Simulation of Heat Assisted Magnetic Recording System

A THESIS
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Yipeng Jiao

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Randall H. Victora

November 2018

Acknowledgements

I am really lucky to have a great advisor, professor Randall Victora, who introduced me to the field of magnetic recording and micromagnetic simulation, and grateful for him encouraging me to conquer every challenge in my PhD study, especially during my difficult times. I would like to thank him for his valuable comments and helpful suggestions on both of my research and life in the last five years.

I would also like to thank other professors in my PhD final examination committee, Professor Jian-Ping Wang, Professor Steven Koester and Professor Xiaojia Wang, for their time, help and efforts of serving in my committee.

My managers, Chris Rea and Chindu Satheesh, gave me a lot of help and trust during my internships at Seagate and MathWorks, respectively. Thank you!

It has been a memorable experience with members of Professor Victora's group and other students in MINT center. I would like to express my special thanks to Pin-Wei Huang, Yao Wang, Sumei Wang, Tao Qu, Ali Ghoreyshi, Zengyuan Liu, Wei-Heng Hsu, Ziran Wang, Rizvi Ahmed, Aneesh Venugopal, Niranjan Natekar, Kai Wu, Zhengyang Zhao, Hongshi Li, Yi Wang, Jinming Liu, Yinglong Feng, Xiaohui Chao, Eliot Estrine, Xiaowei Zhang, Todd Klein and many other friends for their help.

Finally, I would like to express my love and thanks to my family——it is they who make me stand where I am today.

Dedication

With gratitude to my family

Abstract

Heat-assisted magnetic recording (HAMR) is a promising approach to increase the areal density of hard disk drives (HDDs). The first part (Chapter 2~3) of this dissertation provides a computational framework/program for HAMR simulation and a system-level optimization is then discussed in the second part (Chapter 4~6).

Anisotropic exchange has been incorporated in a description of magnetic recording media near the Curie temperature, as would be found during HAMR. The new parameters were found using a cost function that minimized the difference between atomistic properties and those of renormalized spin blocks. Interestingly, the anisotropic exchange description at 1.5 nm discretization yields very similar switching and magnetization behavior to that found at 1.2 nm (and below) discretization for the previous isotropic exchange. This suggests that the increased accuracy of anisotropic exchange may also reduce the computational cost during simulation.

We then studied several factors that could affect the playback signal-to-noise ratio (SNR), such as the switching time of the writing field, the peak temperature of the heat spot, the standard deviation of anisotropy field and exchange parameters, and the size of the reader element. We found that bit length (BL) becomes the key factor for the value of SNR for small, e.g., 10 nm, BLs. We established a relationship between the SNR and the bit error rate (BER) so that it will be easy to predict the storage capacity based on the

playback SNR. The capacity prediction is performed with current technologies as well as those expected to be available in the near future, since the ultimate user storage-area density for heat-assisted magnetic recording is an important consideration both from research and commercial perspectives. Micromagnetic shows a result of 1.8 Tb/in^2. Optimization of track pitch in the absence of track misregistration increases the density to 5 $TBit/in^2$ for perfect recording.

Finally, transition noise and remanence noise are the two most important types of media noise in HAMR. We examine two methods (spatial splitting and principal components analysis) to distinguish them: both techniques show similar trends with respect to applied field and grain pitch (GP). It was also found that PW_{50} can be affected by GP and reader design but is almost independent of write field and bit length (larger than 50 nm). Interestingly, our simulation shows a linear relationship between jitter and $PW_{50}NSR_{rem}$, which agrees qualitatively with experimental results.

List of Contents

Acknowledgements i
Dedication ii
Abstract iii
List of Contents v
List of Figures vii
List of Tables xi
List of Abbreviations xii
1 Introduction 1
1.1 A Brief Introduction to Magnetic Storage Technology 1
1.2 Structure of Modern HDD 3
1.2.1 Recording Media 5
1.2.2 Write Head 6
1.2.3 Read Head 8
1.3 Challenges for HDD above 1Tb/in^2 10
1.4 "The Next Leap Forward is Now" 12
1.4.1 Optical system 14
1.4.2 Recording Media (HAMR) 17
1.4.3 Recording Physics and Noise 20
1.5 Summery and Outline of this Work 25
2 Computational and Modeling Methods 27
2.1 Simulation program overview 27
2.2 Media generation 28
2.3 Recording Simulation: Micromagnetics and Renormalized LLG model 35
2.3.1 Landau-Lifshitz-Gilbert (LLG) equation 35
2.3.2 Renormalized LLG model 42
2.3.3 Recording Simulation Program 45
2.4 Playback and State Detection 47
3 Renormalized Anisotropic Exchange for Representing Heat Assisted Magnetic Recording Media 54
3.1 Introduction 54
3.2 Method and Model 55

3.3 Results and Discussions 61
3.4 Conclusion 64
4 A Study of SNR and BER in Heat Assisted Magnetic Recording 67
4.1 Introduction 67
4.2 Simulation Models and Methods 68
4.2.1 Recording Media and Writing Head 68
4.2.2 Playback and Detection System Architecture 69
4.3 Results and Discussion 71
4.3.1 Noise 71
4.3.2 Switching Time 72
4.3.3 Peak Temperature of the Heat Spot 73
4.3.4 Standard Deviation of Anisotropy Field and Exchange 73
4.3.5 Reader Width 76
4.4 Conclusion 78
5 Dependence of Predicted HAMR Areal Density on Common Optimization Strategies 80
5.1 Introduction 80
5.2 HAMR Optimization with Micromagnetic Simulation 80
5.2.1 Optimization of NFT and reader size 81
5.2.2 Effects of Grain Size 82
5.2.3 Distance between NFT and Writing Pole 84
5.2.4 Write Field Compensation and Pulse Laser Recording 84
5.3 HAMR User Areal Density Prediction 87
5.4 Conclusion 90
6 Understanding Transition and Remanence Noise in HAMR 92
6.1 Introduction 92
6.2 Simulation Parameters and Method 92
6.3 Transition Noise and Remanence Noise 95
6.4 Relationship between SNR and Jitter 102
6.5 Conclusion 104
7 Summary 106
References 108

List of Figures

Figure 1.1: Areal density of hard disk drive (HDD) since 1956 [1]. 1
Figure 1.2: The structure of IBM 3340 [18] (left); the structure of today's HDD with labels of the main component [19] (right). 3
Figure 1.3: Magnetic image (by MFM) of recorded patterns and tracks on a disk [21]..... 5
Figure 1.4:Longitudinal magnetic recording (left) and perpendicular magnetic recording (right). [32]..... 8
Figure 1.5: Illustration of CPP GMR structure [42]. The parallel state is on the left and the anti-parallel state is on the right with equivalent electric circuit representations, respectively. 10
Figure 1.6: Trilemma among media, writer and reader in HDD. 12
Figure 1.7: A 3d model of HAMR recording system, only showing the slider and medium [46]. 14
Figure 1.8: Schematic of a HAMR head (left); Common blocks of the light delivery system in a HAMR head (right). [54] 15
Figure 1.9: A planar solid immersion mirror with a dual offset grating used to focus a waveguide mode onto the lollipop NFT [51]..... 16
Figure 1.10: HAMR media structure [46]..... 17
Figure 1.11: Phase diagram of Fe-Pt binary alloy (left) and Lattice structure of A1 phase FePt and $L1_0$ phase FePt (right). [66] 20
Figure 1.12: Illustration of the working principle of HAMR [46]..... 21
Figure 1.13: The medium SNR as a function of normalized standard deviation for H_k or T_c variation with different thermal gradients [69]. 22
Figure 1.14: The dependence of curie temperature on grain size [67]. 23
Figure 1.15: Absorption power inside (a) granular and (b) effective medium approximation of HAMR media. [68] 24
Figure 1.16: Recording magnetization patterns w/o T_c variations when H = 10 kOe. Some grains aren't switched successfully as expected because of thermal fluctuations. ... 25
Figure 1.17: Track width normalized to that at 300 KFCI for a PMR head (green line), and two different HAMR designs. Inset: the measured curvature for all three designs, showing higher curvature for the HAMR heads, with HAMR B (red line) being worse than HAMR A. [74]..... 26
Figure 2.1: Working principles of an HDD interpreted with the "write/read" channel model..... 28
Figure 2.2: Basic modules in our HAMR simulation program..... 28
Figure 2.3: Bowyer–Watson algorithm: (a) a new point P is added into Delaunay mesh; (b) P is in the two circumcircles of ⊙ ABD and ⊙ ABC; (c) △ ABD and △ ABC are broken; (d) regenerate Delaunay tringles..... 31
Figure 2.4: Relationship between Delaunay triangles (black) and Voronoi cells (red): Voronoi seeds is the vertices of Delaunay triangles (left) while circumcenters of Delaunay triangles is the vertices of Voronoi polygons (right). [82] 32

