

      Calculate the local girth of variable node $v_{(k-1)p}$ and v_{kp}. Then, collect the local girth and the check nodes that connect to variable nodes.

   end

   Select the check nodes with the maximum local girth property then establish the edges incident to variable node $v_{(k-1)p}$ and v_{kp}.

end

58

4.4.3. Simulation Results

We first study the local girth of PEG, PEG-QC, PEG-QC-MAX and the other QC-LDPC codes in [32]. The local girth property of each code is shown in Table 4.4. All codes have the block size of 1944 as the IEEE 802.11n standard. The dimension of the circulant matrix is 54 and the degree of each variable node is 3. The code rates considered are 1/2, 2/3, 3/4 and 5/6. For PEG-QC codes, the results confirm that the multiple choice situation will affect the local girth more than the situation in PEG algorithm. For the code rates are 1/2, 2/3 and 3/4, QC-PEG codes generate local girth of 6 while that the PEG codes do not create. The local girth property of PEG-QC-MAX codes that can solve the problem for selected check nodes is superior to all codes. However, for medium code rates, PEG-QC-MAX codes are comparable to the original PEG codes that is a random codes. The other codes in the table provide the poor local girth because the most existing LDPC codes are designed to avoid the girth of 4.

Table 4.5 shows the local girth of codes at the block size is longer. All codes is generated that the length of code word is 4608 (approximate the sector size in magnetic recording system). In intuition, for a long block length, the random LDPC code can be easily generated with the large girth property when the degree of each variable node is a low. The PEG codes still provide the large girth than PEG-QC and conventional QC-LDPC codes at every code rates. At the code rate is 27/32, the PEG codes create 0.89% of the variable nodes that have a local girth of 6 but the PEG-QC and conventional QC-LDPC codes have 54.09% and 94.12 % of the variable nodes, respectively. The local girth of the proposed codes still perform well at various code rates. At the code rate of 1/2, PEG-QC-MAC have no the local girth of 6 and 8 except PEG, PEG-QC and conventional QC-LDPC codes.

In Fig. 4.30, the performance between the PEG, PEG-QC and the PEG-QC-MAX codes are presented. The performance is plotted in terms of the frame error rate (FER) and the bit error rate (BER) versus signal-to-noise ratio (SNR). In this simulation, the block size is 1944 bit and the code rate is 1/2. the dimension of the circulant matrix is 54 and the degree of each variable node is 3. The PEG codes has a slightly better performance than the PEG-QC codes since it has a great local girth. For the PEG-QC-MAX, the performance are superior to other codes. The proposed PEG-QC codes achieves about 0.05 dB gain over PEG-QC codes at the BER of 2×10^{-6}. In Fig. 31, for the code rate is 5/6, the SNR gain of the PEG-QC-MAX codes over the PEG-QC codes is about 0.12 dB at the FER is 5×10^{-6}.

Table 4.4 Local girth property for PEG, PEG-QC and proposed PEG-QC codes at the block size is 1944.

Code rates	LDPC codes	Local Girths (%)			
		4	6	8	10
1/2	PEG	0	0	0.07	99.93
	PEG-QC	0	2.12	45.55	52.33
	PEG-QC-MAX	0	0	0.06	99.94
	Other codes [32]	0	0	100	0
2/3	PEG	0	0	99.99	0.01
	PEG-QC	0	8.39	91.61	0
	PEG-QC-MAX	0	0	100	0
	Other codes [32]	0	87.50	12.50	0
3/4	PEG	0	0	100	0
	PEG-QC	0	22.70	77.30	0
	PEG-QC-MAX	0	0	100	0
	Other codes [32]	0	90.91	9.09	0
5/6	PEG	0	96.45	3.55	0
	PEG-QC	0	96.26	3.74	0
	PEG-QC-MAX	0	89.50	10.50	0
	Other codes [7]	0	94.12	5.88	0

Table 4.5 Local girth property for PEG, PEG-QC and proposed PEG-QC codes at the block size is 4608.

Code rates	LDPC codes	Local Girths (%)				
		4	6	8	10	12
1/2	PEG	0	0	0.04	98.47	1.49
	PEG-QC	0	1.93	33.95	63.20	0.92
	PEG-QC-MAX	0	0	0	88.02	11.98
	Other codes [32]	0	0	100	0	0
11/16	PEG	0	0	92.67	7.33	0

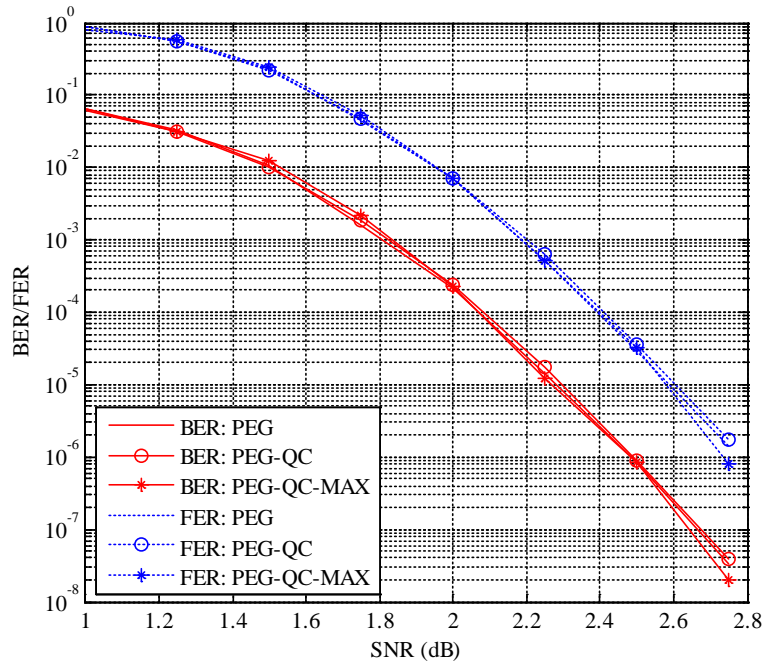|							
	PEG-QC	0	7.28	90.17	2.55	0	
	PEG-QC-MAX	0	0	75.57	24.43	0	
	Other codes [32]	0	87.36	12.48	0	0	
3/4	PEG	0	0	100	0	0	
	PEG-QC	0	13.70	86.30	0	0	
	PEG-QC-MAX	0	0	100	0	0	
	Other codes [32]	0	90.91	9.09	0	0	
27/32	PEG	0	0.89	99.11	0	0	
	PEG-QC	0	54.09	45.91	0	0	
	PEG-QC-MAX	0	0	100	0	0	
	Other codes [32]	0	94.44	5.56	0	0	

Fig. 4.30. Performance of PEG, PEG-QC and PEG-QC-MAX codes with block size of 1944 bits and the code rate of 1/2 after 25 iterations.
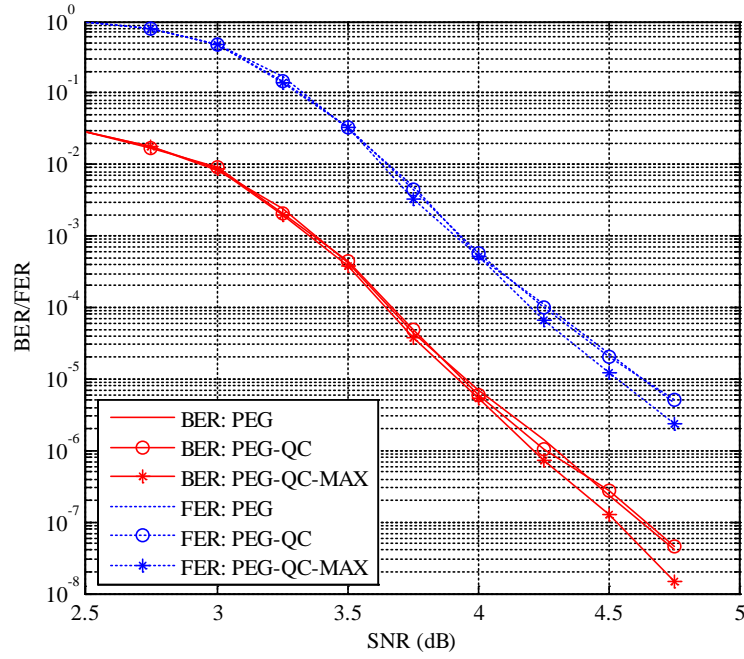
61

Fig. 4.31 Performance of PEG, PEG-QC and PEG-QC-MAX codes with block size of 1944 bits and the code rate of 5/6 after 25 iterations.