Figure 2.5: A medium sample using Voronoi cells with grain boundaries. The parts in green and yellow color are repetitions of the leftmost part in white. Therefore, the edges on both sides must match with each other. The units for both axes are angstroms. 33
Figure 2.6: A short edge may disappear after the polygon is shrunk. The original polygon is filled with dark blue and the shrunk one is filled with light blue..... 33
Figure 2.7: A case where several edges or their extension lines cross with the edge α at A, B, C and D after shrinking, whose directions are also marked. In order to pick up the two points after offsetting, we divided these crossing points into two groups, based on their cross products, i.e. $\mathrm{CB} \times \mathrm{FB}$, $\mathrm{CB} \times \mathrm{FA}$ fall into one group and $\mathrm{CB} \times \mathrm{EC}$, $\mathrm{CB} \times \mathrm{ED}$ make another one. Indeed, the first group means the edges will "leave" from edge α while the second means they will "arriving" at edge α. Then we can pick up the point in group one which has the minimum distance to any points in group two and vice versa. 34
Figure 2.8: The atomistic spins (a) are divided into renormalized spin block (b). Granular medium which is represented by Voronoi diagram are discretized using this spin block (c). [89]..... 44
Figure 2.9: Renormalized Parameters as functions of temperature and length scale. (a) Gilbert damping parameter; (b) Saturation magnetization (c) Anisotropy constant. (d) Exchange stiffness parameter. d represents different length scales; "agg" is equal to the number of atomic spins in a renormalize spin. Dash lines indicate the corresponding values in atomistic (single spin) model, which are constants. [89]... 44
Figure 2.10: The workflow of the recording simulation module. All blocks/functions in the red box are executed on GPGPU which are also called kernel functions. And four of them in the blue box, marked with dotted borders and no arrows, can also be executed at the same time on different devices (GPGPUs or CPUs). 46
Figure 2.11: A general memory model of GPGPU. The access time for shared memory is 20 ~ 40 Cycles while for global memory it is 400 ~ 600 Cycles. The capacity of shared memory is usually tens of Kilobyte while it is several Gigabytes for global memory. [94]..... 47
Figure 2.12: Layout of MR read head (a) and a sample the read field profile (b). 48
Figure 2.13: The Viterbi decoder used in our simulation, with four states and a path length of 4. Discarded states are shown in grey..... 52
Figure 2.14: Outputs from each step. (a) recorded patterns on a Voronoi medium; (b) original playback signals with a reader; (c) sampled playback signals; (d) signals after equalization; (f) binary states -1,1 after decision, 0 is reserved for invalid value here..... 53
Figure 3.1: Panel (a) shows the atomic locations in the FePt $L1_0$ that are replaced by renormalized blocks (b) in the simulation. 55
Figure 3.2: Renormalized parameters as functions of temperature. Circles are the results using this new method, squares are the result from [89], and solid lines show the fitting functions. 61
Figure 3.3: Switching probability for different discretizations and exchange, solid symbols are the results with a field of 5 kOe and open ones are 10 kOe. 64

Figure 3.4: Average M_z (magnetization along z direction) variation with time under a given heat pulse. In 4(a) the AEP is considered with block size of $1.5\times1.5\times1.5$ nm^3, and in (b) only isotropic exchange parameters are considered with block size of $1.0\times1.0\times1.0$ nm^3. In (c) the different models are compared for an applied field H = 5 kOe and θ_H (the angle between the z axis and H) of 0°to show the differences in M_z during the switching process. 66
Figure 4.1: A recorded medium sample with PRBS pattern (11100010110100111101011 10 68
Figure 4.2: System diagram of our simulation program. 69
Figure 4.3: Readback signals (top) and noise (bottom) for one case with 8 runs. H = 8kOe, T_{max} = 850K, and PRBS is used. 72
Figure 4.4: SNR curves against the switching time t_p of the writing field (a). The definition of t_p is illustrated in (b). H = 8kOe, BL = 10.5nm, and T_{max} = 850. 73
Figure 4.5: SNR curves against the maximum temperature of the heat spot T_{max}. H = 8kOe, switching time of field t_p = 0ps, and BL = 10.5nm. 74
Figure 4.6: SNR curves against the standard deviation of anisotropy field H_k (a) and average exchange parameter A_{ex} (b). H = 8kOe, θ_H = 0, switching time of field 0ps, BL = 10.5nm, and T_{max} = 850K. 75
Figure 4.7: SNR curves obtained with different width of readers. H = 8kOe, switching time of field 0ps, BL = 10.5nm, and T_{max} = 850K. 76
Figure 4.8: The relationship between SNR and BER. The black line is the fitting function. 78
Figure 5.1: Media SNR with different reader width. 30 nm FWHM gives the best SNR performance. 82
Figure 5.2: Media SNR with different BL, GP and its standard deviation. FWHM = 50nm. 83
Figure 5.3: SNR vs distance between write head and NFT with different thermal spot size. 84
Figure 5.4: (a) Write field pulse including overshoot. (b) Media SNR performance with different overshoot fields. t_D = 1ns (BL = 10.5 nm), H_W = 10 kOe. "Ideal" means t_s = 0, H_{OA} = 0, t_{OD} = 0; case 1: t_s = 0.3 ns, H_{OA} = 0, t_{OD} = 0; case 2: t_s = 0.3 ns, H_{OA} = 2 kOe, t_{OD} = 0.2 ns; case 3: t_s = 0.3 ns, H_{OA} = 2 kOe, t_{OD} = 0.4 ns; case 4: t_s = 0.3 ns, H_{OA} = 5 kOe, t_{OD} = 0.2 ns; case 5: t_s = 0.3 ns, H_{OA} = 0, t_{OD} = 0, H_W = 15 kOe. 86
Figure 5.5: Media SNR performance w/ and w/o pulsed laser. The inset shows the wave shape of a pulsed laser, which was calculated by [119] (top) and its corresponding temperature changes on the media (bottom). 87
Figure 5.6: Media SNR (with solid symbols) and UAD (with open symbols) predication with different BL and FWHM with perfect recording simulation. 89
Figure 5.7: UAD with squeezed track, prediction using perfect recording simulation. ... 90
Figure 6.1: (a) Averaged playback signal in one period (dibit) with reader 2 when H = 5kOe, GP = 5.5nm. (b) The derivative of (a). PW_{50} is also shown. 96
Figure 6.2: (a) PW_{50} vs writing field. Parameters of Reader 1: SSS (shield-to-shield spacing) = 11nm, reader width = 42nm and fly height = 6nm; Parameters of Reader 2: SSS = 22nm, reader width = 36nm and fly height = 10nm. (b) Calculated SNR

using the method in [8] with BL = 50nm. Open symbols are the results at GP = 5.5nm, solid ones are at GP = 6.5nm and half open ones are at GP = 8.5nm. 97
Figure 6.3: Reconstructed playback signals with different number of PCs. (a) R = 1; (b) R = 2; (c) R = 3; (d) R = 10. 100
Figure 6.4: Different PCs of Figure 6.3 in one period. The largest blue curve is the first PC (r = 1). (a) r = 1, 2, 3, 4; (b) r = 1, 6, 7, 8. 101
Figure 6.5: Calculated SNR using PCA with BL = 25nm. Open symbols are the results at GP = 5.5nm, solid ones are at GP = 6.5nm and half open ones are at GP = 8.5nm. 102
Figure 6.6: Relationship between remanence jitter (or jitter' in method (1)) and the product of PW_{50} and remanence NSR using the (a) first and (b) second method, respectively. Dash lines are the fitting functions. 105

List of Tables

Table 1 High K_u magnetic materials [59] 17
Table 2 A summary of FePt granular media fabrication..... 20
Table 3 Temperature categories and corresponding fitting rules of renormalized material parameters 59
Table 4 Calculated results with equation (4-3) and (4-4) 79
Table 5 Important parameters in micromagnetic simulation 81
Table 6 Storage capacity prediction with micromagnetic simulation..... 88
Table 7 Important parameters in micro-magnetic simulation..... 94