4.5 Flash Storage

4.5.1 Simulation results of flash channels

First, we simulate the NAND flash channel for 2 bits/cell as shown in Fig. 4.32. The Erase state has White Gaussian distribution with the mean of 1.4 and standard deviation of 0.3 while the 3 Program states have ΔV_{pp} = 0.2. The target voltages V_p of the state 1 to 3 are 2.6V, 3.2V and 3.93V, respectively.
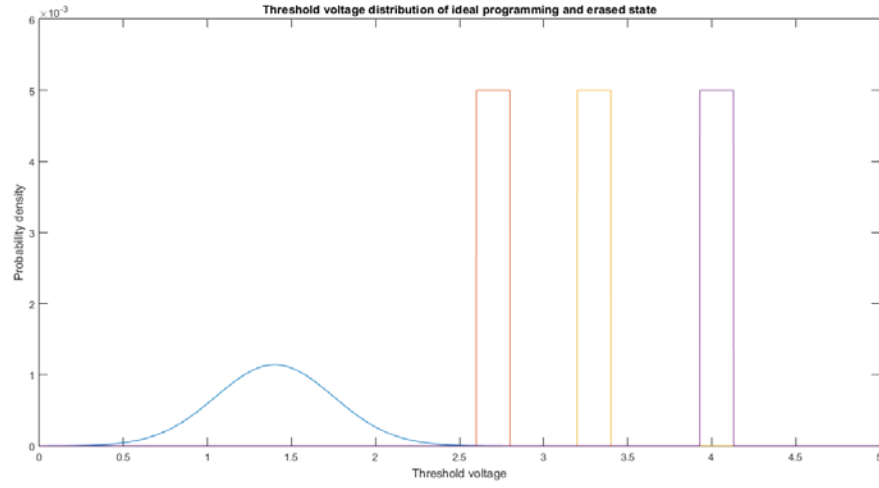
Fig 4.32 Distribution of the Erase state and three Program states

Figures 4.33 to 4.34 show the distribution of the RTN noise with an exponential distribution with K_λ = 0.00025.
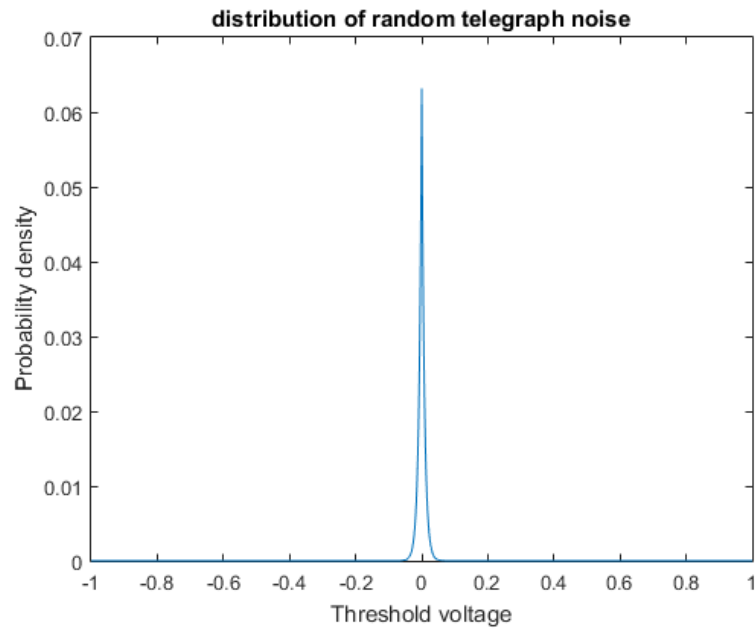

Fig 4.33. Distribution of RTN noise with 1000 times of erase and programming
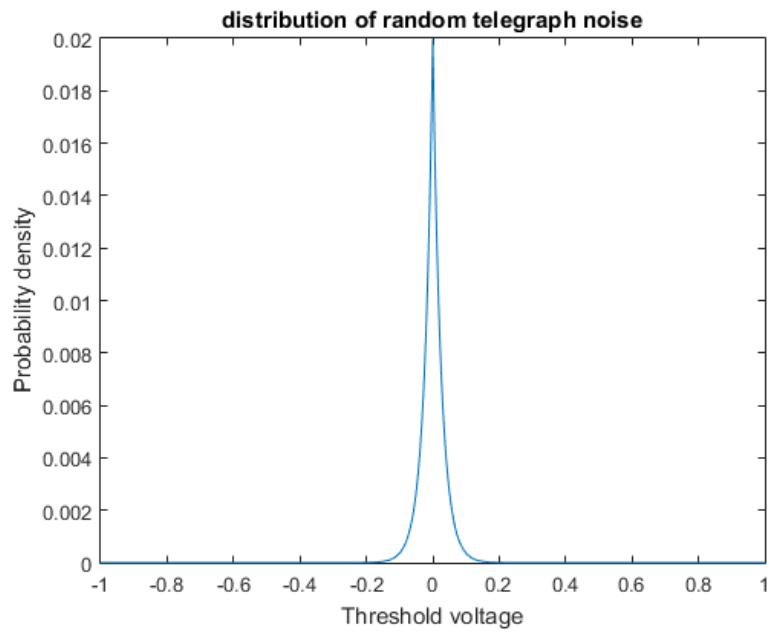
Fig 4.34. Distribution of RTN noise with 10,000 times of erase and programming

After including the RTN noise, the aggregate distributions are shown in Fig. 4.35 and Fig. 4.36
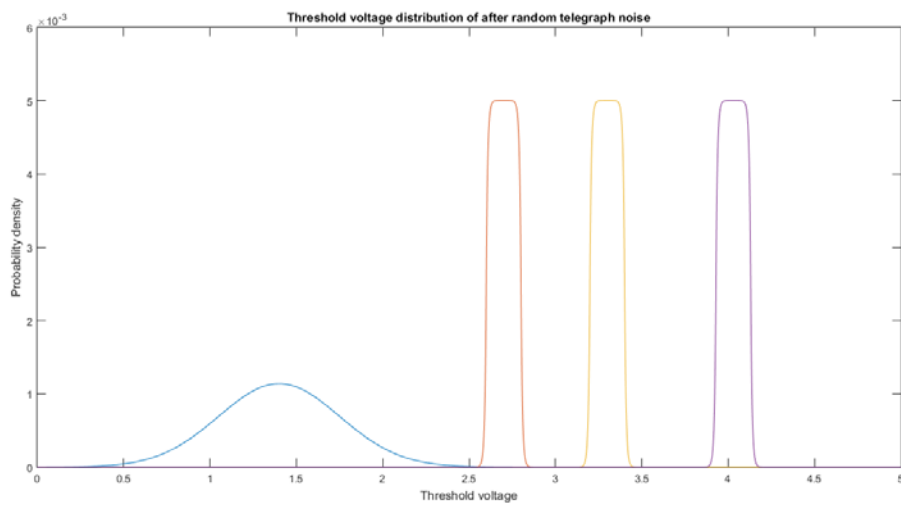
Fig 4.35. Distribution of voltages after RTN noise with 1,000 times of erase and programming
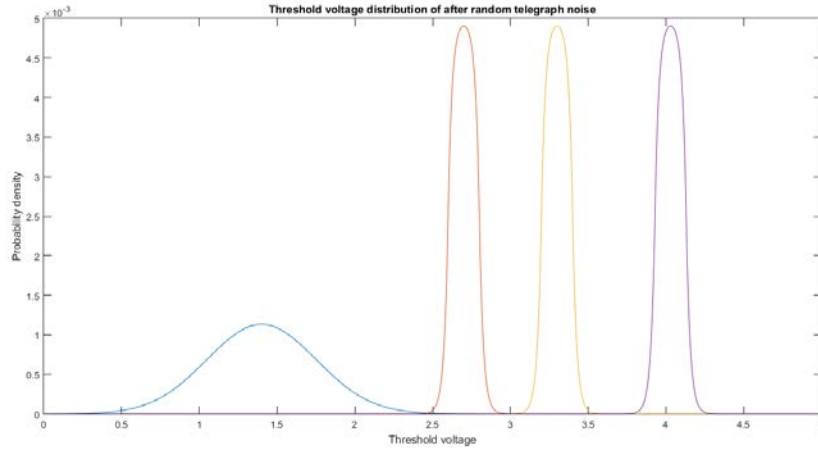
Fig 4.36. Distribution of voltages after RTN noise with 10,000 times of erase and programming

For cell-to-cell coupling, the victim cell is affected by 3 adjacent cells, in the diagonal, left and right side. Let γ_y and γ_{xy} be Gaussian distributed with the mean of 0.008s and 0.0008s, respectively, where s is the cell-to-cell coupling strength = 1.5. After including the cell-to-cell interference, the distribution of the four states is shown in Fig. 4.37.

Fig 4.37. Distribution of voltages after cell-to-cell interference with s = 1.5.

Finally, we simulate the effects of retention. Fig. 4.38 and Fig. 4.39 show the distribution after 10 years of storage and with N = 10,000 and 100,000, respectively.
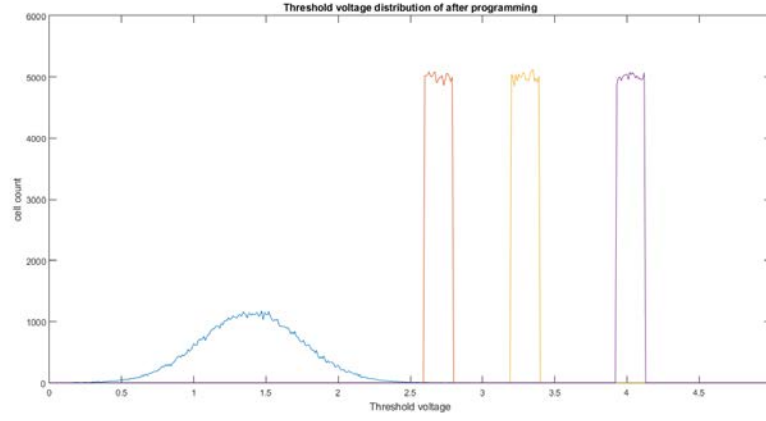
Fig 4.38. Distribution of voltages after retention noise of 10 years (N = 10,000).
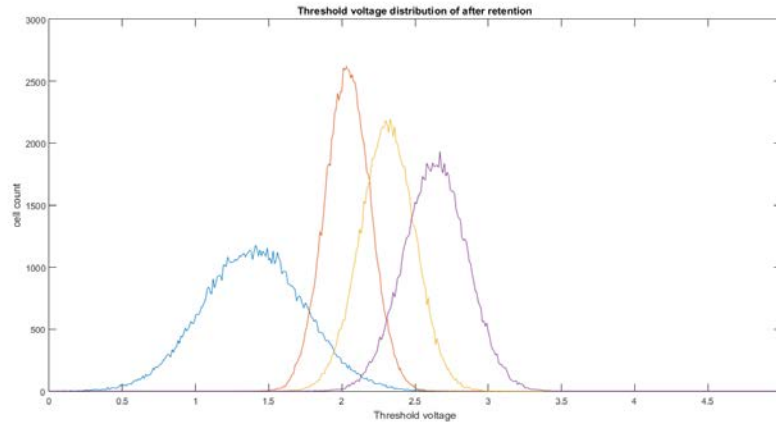
Fig 4.39 Distribution of voltages after retention noise of 10 years (N = 100,000).

4.5.2 Soft decoding of flash channels

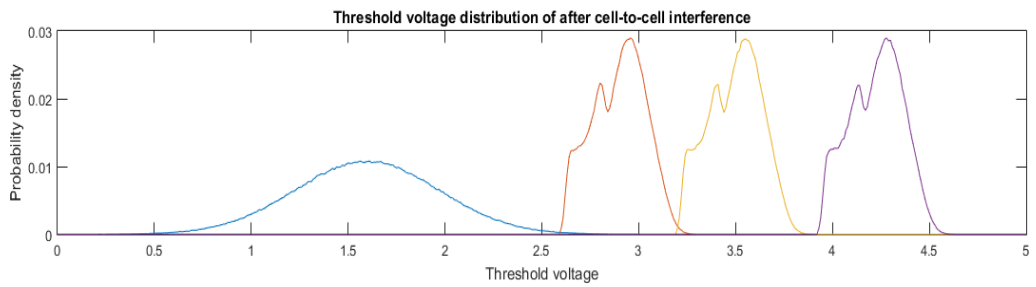Previously, reading the memory cell relies on hard decision based on one single voltage reference. The hard-decisioned bits are protected by BCH codes [52]. However, recent works focus on the use of high-performance error correcting codes such as LDPC codes. Using many levels reference voltages, however, causes latency. In [53], G. Dong proposed irregular reference levels with more resolution around the overlapping pdf values. In [54], the authors proposed the reference voltage levels using maximized mutual information (MMI) optimization. A benefit of the MMI method is that each cell can be read with any amount of time, the lowest bit error rate may not be obtained. In 2016, C.A. Aslam [55] alternatively used the entropy approach together with the additional optimization to give the optimal performance. However, a disadvantage is that it is quite difficult to use especially for non-Gaussian pdf.

In this work, we analyze the LDPC performance together with the reference voltage reading using the *density evolution* method. The reference levels can be mapped to a Discrete Memoryless Channel (DMC), in particular, Binary Erasure Channel (BEC), as shown in Fig. 4.40.

Fig. 4.40. Binary Erasure Channels.
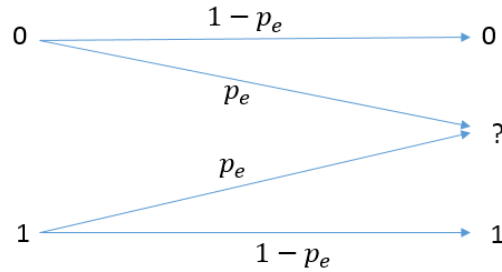
An analysis of LDPC decoder starts from considering an example of regular parity-check matrix in Fig. 4.41 with the corresponding Tanner graph in Fig. 4.42.

Fig. 4.41. An example of a parity-check matrix
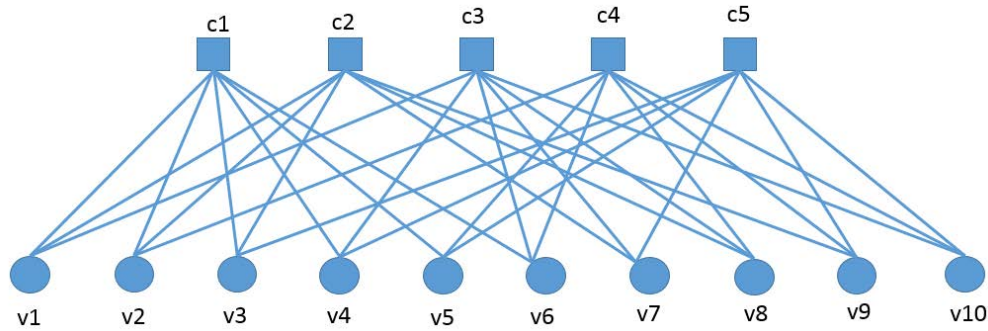
Fig. 4.42. The corresponding Tanner graph of Fig. 4.41.

Each bit node is connected to d_v check nodes, but each check node, on the other hand, is connected to d_c bit nodes.
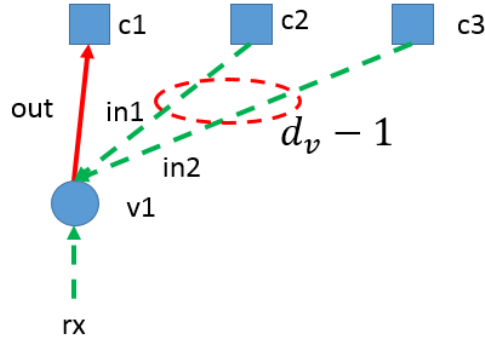
Fig. 4.43. The check node connection.

Consider the bit node v_1 with three incoming connections from 2 check nodes: in_1, in_2 and the received signal r_x. The outgoing value *out* at this bit node is considered *erasure* if and only if all the three incoming paths are *erasures.* The probability that each incoming branch is *erasure* is shown in Fig. 4.44.
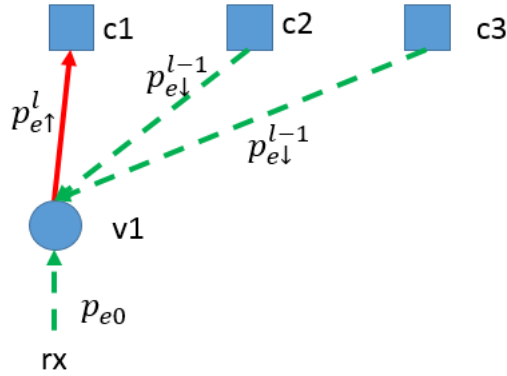
Fig. 4.44. The probability that each incoming branch is *erasure*.