List of Abbreviations

AAMR: acoustic-assisted magnetic recording

AEP: anisotropic exchange stiffness parameter

BER: bit error ratio

BL: bit length

ECC: exchanged coupled composite

FIFO: first-in-first-out

FWHM: full width at half-maximum

GP: grain pitch

GPGPU: general-purpose graphics processing unit

HAMR: heat-assisted magnetic recording

HDD: hard disk drive

HMM: hidden Markov model

ISI: inter-symbol interference

ITI: inter-track interference

LLB: Landau–Lifshitz–Bloch

LLG: Landau–Lifshitz–Gilbert

LMR: longitudinal magnetic recording

MAMR: microwave-assisted magnetic recording

MLE: maximum likelihood estimation

MLSD: maximum-likelihood sequence detection

MMSE: minimum mean-square error

MR: magnetoresistance

NFT: near field transducer

NSR_{rem}: remanence noise-to-signal ratio

PC: principal component

PCA: principle components analysis

PCRAM: phase-change random-access memory

PMR: perpendicular magnetic recording

PR: partial-response

PR4: PR Class IV

PRBS: pseudo-random binary sequence

PRML: partial-response maximum likelihood

RRAM: resistive random-access memory

SFD: switching field distribution

SMR: shingled magnetic recording

SNR: signal to noise ratio

SP: surface plasmon

SPD: switching probability distribution

SSD: solid state drive

SSS: shield to shield space

SUL: soft underlayer

SVD: singular value decomposition

UAD: user areal density

WAS: wrap around shield

1 Introduction

1.1 A Brief Introduction to Magnetic Storage Technology

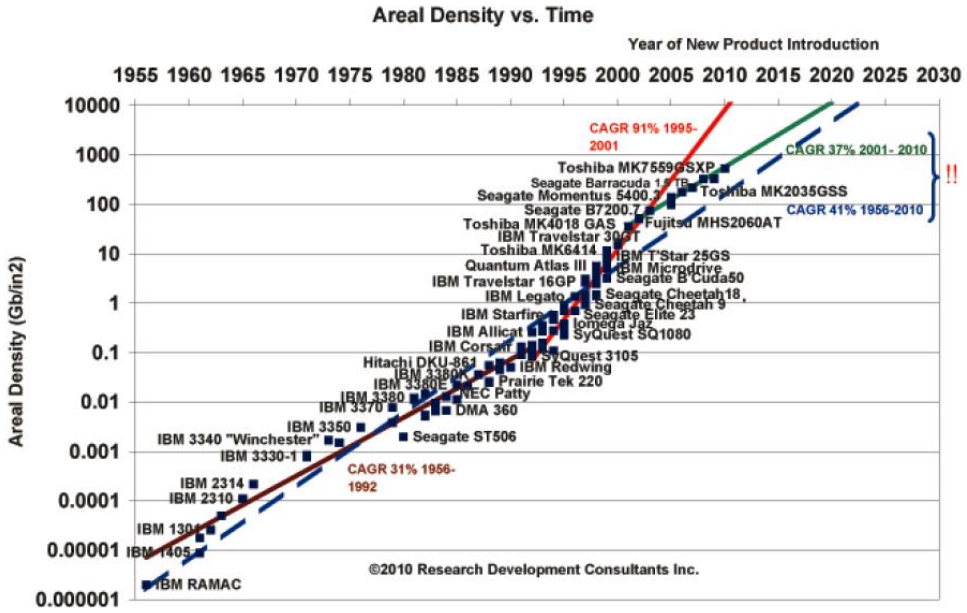


Figure 1.1: Areal density of hard disk drive (HDD) since 1956 [1].

The first product in the magnetic recording history is the magnetic wire recorder, developed by Valdemar Poulsen, a Danish engineer, in 1898 [2]. During the following 120 years (1898 - 2018), people made revolutionary progress and many breakthroughs in the data storage field based on magnetic recording technology, such as magnetic tape (1928) [3], drum memory (1932) [4], etc. Among all of them, Hard-disk drive (HDD) is the most influential invention. The first commercial hard disk drive was built by IBM in 1956, which was called the random-access method of accounting and control (RAMAC) [5] at that time.

Its capacity was less than 5MB, but its size was almost equal to two medium-sized refrigerators, and its areal density was only 2KB/in^2 [6]. After that, the areal density of HDDs increased exponentially and by 2015, this value has reached about 1TB/in^2, far greater than the scaling rate in semiconductor industry predicted by Moore's Law [7]. Figure 1.1 shows the areal density of HDD increases as time changes, with important products also marked [1].

Magnetic recording is still the dominant technology for data storage today, though solid-state drive (SSD) [8] is becoming a big threat, especially in mobile devices. The advantage of SSD is its high I/O speed and low energy cost. However, because one single storage unit of SSD is a transistor-based device with complex structure and operation method, it is very difficult to reduce its fabrication cost, especially as silicon technology is approaching its physical limit (around 5 ~ 7nm) currently. Another inherent problem caused by the complex device structure and working principle is the relative bad data durability and recoverability [9]. Although people are trying to develop novel devices to replace this silicon-based device, such as resistive random-access memory (RRAM) [10], phase-change random-access memory (PCRAM) [11], [12] etc., or using 3D structures [13], they cannot solve these two problems essentially. In the opposite way, the data in HDD is stored among the grains of the media, which have very simple structure and are also easy to fabricate. Most importantly, this data does not disappear after millions of read and write, even if the head (slider), which is the most complex element in an HDD, breaks. As long as we have a new head or other professional equipment, we can recover this data from the media directly and this is actually what data recovery companies do [14]. Much more

amazing is that the other data could still be recovered even though a part of the medium is broken. In addition, magnetic recording industry is making progress continuously these years: in 2013, the Western Digital subsidiary HGST introduced the world's first commercial helium-filled hard drive in 2013 [15] which reduced the power consumption and increased the number of disks packed in one HDD significantly. In the following year (2014), Seagate built the world's first Shingled Magnetic Recording (SMR) hard drive [16], which improves areal densities by 25%. Seagate also announced its multi-actuator technology last year [17] and its first-generation product, equipped with dual actuators (two actuators), can provide a 480MB/s reading speed, which is 60% faster than the normal speed of HDD. Therefore, with these new techniques, the I/O speed and areal density of HDD are expected to be improved even more in future and magnetic recording will have a very good prospect for development. Until today, it is still the irreplaceable technology for data storage in the world.

1.2 Structure of Modern HDD

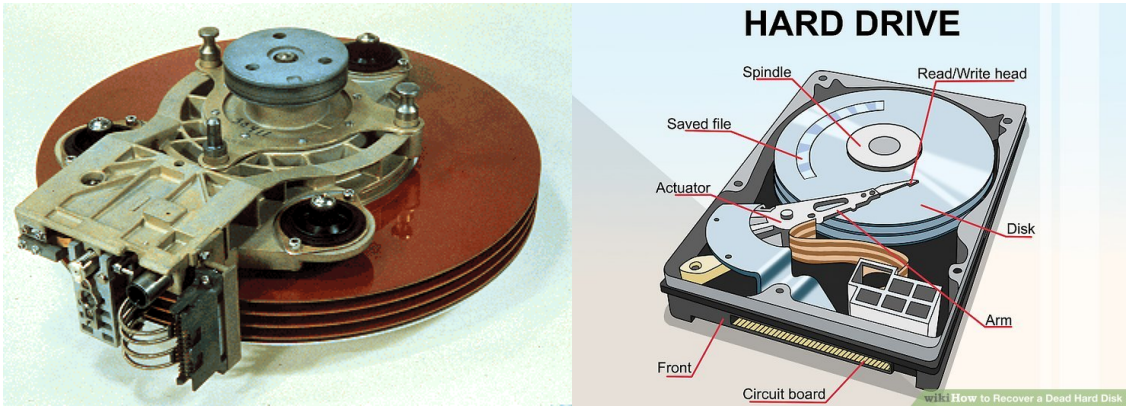


Figure 1.2: The structure of IBM 3340 [18] (left); the structure of today's HDD with labels of the main component [19] (right).

Figure 1.2 shows the structure of a hard disk drive which is also called a "Winchester", following the code name of IBM's groundbreaking product, IBM 3340, in 1973 [20]. Although today's HDDs can hold millions of times more data than the IBM 3340 did, the basic principles and structures are almost the same as the "Winchester". A typical such product consists of platters, motor, spindle, read/write heads, actuator and electronics. The spindle holds flat circular disks, also called platters, in which are stored the recorded data. The platters are made from non-magnetic materials, such as glass or ceramic, coated with layers of magnetic materials on both sides. The read/write heads are assembled together to make a slider, which flies only a few nanometers above the surface of the platters during reading/writing. The slider, controlled by a voice coil actuator, can only move between the center and edge of the disk, and this is why it is called a "slider". When you want to write data, the slider will be moved to some position on the disk. The write head (or writer) will generate a magnetic field to modify the magnetization of the materials on the disk passing immediately under it. Since the direction of magnetic writing field changes according to the data we want to write, there will be alternating patterns on the disk (Figure 1.3), which will form a series of concentric circles later, called tracks, as the disks rotate, and this is because the slider always tries to finish one "circle" as much as possible during writing. Similarly, when you want to read the existing data from the disks, the controller of the actuator will find the track which contains the data, then move the read head (reader) to that track and finally the reader will pick up the magnetization patterns, changing them back to signals.