The probability that the information to a check node is *erasure* during the l_{th} iteration, $P_e^l \uparrow$, depends on the probability that the information to the bit node is *erasure* during the $(l\text{-}1)^{th}$ iteration, $P_e^{l-1} \downarrow$, and the incoming is erasure, P_{e0}. Assuming that each incoming path is independent, we have

$$P_e^l \uparrow = P_{e0}\left(P_e^{l-1} \downarrow\right)\left(P_e^{l-1} \downarrow\right)$$

$$P_e^l \uparrow = P_{e0}\left(P_e^{l-1} \downarrow\right)^{d_v - 1}$$

(4.4)

Next, derive the relationship at the check node as shown in Fig. 4.45.

68
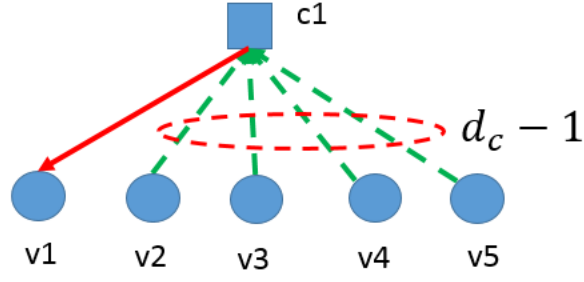
Fig. 4.45. The check node connection.

Consider Fig. 4.46, each check node is connected to bit nodes,
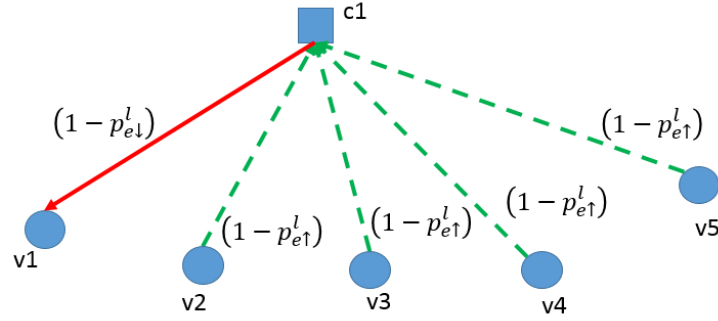
Fig. 4.46. The probability of correct bit node.

so we have the probability of correct incoming bit node

$$\left(1 - P_e^l \downarrow\right) = \left(1 - P_e^l \uparrow\right)\left(1 - P_e^l \uparrow\right)\left(1 - P_e^l \uparrow\right)\left(1 - P_e^l \uparrow\right)$$
$$\left(1 - P_e^l \downarrow\right) = \left(1 - P_e^l \uparrow\right)^{d_c - 1} \tag{4.5}$$

where the probability of a correct bit node, $1 - P_e$, $P_e^l \downarrow$ is the probability that *erasure* information is sent to the bit node at iteration *l* and $P_e^l \uparrow$ the probability that *erasure* information is sent to the check node at iteration *l* and all paths are independent. Rearranging (4.5) gives

$$P_e^l \downarrow = 1 - \left(1 - P_e^l \uparrow\right)^{d_c - 1} \tag{4.6}$$

Replacing (4.6) in (4.4), we have

69

$$P_e^l \uparrow = P_{e0}\left(1-\left(1-P_e^l \uparrow\right)^{d_c-1}\right)^{d_v-1} \tag{4.7}$$

From (4.7), the probability that erasure can be corrected by an LDPC decoder can be obtained. We can plot the result at various iterations as shown in Fig. 4.47.
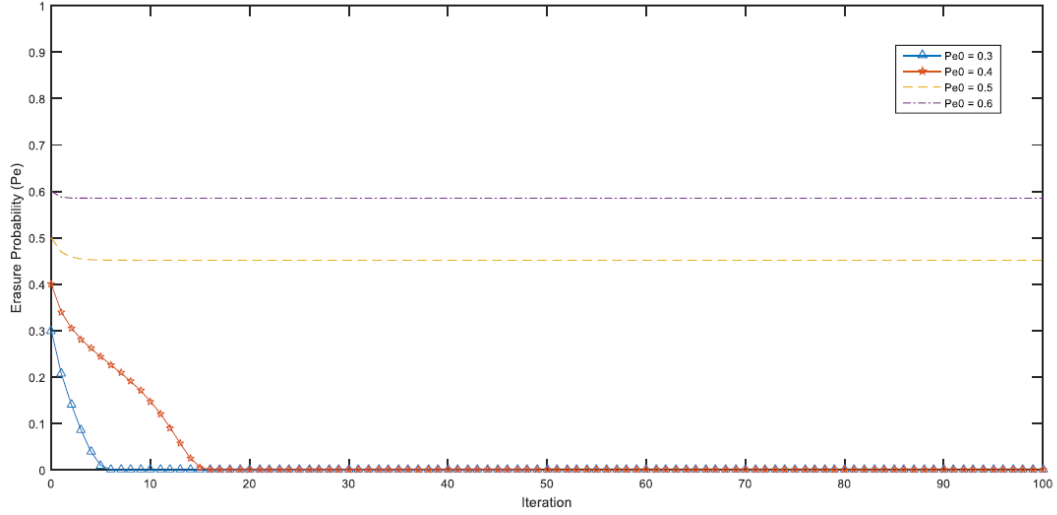
Fig. 4.47. The probability of erasure correction as various iterations.

The blue curve is for P_{e0} = 0.2, while The red curve is for P_{e0} = 0.4. For these two cases, as the iteration number increases, all erasures can be corrected especially after 6 and 15 iterations, respectively. The yellow and purple curves are for the case of P_{e0} = 0.5 and 0.6, respectively, showing that erasures cannot be fully corrected by the LDPC decoder. We can determine that largest amount of erasures which can be fully corrected a parity-check matrix with d_c =3, d_v =6 by using (4.7) and setting the iteration number to be 100 as shown in Fig. 4.48.
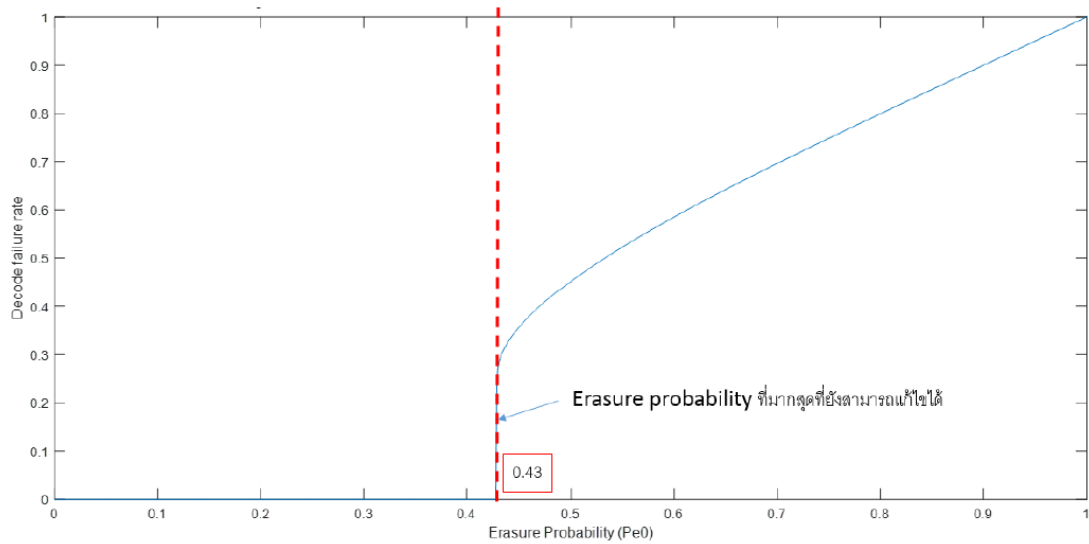
Fig. 4.48. The decoder failure rate $(P'_e \uparrow)$ and P_{e0}

It can be seen that at after P_{e0} =0.43, erasure cannot be corrected. This value is called Maximum corrected Erasure (P^*_{e0}) which depends on d_c and d_v. This result is based on the assumption of infinite codeword length. In practice, the block length is not very long, therefore, to obtain more accurate result, the simulation system in Fig. 4.49 is performed.
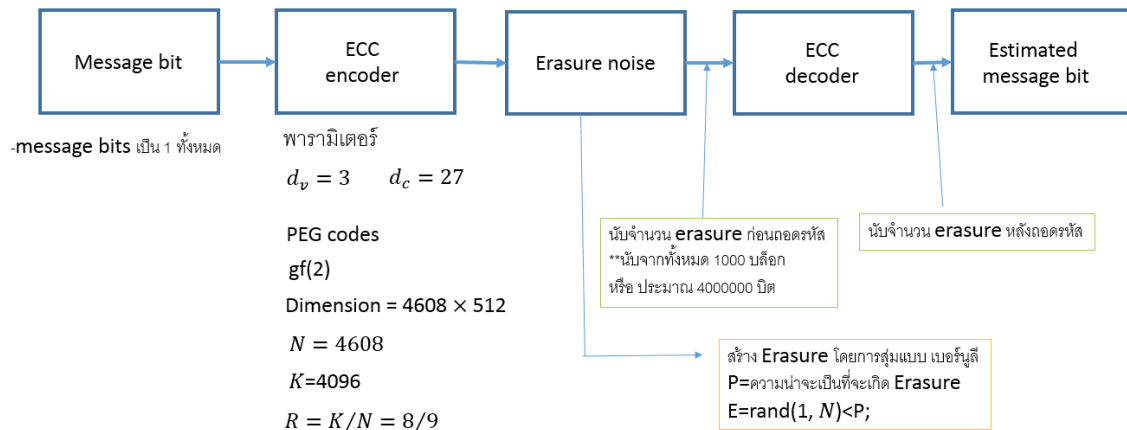
Fig. 4.49. Block diagram of the simulation setup.

In this simulation, all input bits are '1' which are then encoded using the LDPC encoder with the parity-check matrix with d_c = 27 and d_v = 3. The codeword length is

4608 and code rate is 8/9. The erasure is generated at the LDPC encoder and then after the decoder, the remaining erasure is counted. The result from (4.7) is shown in Fig. 4.50.
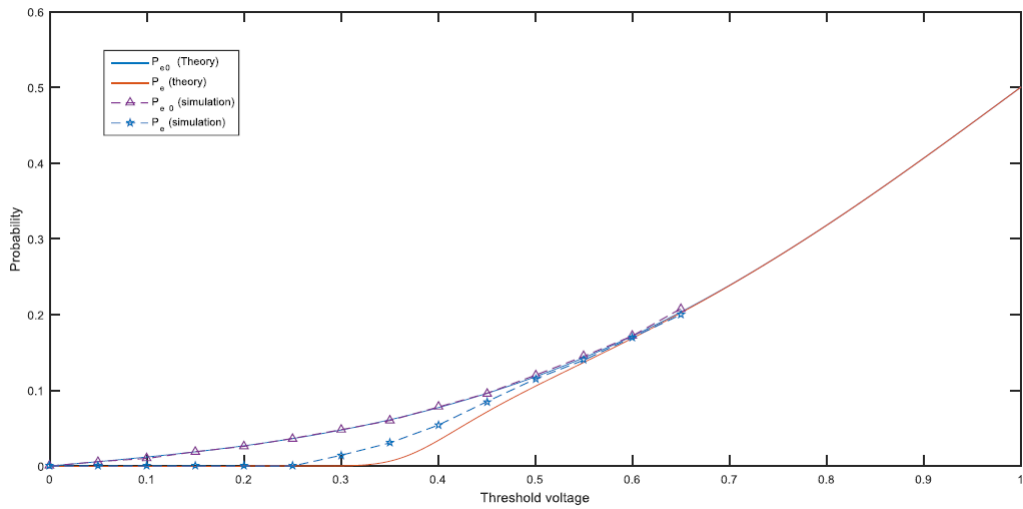
Fig. 4.50. Decoding failure rate for finite-length codewords.

The x-axis is threshold voltage and the y-axis is the probability of erasure. The purple curve is P_{e0} of the erasure probability before decoder and the dotted blue curve is the erasure probability after decoder. The red curve is the result from (4.7). It is evident that the fewer erasures can be corrected than the theoretical result. This is due to the cycle effect.

4.6 Shingled Magnetic Recording (SMR)

SMR is based on the concept of overlapping the previously written tracks and thereby estimating the recording head performance experimentally based on overlapping track pitch or the shingled track pitch using various write width heads.

4.6.1 Conventional Perpendicular Magnetic Recording

In Perpendicular magnetic recording (PMR), the recording bits are recorded to the medium in perpendicular to the plane unlike longitudinal recording where it used to be horizontal to the plane as shown in Fig. 4.51.
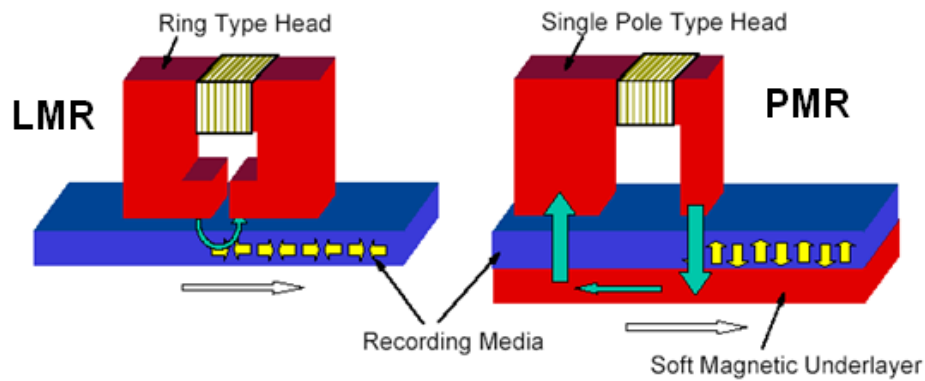
Fig 4.51. LMR Vs PMR basic head design.

For understanding the conventional PMR recording, a basic track profile is shown in Fig. 4.52 (a). The basic recording concept of the conventional recording on the medium is self-explained. They are simulations of three written tracks shown as Track N colored as pink, Track N+1, colored as yellow and Track N+2 as blue. Each written track is simulated with a track pitch of 4 µin. Each and every track has a clear separation. In this simulation, it is assumed that a writer width of 4 µin is used for recording every track sequentially in one direction. From Fig 4.52, the direction is from the left view to the right view of the picture. The y-axis denotes the amplitude level of the recording (also can be assumed as the circumferential direction of recording). The x-axis is assumed to be the radial direction. Fig. 4.52(b) shows the simple concept for the SWR. The rightmost track (3rd track – blue colored) is the full track, while the preceding left tracks (2 tracks – pink and yellow colored) are the shingled tracks. In the conventional recording, all the tracks will be separated when it is written sequentially based on the writer width and the specified product track pitch. In SWR, the concept is to overlap the previously written tracks and the data will be read based on the remaining track information. The shingled recording writes the information in the sequential manner overlapping the previously written track. This will be particularly helpful in the case of storing bulk information such as pictures or movies. For random storage, the disadvantage lies in the update-in-place issue which means the recorded data in a particular track could not be erased as it used to be in the conventional recording. To erase a single track, the entire block of data should be erased. Some system-level changes similar to flash storage technology will be required to overcome this.
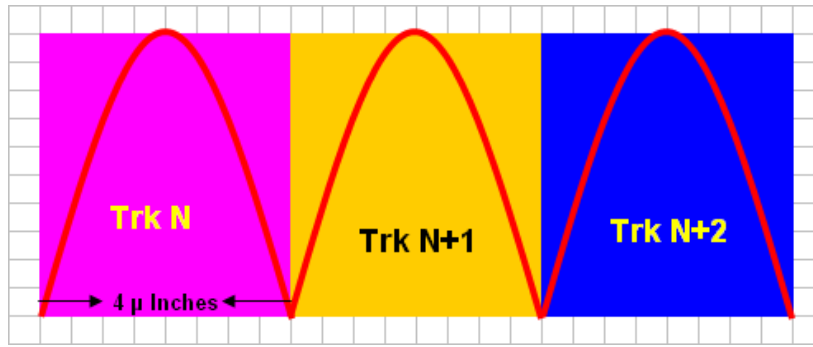
(a)
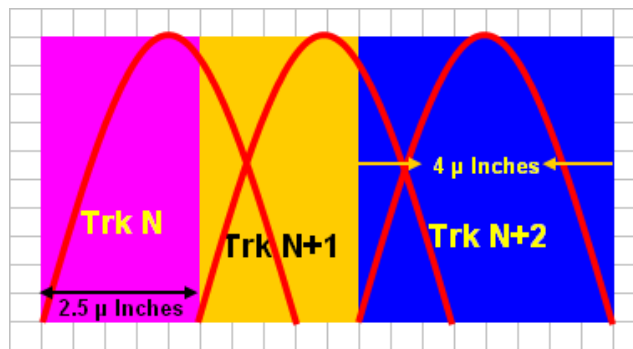
(b)

Fig 4.52. An understanding of (a) the conventional PMR and (b) SMR track profile.