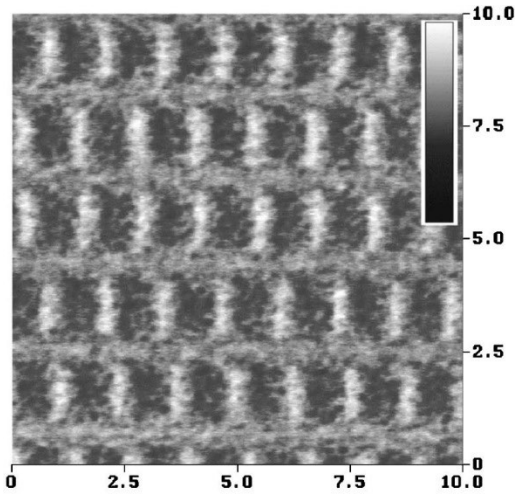


Figure 1.3: Magnetic image (by MFM) of recorded patterns and tracks on a disk [21].

Obviously, the most important components in an HDD are the disks (or media, which is a more preferable term in magnetic recording physics) and sliders (including both the writers and readers). Their performances will determine the areal density or capacity of HDD directly.

1.2.1 Recording Media

The first generation of media was particulate media. And around 1990, people developed thin film media with the help of vacuum technology, which has a smoother surface and higher packing density comparing to particulate media. Thin film media have multilayer structures comprising of glass or aluminum substrate, amorphous NiP undercoat, chromium or Ni-based alloy underlayer, Co-based alloy magnetic film, carbon overcoat, and lubricant. The materials of magnetic recording layer are cobalt alloys including CoCrTa, CoCrPt, CoCrPtTa, CoNiCrPt, etc. [22]–[24]. The added elements such as Cr, Ta

and Pt can increase coercivity, reduce the intergranular exchange interaction and enhance corrosion resistance of the media.

Until 2004, longitudinal magnetic recording (LMR) was the standard methodology in the magnetic recording industry, where the magnetizations of grains lie in the plane of a thin film medium [25]. These magnetizations are written by the longitudinal magnetic fields generated from the ring inductive writing head and this process is shown in the left picture of Figure 1.4. As the areal density increases and grain size decreases, due to the demagnetization field among grains and superparamagnetic effect, it is harder to write more information with a smaller bit length (BL). Around 2005, perpendicular magnetic recording (PMR) [26], [27] took the place of LMR in HDD market. Different from LMR, the easy axis of magnetic particles in PMR aligns perpendicular to the media planes and a perpendicular field is used to write data. In addition, a highly permeable soft underlayer (SUL) is placed beneath the magnetic materials to maintain its strength in perpendicular direction around the writing region, and that is why SUL is also called a "keeper" [25] (right picture in Figure 1.4). Another bonus for PMR is the introduction of exchanged coupled composite (ECC) media at almost the same time (2004) [28], [29], which can reduce its switching field significantly.

1.2.2 Write Head

A writer is generally made from soft magnetic materials with high permeability and electrical resistivity. In the LMR era, people used a ring type head to create the longitudinal transitions on the media (left picture in Figure 1.4). The field generated from the write gap

is relatively small and therefore limits the coercivity of magnetic materials that can be used for the media. In PMR, the monopole structure is introduced to enhance the field strength and gradient along the down track direction (right picture in Figure 1.4). The field between the main pole and SUL are kept perpendicular to the media due to the mirror effect of the SUL. The main pole (P2), return pole or trailing shield (P1) and SUL forms a closed magnetic circuit. Because

$$\Phi = A_{P1}B_{P1} = A_{P2}B_{P2} = A_{SUL}B_{SUL} \tag{1-1}$$

where A_{P2} is the bottom surface of the main pole, B_{P2} is the flux density at the tip of main pole, A_{P1} is the bottom surface of the return pole, B_{P1} is the flux density at the return pole, A_{SUL} and A_{SUL} are the relevant area and flux in the SUL [25], [30], the tip of the main pole should be made as small as possible and the surface of return pole should be larger to avoid overwriting the recorded bits. Besides, inter-track interference (ITI) effect will also become a problem as we increase the areal density and reduce the track width, so it is important to obtain a low adjacent track erasure field in the cross-track direction. Therefore, modern HDDs are using wrap around shield (WAS) to guarantee high field gradients on both down-track and cross-track directions and lower adjacent bit/track erasure fields [31].

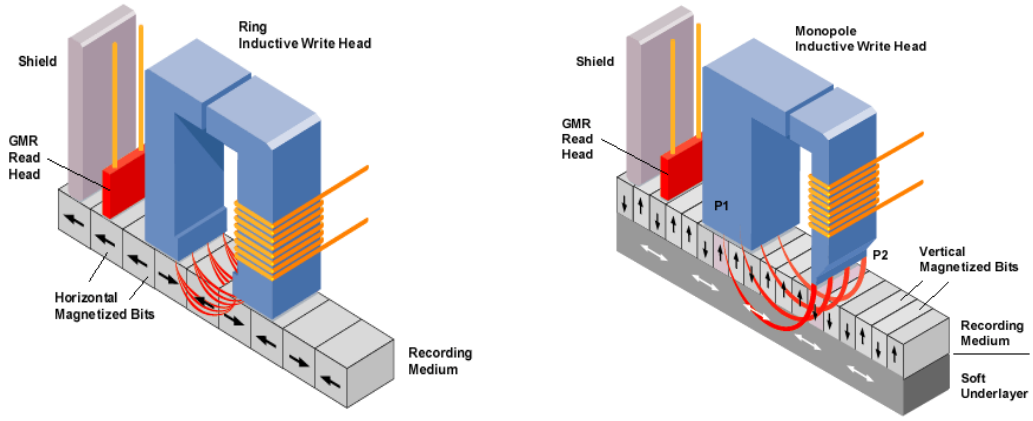


Figure 1.4:Longitudinal magnetic recording (left) and perpendicular magnetic recording (right). [32]

1.2.3 Read Head

Similar to the media and writer, the development history can also be divided into two stages.

Prior to 1990s, an inductive head was used to read the information back in all HDDs. The magnitude of playback signals is proportional to the number of turns of coils wrapped around the head, the magnetic flux from the patterns on the media and the relatively velocity between the head and disk. However, this kind of reader becomes less sensitive as the areal density increases, because there is not enough magnetic flux produced by a smaller bit. The second generation of reader is based on magnetoresistance (MR) effects, in which the playback signals are directly proportional to the MR ratio and independent of the head velocity. Giant magnetoresistance (GMR) is one of such effects that can provide high MR. In 1988, Grünberg and Fert observed GMR effect for the first time in a sandwich structure made up of Fe/Cr/Fe [33], [34]. Its principle can be briefly explained with a two-channel model as shown in Figure 1.5: in ferromagnetic metals the scattering rates of the

up-spin and down-spin electrons are quite different, but both are proportional to their corresponding density of states at the Fermi level. Therefore, resistance for electrons depends on their spin direction and the magnetization direction of the material. The behavior of a general current in such a material, where the elections have two different spins directions, is similar to a parallel circuit with two paths, if we assume the spin relaxation length is much larger than the free mean path, i.e., electrons are not likely to change their spin directions after scattering. When two ferromagnetic layers are put together, the total resistance depends on the magnetization directions of these two layers, and obviously, the parallel case on the left of Figure 1.5 has a lower resistance. In a GMR read head, the magnetization of one ferromagnetic layer is usually pinned by a neighboring anti-ferromagnetic layer via exchange coupling effect, while the magnetization of the other one is free to rotate with external field, which is generated by the magnetic patterns on the media.

There are two types of GMR structures, current in plane (CIP) and current perpendicular to the plane (CPP), depending on the direction of current flow in the device [35], [36]. In the current in plane (CIP) structure, the current flows along the layers, and the electrodes are located on one side of the structure while in the current perpendicular to plane (CPP) configuration the electrodes must be located on the top/bottom of the multi-layer structure. Comparing with CIP, CPP has several advantage: First, read signal strength in CPP head is basically independent of read track width [35]. Second, the intrinsic GMR ratio of a spin valve in CPP geometry is certainly larger than that in CIP geometry [37]–[39]. Third, by directly contacting the GMR film stack with its two shields instead of

leaving a gap between them, we can reduce its size (length) and improve heat reliability, which is extremely important in HAMR [40]. However, the CPP structure has smaller output and is harder to fabricate than CIP, so the CIP GMR head was delivered first in late 1990s, and the CPP GMR are considered to be promising for the future read sensors [41]. In order to enhance the SNR of the readback signal for MR head, usually the GMR structures are placed between two soft magnetic layers to prevent sensing of magnetic flux from adjacent bits. In a reader design, the GMR device is also called an "element" and the soft magnetic layer is called a "shield". The size of element and distance of two shields need to be optimized carefully to achieve the best SNR for the corresponding patterns.