4.6.2 Experiment structure and data analysis

The experiment for this thesis and the data analysis is performed at Western Digital Bangpa-in plant and KMITL. All experiments related to conventional recording and shingle recording is performed on commercially available Guzik HGA based dynamic electrical tester. The experiment structure and data analysis is as follows:

1. The pre selected HGA samples based on the conventional recording data will be subjected to the shingle testing with a fixed shingled track pitch of 2.5 μin. Key parameters such as ROW, SNR and BER are measured. The test results are compared to the conventional recording data.

2. The results from step 1 are analyzed. This result will help us in understanding the feasibility of the SWR. The details are covered in Chapter 4.

3. The same samples from step 1 are again subjected to test on multiple squeeze tests or multiple overlapping shingled TPI in the range of 1.5 μin to 3.5 μin. The same key parameters ROW, SNR and BER are measured. Also the AD is also calculated.

4. The results from step 3 are analyzed. This result will help us in determining the optimal track pitch and the optimal write width for the SWR. Also the AD could also be determined based on the optimal selection point.

5. Based on the different criteria of the selected writer width and the multiple shingled TP, the ratio will be obtained based on the write width and the squeeze/overlapping shingled TP (Shingled TP/MWW). The optimal shingled ratio will be determined based on the desired AD.

Based on the limitations of the conventional recording due to the so called super paramagnetic effects, there needs to be a new concept to increase the areal density. Several candidate technologies are being explored. One of the promising concepts without a major cost impact or major technological impact is SWR, which is being seriously considered as an alternative. This technology is perceived as an extension of the current perpendicular recording with probable minor changes compared with the conventional recording. In this section, we study and analyze the performance of the same selected samples that were studied for the conventional recording on the new SWR concepts. The same key parameters such as ROW, BER and SNR are measured using the Guzik spin stand testers. For shingled recording, a specifically designed test module solely owned by Western Digital is used to test the samples for shingled recording. The samples were tested with a fixed overlapping track pitch (squeeze track pitch) of 2.5 µin.
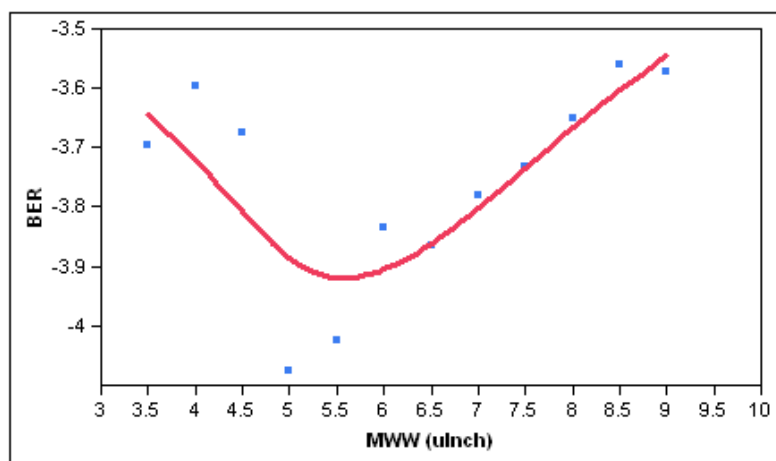

Fig 4.53. BER performance against MWW of shingled write recording.

Figure 4.53 shows the BER performance at various write width samples ranging from 3.5 µin to 9 µin, with the constant shingling track pitch of 2.5 µin. The BER varies

from -3.6 to -4 orders across multiple write width samples. However, the graph shows as a U curve with a function BER = a*MWW^2 + b*MWW + c (or others…). Based on the results, the optimal result is obtained with a write width of 5 to 6 µin, having the best case of BER around -4.1 orders with the 2.5 µin squeeze test. Also this graph suggests that for the write width samples smaller than 5 µin, the BER does not have a fixed pattern or a trend. This is due to the deficiency of the writability for the lower write width samples. For the write width samples greater or equal to 5 µin, we see the performance drops in a linear way, i.e., the higher the write width results to the lower performance, and the impact of drop between two different writer width levels does not have a significant difference and this difference can be attributed to the repeatability of the head, media and the system as theoretically the values are to be flat. The BER values that are obtained across multiple write width samples are acceptable for the recording system.
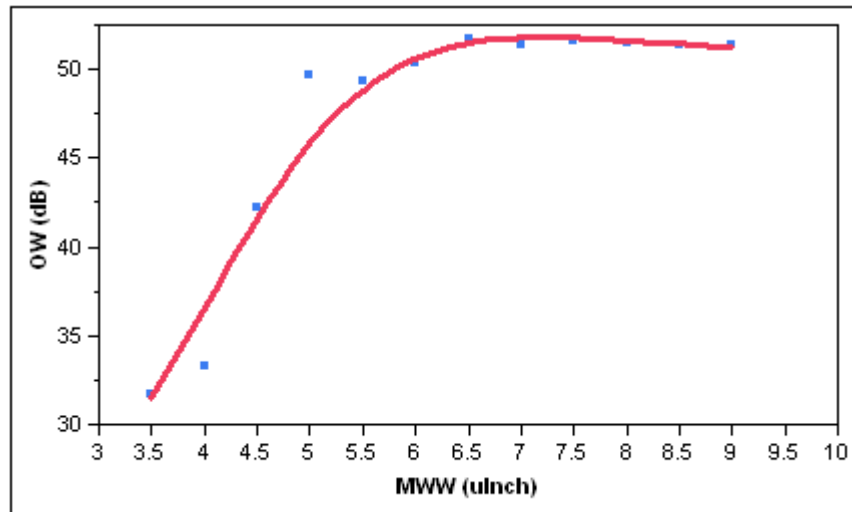
Fig 4.54. Reverse Overwrite (ReOW) performance against MWW on
shingled write recording.

The shingle write performance in terms of the ROW from the Guzik spinstand is shown in Fig 4.54. The track pitch of 2.5 µin was used to overlap the previous written track for shingling. The graph can be divided into two parts based on the write width values. The write width values of less than 5 µin as one part and above 5 µin as the other part. The lower part has the ROW values ranging around 30 to 40 dB which is due to the deficiency of the writability for the lower write width samples, and the other part has the flat ROW performance at around 50 dB. It means that for the write width values above 5 µin, the ROW remains constant and is similar to the theoretical predictions, in other words,

it is saturated to a value. This proves beneficial for the writability since we have seen that with a 2.5 TPI, we cannot achieve a good overwrite performance in the conventional recording, unlike the shingled recording. This is a good sign to overcome the writability deficiency when we need to increase the areal density because for the conventional recording, the only way to increase the areal density is to reduce the writer widths. Shingling can prove it right for increasing the areal density.
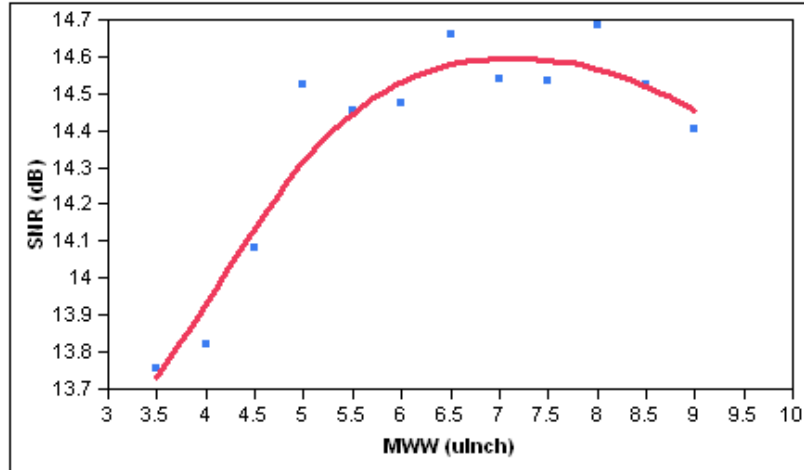
Fig 4.55. SNR performance against MWW on shingled write recording.

Figure 4.55 describes the Signal to Noise Ratio performance with a constant overlapping track pitch of 2.5 µin across various write width samples (3.5 µin to 9 µin). Similar to the ROW results shown previously, the SNR results are different for the two ranges of the MWWs. We can see that for the samples with the write width higher than 5 µin, the performance of the SNR seems to be almost flat averaging in the range of 14.5 dB (+/- 0.1 dB) which matches to the theoretical values which are supposed to be flat. These SNR values are higher than those found in the current recording systems and are well within the nominal distribution values. For the write widths below 5 µin, the SNR values drop significantly. For every decrement step of 0.5 µin, the SNR drops sharply by about 0.2 dB to 0.3 dB. This is due to the deficiency of the writability for the lower write width samples.