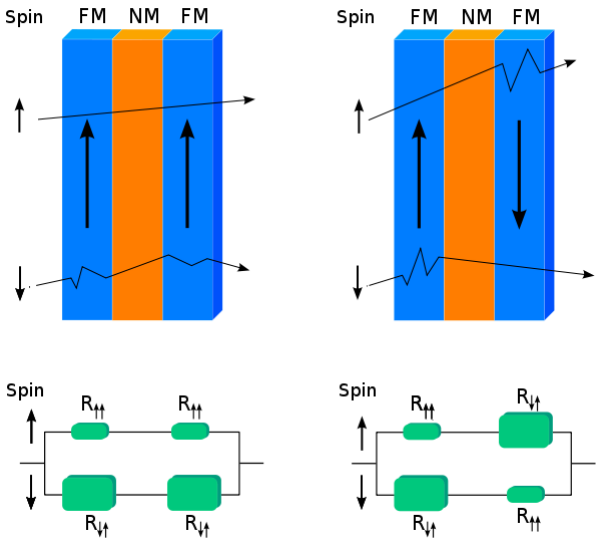


Figure 1.5: Illustration of CPP GMR structure [42]. The parallel state is on the left and the anti-parallel state is on the right with equivalent electric circuit representations, respectively.

1.3 Challenges for HDD above 1Tb/in^2

In order to improve the areal density of HDD further, it is essential to fabricate media with smaller and more uniform grains to maintain sufficient SNRs. However, this is limited by the superparamagnetic effect [43] as well as an intrinsic trilemma in HDD:

Consider a magnetic grain with uniaxial anisotropy and energy barrier E_B equal to the product of the anisotropy constant, K_u, and the grain volume, V, from the Stoner–Wohlfarth model [43]. With the help of continuous thermal agitations, the magnetization keeps trying to switch between the two local minimum energy states separated by the energy barrier. The reversal rate r at any time will be proportional to the frequency factor f_0 and to the Boltzmann factor $e^{-\frac{K_u V}{k_B T}}$, where k_B is Boltzmann's constant, 1.38×10^{-16} erg/K,

$$r = f_0 e^{-\frac{K_u V}{k_B T}} \tag{1-2}$$

To confirm that the magnetism of the grains could be kept for more than 10 years, which means the reversal rate r should be smaller than 1/10 years, we have,

$$\frac{K_u V}{k_B T} > 60 \tag{1-3}$$

Hence, in order to maintain signal to noise ratio (SNR) while down-scaling bits, grain size must be reduced. If the hard disk works in room temperature, the only option is to increase the anisotropy of the recording media. But as we know, coercivity for a single domain particle is proportional to K_u. As a result, the recording heads must provide larger magnetic field as the grain size becomes smaller. However, the magnetic field generated by the write

head is limited by the materials and is theoretically around 2 Tesla [2]. Indeed, this can be understood as a "conflict of interest" among the three key elements, writer, reader and medium, in an HDD: the writer hopes it can write easily, i.e. the switching field or coercivity should be small; the reader hopes it can read easily, i.e. there should be enough grains providing the right information (SNR) in one bit; and the medium hopes it can store the information as long as possible (Figure 1.6). Therefore, the further scaling of PMR will be challenging and its areal density would not reach far beyond 1 Tb/in^2 [27], [44], [45].

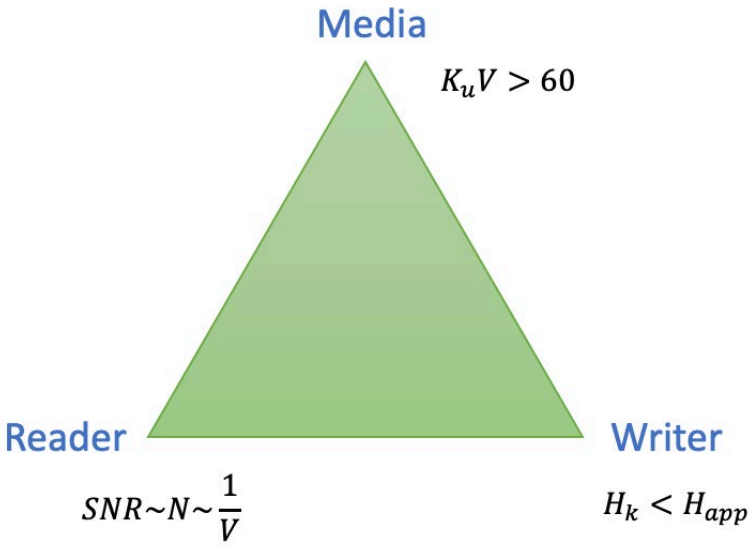


Figure 1.6: Trilemma among media, writer and reader in HDD.

1.4 "The Next Leap Forward is Now"

In recent years, quite a few novel recording schemes and techniques have been proposed to break the trilemma, such as Heat-Assisted Magnetic Recording (HAMR) [46],

Microwave-Assisted Magnetic Recording (MAMR) [47], Acoustic-Assisted Magnetic Recording (AAMR) [48], etc. Among these new features or new schemes, HAMR has been widely agreed to be the leading candidate for the next generation storage. In Intermag 2015, Seagate Technology, who has been the leader in the development of HAMR, demonstrated a prototype of HAMR for the first time [49]. They also announced that the first commercial HAMR drives will enter the market around 2019 [50].

The structure of the HAMR system is shown in Figure 1.7. The most significant difference is the introduction of an optical system, including the laser, light delivery system and near field transducer (NFT), which are used to heat the medium. Another difference is the improvement of media to handle the thermal effect: To write data successfully, the medium must withstand high temperatures and should be cooled down as soon as the head leaves the writing position, otherwise this part will be demagnetized, and the data will be lost. Therefore, a robust head disk interface (overcoat), and a heat sink is needed (Figure 1.10).

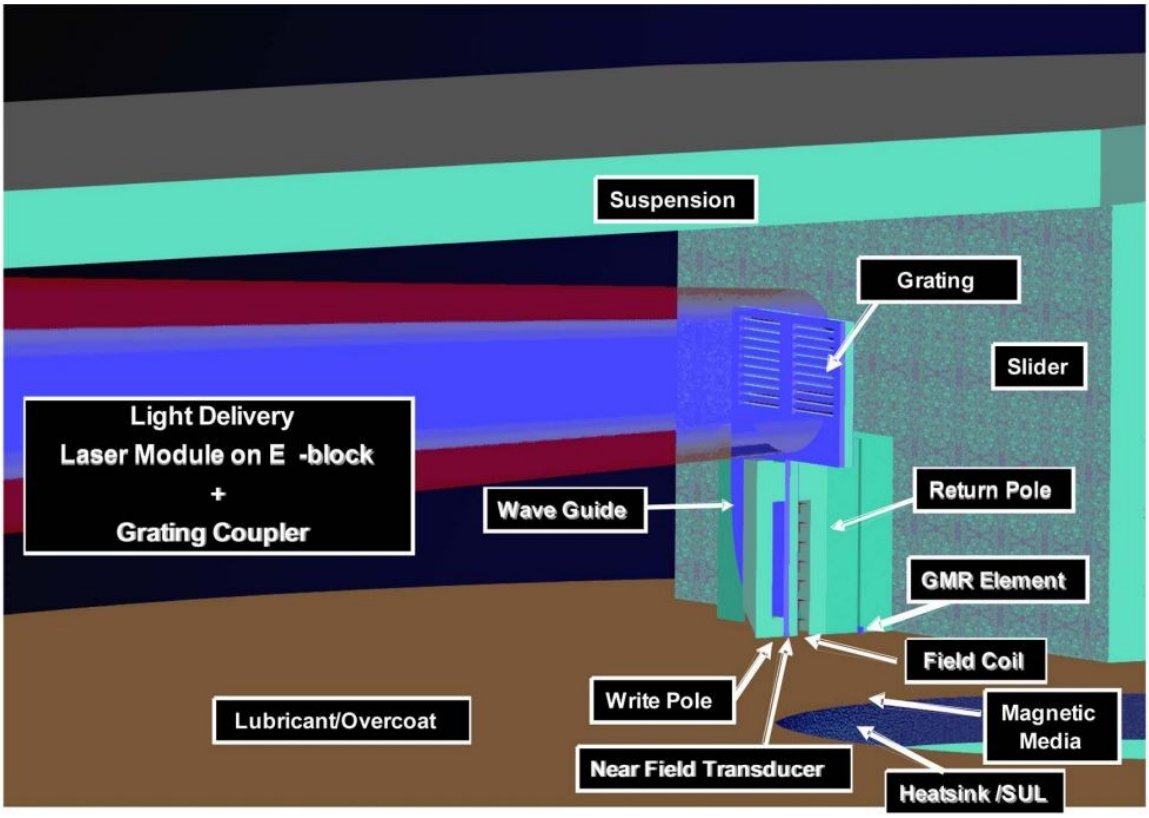


Figure 1.7: A 3d model of HAMR recording system, only showing the slider and medium [46].

1.4.1 Optical system

Figure 1.8 (left) shows a general schematic of the HAMR write head. The laser beam is transferred to NFT by the light delivery system. There are different implementations of light delivery system [51]–[53] but a workflow with the most common elements is shown in the right picture. The light is first coupled to a waveguide by the mode coupler and then propagates down the waveguide. Finally, it couples into the NFT which is the most important part in the write head of HAMR because it focuses and converted the light beam into thermal energy on the medium which is one of the key factors to record data successfully.

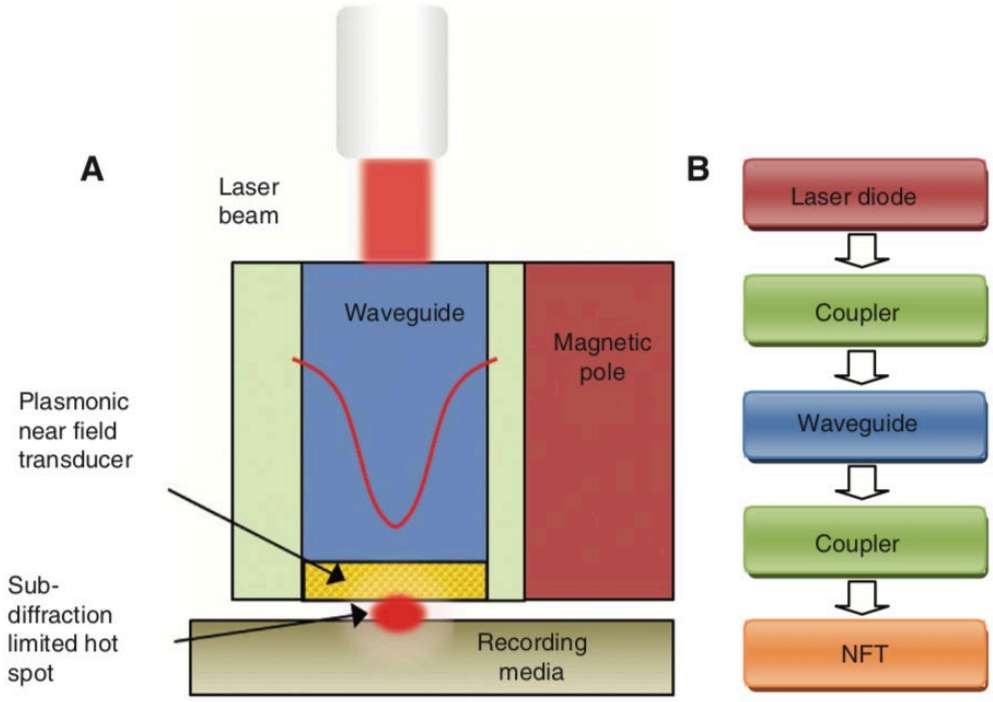


Figure 1.8: Schematic of a HAMR head (left); Common blocks of the light delivery system in a HAMR head (right). [54]

The diffraction limit for the full width at half-maximum (FWHM) optical spot size d as estimated from scalar diffraction theory is

$$d = \frac{0.51\lambda}{N.A.} \tag{1-4}$$

where λ is the wavelength and $N.A.$ is the numerical aperture of the focusing lens. Commercially used short-wave length lasers, which could be installed in a hard disk, have a wavelength of 405nm, and assume $N.A = 0.85$, so we get a focused spot size of about 240 nm. This value is much more than the recording area per bit, about $(25\text{nm})^2$, when assuming the areal density is 1Tb/in^2.

Therefore, the NFT must be able to localize the optical spot to a very small dimension in order to achieve high areal density. This can be achieved using surface plasmons (SPs) effect. Therefore, the most important concern in NFT design is how to excite surface plasmons of a nanostructure, so that it can re-radiate and produce a sub-diffraction-limited light spot. [55] reports such a design that can deliver over 50 μW power into a thermal spot with a diameter of 30 nm. Nowadays, the most commonly used design is the 'lollipop' transducer, as shown in Figure 1.9. The NFT is put at the focus of a parabolically shaped waveguide (called planar solid immersion mirror, PSIM), where all the light converges. The peg on NFT is very sharp and therefore provides the lightning rod effect [56] for field confinement. A thin plasmonic metal (i.e. dielectric metal, so it can be polarized by external field easily) is deposited before the recording layer and acts as both a heat sink and an image plane for the electric field. As a result, the recording layer actually lies within the gap of the NFT and its image, resulting in a substantial enhancement in coupling efficiency and further confinement of the electric field [57], [58].

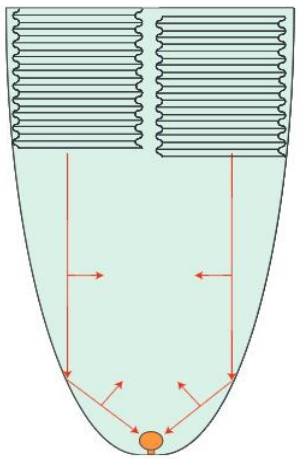


Figure 1.9: A planar solid immersion mirror with a dual offset grating used to focus a waveguide mode onto the lollipop NFT [51].

1.4.2 Recording Media (HAMR)

A schematic of a typical HAMR media structure is illustrated in Figure 1.10. As we see, the structure of the recording media is very complex, and except the recording layer, it also contains a lubricant and overcoat layer, a properly designed heat sink layer, an interlayer for magnetic property and microstructure control and a substrate to support the whole structure.

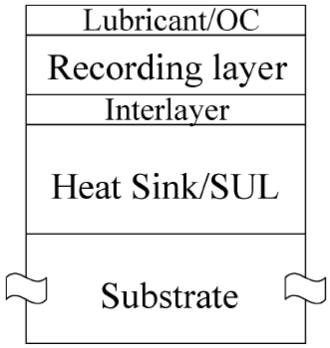


Figure 1.10: HAMR media structure [46].

Table 1 lists several high K_u materials which could be potentially used in recording media. In the table, K_u is the magneto-crystalline anisotropy, M_s is the saturation magnetization, H_K is the anisotropy field, T_c is the Curie temperature and D_P is the critical diameter for stable magnetization state.

Table 1 High K_u magnetic materials [59]

type	material	K_u (10^7erg/cm^3)	M_s (emu/cm^3)	H_k (kOe)	T_c (K)	Dp (nm)
	CoCrPt	0.20	298	13.7	--	10.4
Co alloys	Co	0.45	1400	6.4	1401	8.0
	Co3Pt	2.0	1100	35	--	4.8

$L1_0$ alloys	FePd	1.8	1100	33	760	5.0
	FePt	6.6~10	1140	116	750	3.3~2.8
	CoPt	4.9	800	123	840	3.6
	MnAl	1.7	560	69	650	5.1
Rare-earth transition metals	$Fe_{14}Nd_2B$	4.6	1270	73	585	3.7
	$SmCo_5$	11~20	910	240~400	1000	2.7~2.2

As shown in Table 1, $SmCo_5$ has the largest K_u value among all magnetic materials, about $2x10^8$ erg/cc, and therefore a $SmCo_5$ magnetic particle can be thermally stable, even if its diameter is 2.5 nm. However, $SmCo_5$ has a very large corrosion issue: it is hard to achieve a chemically stable state and as a result, it will decompose gradually in room temperature.

$L1_0$ phase FePt has the second largest magneto-crystalline anisotropy of all common magnetic materials with a critical diameter of about 3 nm. And the most important point is that, unlike $SmCo_5$, it is not only magnetically stable, but also chemically stable. Therefore, it is considered to be the most promising candidate for future extremely high-density magnetic recording media material, especially for the HAMR media.

Figure 1.11 (left) shows the phase diagram of FePt binary alloy. In FePt solid alloy, when the atom ratio of Fe to Pt is around 1 to 1, there are two phases. One is disordered phase with FCC structure, and the other one is ordered phase with FCT structure. Figure 1.11 (right) demonstrates the lattice structure of these two phases. In A1 phase, Fe and Pt atoms are randomly arranged. Its anisotropy is low and cannot be used to store data. In $L1_0$

ordered phase FePt, the Fe and Pt atoms are arranged layer by layer. The c axis of its unit cell is shorter than a and b axis. The asymmetry of the unit cell makes the $L1_0$ FePt have uniaxial magneto-crystalline anisotropy. The easy axis is along c axis. $L1_0$ phase FePt has very large magneto-crystalline anisotropy. To transform A1 phase to $L1_0$ phase, we can either anneal the disordered FePt at high temperature or increase the fabrication temperature of FePt thin film during deposition.