77

5. Conclusions

In this work, we propose various advanced signal detection and error correction code designs for data storage. For HAMR systems with large media noise levels, the target-shaping equalization is designed and shown that the $h_1 = 1$ constraint, monic constraint and the unit-energy constraint, give similar levels of mean-square errors. Although the unit-energy constraint achieves the lowest MSE value, the monic constraint gives the lowest bit error rates. The fixed-target of PR2 shows inferior performances to other constraints. For multi-tracking detection in 2-D interference channels, we proposed two methods to improve the performances of the graph-based detectors. The first method separates the 2-D interference channel into two 1-D targets: horizontal and vertical targets. The second method separates the 2-D interference channel into 2-D targets, with and without cornered ITIs. Both of the proposed methods reduce the number of length-four cycles to give the improved performances. The first method has reduced complexity. As the cornered ITI levels increase, the second method evidently outperforms the first method and the full graph detector.

For the algorithms related to LDPC codes, we improve the current schedulings in decoding by using the mixed scheduling which exploit the left-to-right (LR) and right-to-left reliabilities (RL). The SNR gain of MBP over the traditional layered BP (LBP) method is about 0.25 dB at BER = 10^{-6}. In addition, the QC-LDPC codes with maximized girth property are designed by choosing the best set of variable nodes in the progressive-edge growth algorithms. Hence, the maximized girth is obtained. The SNR gains are over 0.1 dB.

For flash channels, the simulation of multil-leveled cell (MLC) channels is generated. Soft decoding based on maximized mutual information (MMI) optimization is proposed to achieve the lowest bit error rate levels. The probability of erasure correction as various decoding iterations of LDPC decoder is shown. For ideal case of binary erasure channels, the performances will be better than more realistic channels. Although we show theoretical results of the approximate binary erasure channels, the derivation for more realistic channels is not made in this work.

Finally, the experimental results of shingled magnetic recording are obtained for over 1 Tbits/in^2. The shingle performance at various reverse overwrite (ReOW), magnetic

writer widths (MWW) are shown. The suitable MWWs are in the range of 6.5 to 7.5 uinch.

Future works in signal processing and code design for 2-D interference channels are promising as the recording densities are higher. The higher storage will force the storage industry to extend the HAMR system so that it is viable for mass production and the cost per GB is still competitive to flash alternatives. Non-binary LDPC codes and 2-D processing will be key technologies for the future of magnetic and flash storage.

6. Project Outputs

1. **ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ**

    1. S. Chandrasekaran, P. Supnithi, C. Warisarn and D.Bai, "Spinning Disk Test Study on Erase Band and Write Width for Shingled Magnetic Recording, " **Journal of Applied Physics**, vol.115, no.17, May, 2014. (2012 SJR impact factor = 2.21)

    2. T. Sopon, P. Supnithi and K. Vichienchom, "Improved 2-D Graph-Based Detectors for 2-D Interference Channels," **IEEE Transaction on Magnetics**, vol.50, no.11, Nov. 2014, article#3101704. (2013 JCR impact factor = 1.21).

    3. W. Phakphisut, P. Supnithi, and N. Puttarak, "EXIT Chart Analysis of Non-Binary Protograph LDPC Codes for Partial Response Channels," **IEEE Transaction on Magnetics**, vol.50, no.11, Nov. 2014, article#3101904. (2013 JCR impact factor = 1.21).

2. **การนำผลงานวิจัยไปใช้ประโยชน์**
    - เชิงสาธารณะ (มีเครือข่ายความร่วมมือ/สร้างกระแสความสนใจในวงกว้าง)
      ได้จัด Special Session ในหัวข้อ "Signal Processing and Communication for Data Storage" ในงานประชุมวิชาการ APSIPA 2014 ที่จะจัดระหว่างวันที่ 9 – 12 ธันวาคม พ.ศ. 2557 เมืองเสียมราฐ ประเทศกัมพูชา

    - เชิงวิชาการ (มีการพัฒนาการเรียนการสอน/สร้างนักวิจัยใหม่)
      นักศึกษาใน<u>ระดับปริญญาเอก</u> จบการศึกษา **จำนวน 3 คน**
        1. Dr.Selvan Chandrasekaran (ขณะนี้ทำงานอยู่ที่บริษัท Western Digital)
        2. ดร.ถนอมศักดิ์ โสภณ (ขณะนี้ทำงานอยู่ที่มหาวิทยาลัยราชมงคลอีสาน)
        3. ดร.เวธิต ภาคย์พิสุทธิ์ (ขณะนี้ทำงานอยู่ที่สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณ ทหารลาดกระบัง)

      ระดับ<u>ปริญญาโท</u> จบการศึกษา **จำนวน 1 คน**
        1. นางสาวอติญญา ศิริรุ่งสกุลวงศ์ (ขณะนี้ทำงานอยู่ที่บริษัท Seagate technology)

3. อื่นๆ (เช่น ผลงานตีพิมพ์ในวารสารวิชาการในประเทศ การเสนอผลงานในที่ประชุมวิชาการ หนังสือ การจดสิทธิบัตร)

เป็น Technical Program Committee (TPC) ในสาขา Data Storage การประชุมวิชาการ International Conference on Communication 2014 (ICC 2014)

การเสนอผลงานในที่ประชุมวิชาการ **จำนวน 3 ฉบับ**

1. A. Sirirungsakulwong, N. Puttarak and P. Supnithi, "Designed MMSE Equalizers for Nonlinear Magnetic Recording Channels," ECTI-CON 2014, May 14-17, 2013.
2. W. Phakphisut and P. Supnithi, " Decoding Algorithm of LDPC codes for Cascaded BSC-AWGN channels," APSIPA ASC 2014, December 2-5, 2014.
3. S. Khittiwitchayakul, W. Phakphisut, P. Supnithi, "Weighted bit-flipping decoding for product LDPC coders," ECTI-CON 2016, Chiangmai, Thailand.

References

[1] R. Rottmeyer *et al.*, "Heat-assisted magnetic recording," IEEE Trans.Magn., vol. 42, no. 10, pp. 2417–2421, Oct. 2006.

[2] M.A Seigler *et al*, "Integrated Heat Assisted Magnetic Recording Head: Design and Recording Demonstration," IEEE Trans. Magn., vol. 44, no. 1, pp. 119 – 124, Jan. 2008.

[3] B. Terris, T. Thomson, and G. Hu, "Patterned media for future magnetic data storage," *Microsyst. Technol.*, vol. 13, no. 2, pp. 189–196, Nov.2006.

[4] G. Hughes, "Read channel for patterned media," IEEE Trans. Magn., vol. 35, no. 5, pp. 2310-2312, Sep. 1999.

[5] P. W. Nutter, I. T. Ntokas, and B. K. Middleton, "An investigation of the effects of media characteristics on read channel performance for patterned media storage," IEEE Trans. Magn., vol. 41, no. 11, pp. 4327-4334, Nov. 2005.

[6] R. Wood, et al, "The feasibility of magnetic recording at 10 terabits per square inch on conventional media, IEEE Trans. Magn., vol. 45, no. 2, pp. 917-923, Feb. 2009.

[7] P. Kasiraj and M. Williams, "System and Method for Writing Hard Disk Drive Depending on Head Skew," U.S. Patent 6 967 810 B2, Nov 22, 2005.

[8] S. Nabavi, V. Kumar, and J. G. Zhu, "Modifying Viterbi algorithm to mitigate intertrack Interference in bit-patterned media," IEEE Trans. Magn., vol. 43, no. 6, pp. 2274-2276, Jun. 2007.

[9] M. Keskinoz, "Two-dimensional equalization/detection for patterned media storage," IEEE Trans. Magn., vol. 44, no. 4, pp. 533-539, Apr. 2007.

[10] S. Nabavi and V. Kumar, "Two-dimensional generalized partial - response equalizer for bit patterned media," IEEE Int. Conf. Communication (ICC), 2007, pp. 6249-6254.

[11] S. Karakulak, P. H. Siegel, J. K. Wolf and H. N. Bertram, "A new read channel model for patterned media storage," IEEE Trans. Magn., vol. 44, no. 1, pp. 193-197, Jan. 2008.

[12] S. Karakulak, P. H. Siegel, J. K. Wolf and H. N. Bertram, "Equalization and detection for patterned media recording," Intermag 2008, HT10.

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

[14] E. Eleftheriou and W. Hirt, "Noise predictive maximum-likelihood (NPML) detection for the magnetic recording channel," in *Proc. IEEE ICC '96*, June 1996, pp. 556–560.