The transformation from FePt A1 phase to $L1_0$ phase has been found to occur when it is heated to around 400 °C [60]. However, a chemically disordered A1 phase FePt thin film has to be annealed at 600 °C to form a fully ordered $L1_0$ FePt layer [61].

In recent years, the research of fabrication of FePt mainly focuses on two points. One is to reduce the exchange interaction between grains, e.g. the size of grains must be very small and in uniform, and also well isolated to reduce to exchange coupling. So researchers try to deposit these films by sputtering technology with some other materials, such as C, SiO_2, B_2O_3, Ag. The other is to reduce the temperature to transform the FePt A1 phase to $L1_0$ phase. However, as shown in Table 2, a dilemma for the fabrication of FePt media with desirable properties is that a high temperature deposition leads to good $L1_0$ ordering and high coercivity, but results in large grain sizes, while a low temperature deposition leads to small grain sizes, but results in poor $L1_0$ ordering thus low coercivity. Several methods are proposed to solve the problem, for example, to add such other atoms as Cu, Ag, Sn and B and to make the sputtering process in a high pressure Ar environment.

Table 2 A summary of FePt granular media fabrication

	Grain Size (nm)	Hc (kOe)	Fabrication temperature (℃)	Reference
FePt-B(-Ag)	2.5	0.05	30	[62]
FePt-Ag-C	6.2	35	550	[63]
FePt-Ag-C	7.2	48	500-550	[64]
FePt-SiN_x-C	5.5	21.5	380	[65]

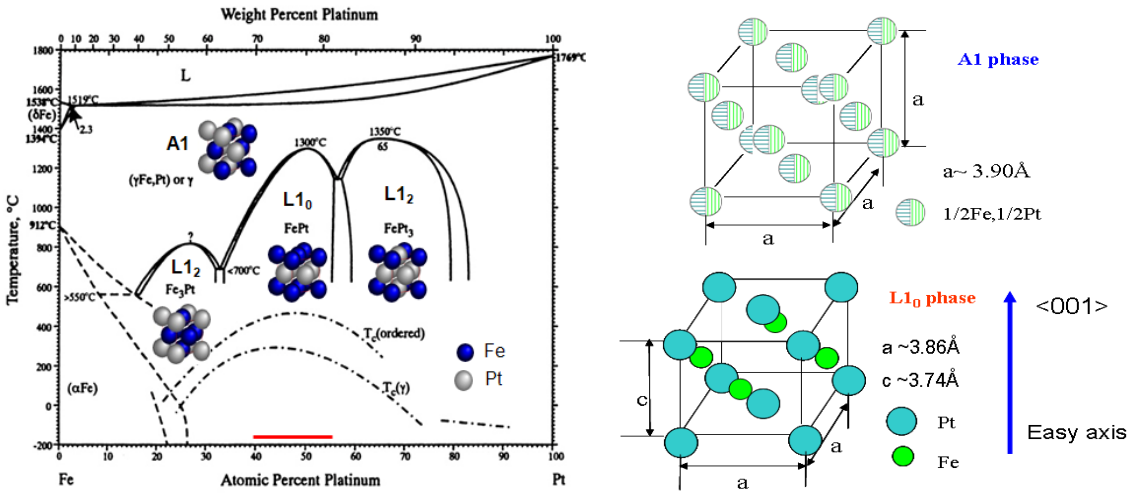


Figure 1.11: Phase diagram of Fe-Pt binary alloy (left) and Lattice structure of A1 phase FePt and $L1_0$ phase FePt (right). [66]

1.4.3 Recording Physics and Noise

As we discussed before, heat assisted magnetic recording (HAMR) technology utilizes laser heat to momentarily lower the coercivity of the material in a localized region below the amplitude of the applied magnetic field from the recording head. Thus, this could enable recording on thermally stable media with extremely small grain size and further extend recording areal density of magnetic data storage. The recording mechanism of

HAMR is illustrated in Figure 1.12. The materials in HAMR have ultra-fine grain size (<5 nm) and a high magneto-crystalline anisotropy which vastly varies across different temperature.

During the recording process, the media needs to be heated to bring down its coercivity below the head field. Then the heat source is removed, and the media cools down quickly to freeze its magnetization for long time storage.

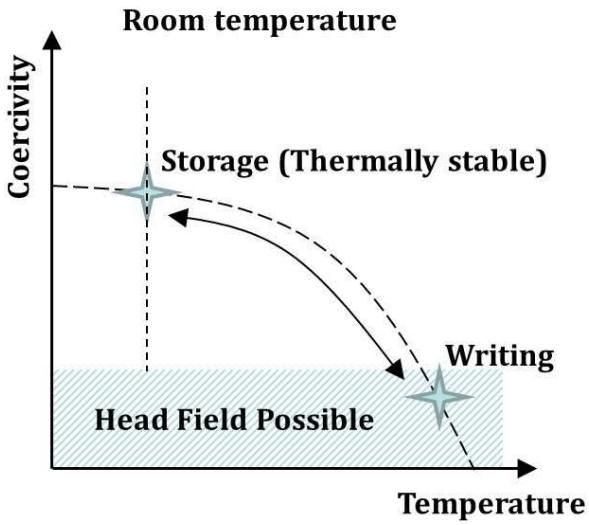


Figure 1.12: Illustration of the working principle of HAMR [46].

Traditionally, medium noise of PMR mainly comes from anisotropy variance σH_k as well as irregular grain sizes and shapes [25]. In HAMR, however, the introduction of thermal effects makes it more complex. First, the writing usually happens at a temperature very close to the Curie temperature of the grains on the media, which is called switching point. However, due to the variations of Curie temperature of grains, the switching point is not constant among different grains (Figure 1.13). As a result, the grains will not switch under the same temperature which will make the pattern boundaries, especially transition

boundaries, more jagged, and this will bring more the noise. Currently, people take Curie temperature variance as a result of the chemical and magnetic ordering, which are caused by doping and grain size fluctuations [67]. The grain size dependence results are shown in Figure 1.14. The dashed line is a guide to the eye and the solid lines are made by fitting the data. We can see that T_c drops significantly after the grain size is smaller than 8nm, which is almost equal to the average value in current industry and this is why T_c variations attract a lot of attention now. In addition, R. H. Victora and Ali Ghoreyshi [68] also proved that structural randomness (i.e. the size, shape and thickness of magnetic grains as well as air gap, grain boundaries etc.) in HAMR media can lead to a significant variation in optical power absorption of the media, which will further result in localized temperature fluctuations on the media (Figure 1.15). This has a similar effect to the Curie temperature variations on the transition noise increase.

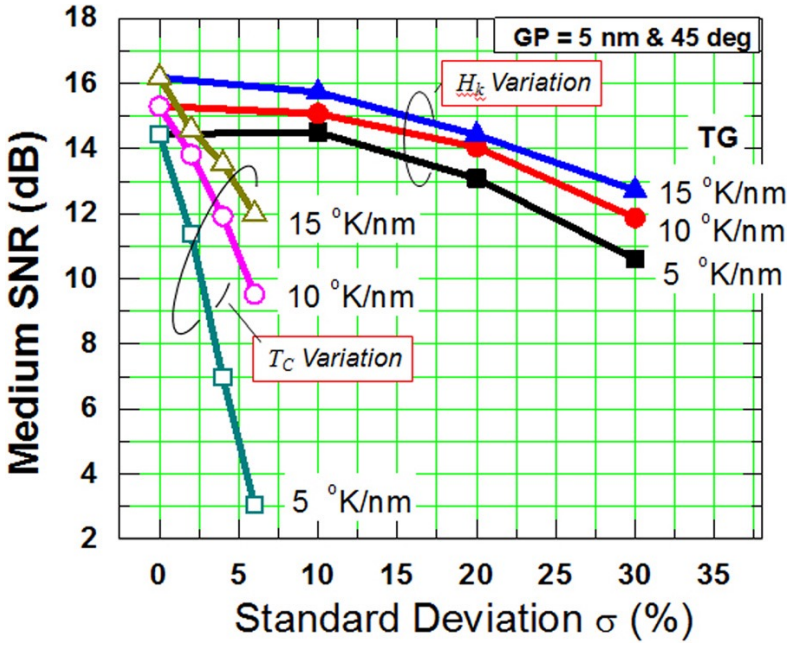


Figure 1.13: The medium SNR as a function of normalized standard deviation for H_k or T_c variation with different thermal gradients [69].