[15] Bahl, Cocke, Jelinek, Raviv, "Optimal decoding of linear codes for minimal symbol error rate", *IEEE Trans. On Inform. Theory*, vol. IT-20, pp.284-287, March 1974 .

[16] J. Hagenauer, "A Viterbi Algorithm with Soft-Decision Output and its Applications", *Proc. GLOBECOM'89*, Dallas, TX, Nov. 1989, pp.47.1.1-47.1.7.

[17] L.M.M. Myint, P. Supnithi and P. Tantasawasd, "An Inter-Track Interference Mitigiation Technique Using partial ITI Estimation in Patterned Media Storage, " To appear in IEEE Trans. on Magn., vol. 5, no. 10, October 2009.

[18] J. Hu, T.M. Duman and M. F. Erden, "Graph-based channel detection for multitrack recording channels, *EURASIP Journal on Advances in Signal Processing*, Vol. 2008, ID 738281, 2008.

[19] O. Shental, A. J. Weiss, N. Shental, and Y. Weiss, "Generalized belief propagation receiver for near-optimal detection of two-dimensional channels with memory," in *Proceedings of IEEE Information Theory Workshop (ITW '04)*, pp. 225–229, San Antonio, Tex, USA, October 2004.

[20] B. M. Kurkoski, P. H. Siegel, and J. K. Wolf, "Joint message-passing decoding of LDPC codes and partial-response channels," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1410–1422, 2002.

[21] G. Colavolpe and G. Germi, "On the application of factor graphs and the sum-product algorithm to ISI channels," *IEEE Transactions on Communications*, vol. 53, no. 5, pp. 818–825, 2005.

[22] R. G. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. IT-18, pp. 21-28, Jan. 1962.

[23] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electronics Letters*, vol. 32, No. 18, pp. 1645-1646, 1996.

[24] T. J. Richardson, M. A. Shokrollahi, and R. Urbanke,"Design of capacity-approaching irregular low-density parity-check codes," *IEEE Transactions Information Theory*, vol. 47, pp. 619-637, Feb. 2001.

[25] X.-Y. Hu, E. Eleftheriou, and D.-M. Arnold, "Regular and irregular progressive edge-growth tanner graphs," *IEEE Trans. Inform. Theory,* vol. 51, pp. 386-98, 2005.

[26] S. Lin, L. Chen, J. Xu and I. Djurdjevic, "Near Shannon limit quasi-cyclic low-density parity-check codes," in *Proc. IEEE Global Communications Conference*, Vol. 4, pp. 2030 – 2035, Dec. 2003.

[27] Z. Li and B. V. K. V. Kumar, "A class of good quasi-cyclic low-density parity check codes based on progressive edge growth graph," in *Proc. 38th Asilomar Conf. Signals, Systems and Computers*, Monterey, CA, Nov. 7–10, 2004, vol. 2, pp. 1990–1994.

[28] H. Xiao and A. H. Banihashemi, "Improved progressive-edge-growth (PEG) construction of irregular LDPC codes," *IEEE Communication Letters*, vol. 8, no. 12, pp. 715–717, Dec. 2004.

[29] Z. Zhou, X.Li, D.Zheng, K.Chen and J. Li, "Entended PEG Algorithm for high rate LDPC codes," in *Proc. IEEE International Symposium on Parallel and Distributed Processing with Applications*, pp. 494-498, 2009.

[30] "IEEE standard for information technology--telecommunications and information exchange between systems--local and metropolitan area networks--specific requirements part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 5: enhancements for higher throughput," *IEEE Stud 802.11n-2009*, pp. c1-502, 2009.

[31] R.L. Galbraith*, et al.*, "Architecture and implementation of a first-generation iterative detection read channel," *IEEE Trans. Magn.,* vol. 46, pp. 837-843, 2010.

[32] D.E. Hocevar, "A reduced complexity decoder architecture via layered decoding of LDPC codes," in *Proc. IEEE Workshop Signal Process. Syst. (SIPS)*, pp. 107-12, 2004.

[33] E. Yeo*, et al.*, "High throughput low-density parity-check decoder architectures," in *Proc. IEEE GLOBECOM,* vol.5, pp. 3019-24, 2001.

[34] M.M. Mansour and N. R. Shanbhag, "High-throughput LDPC decoders," *IEEE Trans. Very Large Scale Integ. Syst.,* pp. 976-96, 2003.

[35] H. Kfir and I. Kanter, "Parallel versus sequential updating for belief propagation decoding," *Physica A,* vol. 330, pp. 259-70, 2003.

[36] Z. Juntan and M. Fossorier, "Shuffled belief propagation decoding," in *Proc. 36th Asilomar Conf. Signals, Syst. and Comput.*, pp. 8-15, 2002.

[37] S. Wu, X. Jiang and Z. NIE, "Alternate iteration of shuffled belief propagation decoding," in *Proc. International Conference on Communications and Mobile Computing.*, pp. 278-281, 2010.

[38] G. Elidan, I. McGraw, and D. Koller, "Residual belief propagation: informed scheduling for asynchronous message passing," in *Proc. Uncertainty in Artificial Intelligence*, 2006.

[39] A.I. Vila Casado, M. Grit and R.D. Wesal, "LDPC decoders with informed dynamic scheduling," *IEEE Trans. Comm.*, vol. 58, pp. 3470-79, 2010.

[40] K. Jung-Hyun, *et al.*, "Variable-to-check residual belief propagation for informed dynamic scheduling of LDPC codes," in *Proc. IEEE Int. Symposium on Information Theory and Its Applications*, pp. 1-4, 2008.

[41] G.-j. Han and X.-c. Liu, "Dynamic schedules based on variable nodes residual for LDPC codes," in *Proc. Wireless Communications, Networking and Mobile Computing*, pp. 1-3, 2009.

[42] W. Phakphisut, *et al.*, "Serial belief propagation for the high-rate LDPC decoders and performances in the bit patterned media systems with media noise" *IEEE Trans. Magn.*, vol. 47, pp. 3562-65, 2011.

[43] H. Muraoka and S.J. Greaves, "Statistical modeling of write error rates in bit patterned media for 10 Tb/in2 recording," *IEEE Trans. Magn.*, vol. 47, no. 1, pp. 26-34, Jan. 2011.

[44] J. Hu, *et al.*, "Bit-patterned media with written-in errors: modeling, detection, and theoretical limits," *IEEE Trans. Magn.*, vol. 43, pp. 3517-24, 2007

[45] A.R. Iyengar, P.H. Siegel and J.K. Wolf, "LDPC codes for the cascaded BSC-BAWGN channel," in *Proc. Allerton Conf. on Comm., Control and Computing*, pp. 620-7, 2009.

[46] A. Risso, "Layered LDPC decoding over GF(q) for magnetic recording channel" *IEEE Trans. Magn.*, vol. 45, pp. 3683-6, 2009.

[47] R. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory*, vol. 27, pp. 533-47, 1981.H. Xiao-Yu, *et al.*, "Efficient implementations of the sum-product algorithm for decoding LDPC codes," in *Proc. IEEE GLOBECOM*, vol.2, pp. 1036-1036, 2001.

[48] X.Y. Hu, E. Eleftheriou and D.M. Arnold, "Regular and irregular progressive edge-growth Tanner graphs," *IEEE Trans. Inform. Theory*, vol. 51, pp. 386-98, 2005.

[49] G. Dong, N. Xie, T. Zhang, "On the use of soft-decision error-correction codes in NAND flash memory," IEEE Trans. On Circuits and Systems, vol. 58,, No. 2, Fe. 2011, pp. 429 – 439.

[50] C. Yang et al., "Flexible product code-based ECC schemes for MLC NAND Flash Memories," SiPS 2011, Oct. 4-7, 2011, Beirut, Lebanon.

[51] Y. Cai et al., "Error patterns in MLC NAND flash memory: measurement, characterization and analysis.

[52] R. E. Blahut, Algebraic Codes for Data Transmission. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[53] G. Dong, N. Xie, and T. Zhang, "On the use of soft-decision error-correction codes in NAND flash memory," *IEEE Trans. Circuits Syst. Iregul. Papers,* vol. 58, no. 2, pp.429-439, Feb. 2011.

[54] J. Wang, T. Courtade, H. Shankar, and R.D. Wesel, "Soft Information for LDPC Decoding in Flash: Mutual-Information Optimized Quantization", publication in the IEEE Globecom 2011 proceedings.

[55] C.A. Aslam, Yong Liang Guan, and Kui Cai "Read and write voltage signal optimization for multi-level-cell (MLC) NAND flash Memory" *IEEE TRANSACTIONS ON COMMUNICATIONS,* VOL. 64, NO. 4, APRIL 2016.