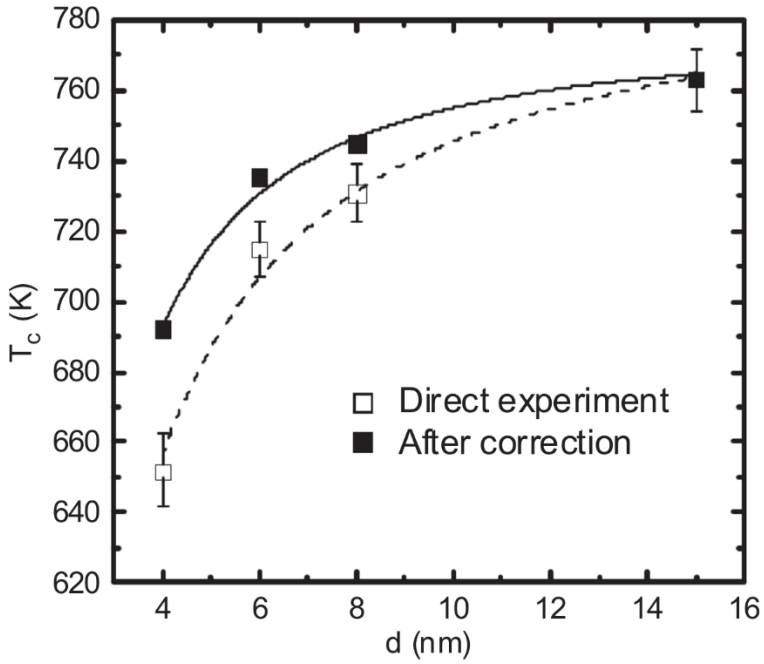
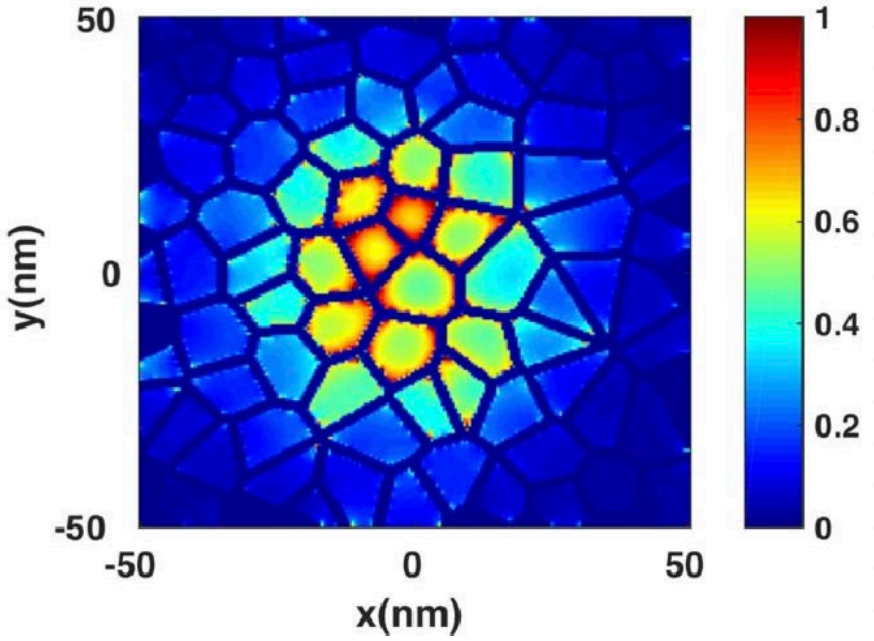
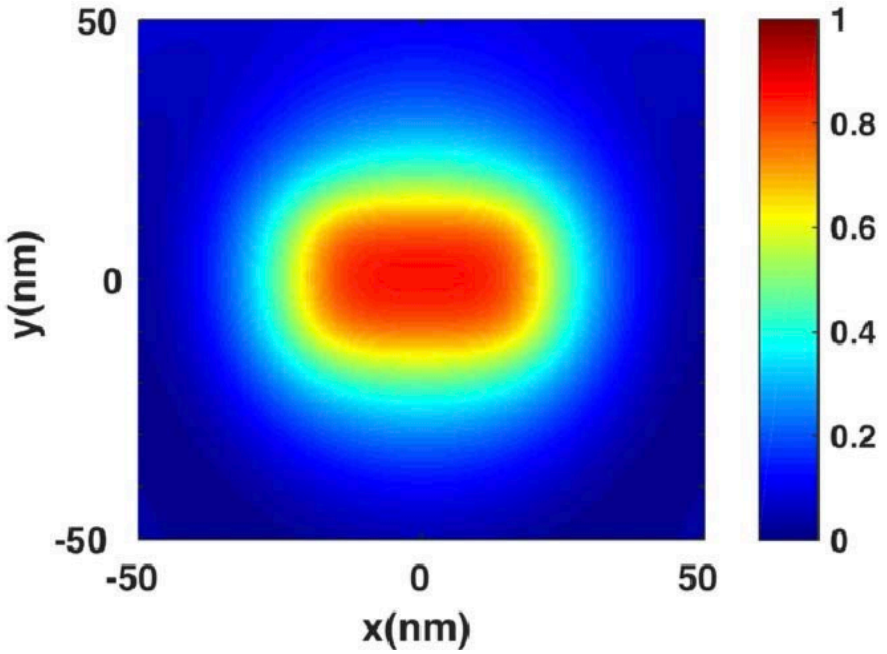


Figure 1.14: The dependence of curie temperature on grain size [67].



(a)



(b)

Figure 1.15: Absorption power inside (a) granular and (b) effective medium approximation of HAMR media. [68]

Second, since the thermal fluctuation field can help the write field to switch the magnetization direction in a grain at high temperatures, it can also make it switch back occasionally due to the stochasticity of thermal fluctuation process. Therefore, the probability for the grains to reverse their magnetization direction follows a probability distribution function, which is called switching probability distribution (SPD) function [70], [71]. [72] gives one probable form of SPD. There does exist a maximum switching probability in the SPD, which can still be considered as the "switching point", though no points can reach 100%. Similar to the first point, this issue will also introduce noise which can hardly be eliminated (Figure 1.16). In other words, the ultimate areal density is limited by the thermal stability of the recording grains during the dynamic writing process [73].

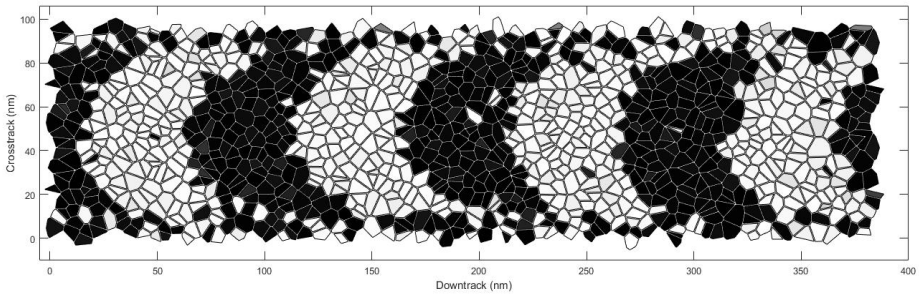


Figure 1.16: Recording magnetization patterns w/o T_c variations when H = 10 kOe. Some grains aren't switched successfully as expected because of thermal fluctuations.

Third, unlike traditional PMR, the recorded shapes are controlled by both the thermal profile and writing field on the media. Currently, most pegs in NFT are round or ellipse ones, so the thermal profiles also have similar shapes, which, in combination with the writing field, usually a spatially uniform magnetic profile in the recoding area, always creates transitions with a significant curvature as Figure 1.17 shows [74]–[77]. As a result, the pulse width in readback signals will be increased, as measured by PW_{50} (the read pulse width at 50% amplitude) and therefore, there will be more overlap regions and inter-symbol interferences (ISIs).

1.5 Summery and Outline of this Work

Form the discussions above, we can see the introduction of heat brings lots of issues though it can "break" the trilemma in magnetic recording and increase the areal density further. Even worse, there will be more parameters to optimize in this system, making the calculation complexity grow exponentially. Most of the theories or equations established

for PMR may not suitable for HAMR. As a result, numerically based optimization is extremely necessary for HAMR development and we will organize the thesis as follows:

Chapter 2 and 3 will describe the details of our modeling methods as well as some considerations during simulation, including computational geometry algorithms, micromagnetic models and signal processing algorithms. We will also discuss how to use general-purpose graphics processing units (GPGPU) to accelerate our calculation.

Based on our simulation program, in Chapter 4 and 5 we will optimize HAMR at the system level. A general relationship between SNR and bit error rate (BER) is obtained, which is used to predict the area density later. We will give some feasible suggestions to optimize HAMR, compared with the results from perfect recording.

Chapter 6 will present two methods to split different types of noise in playback signals of HAMR. We will try to further understand these noises and their relationship with jitter.

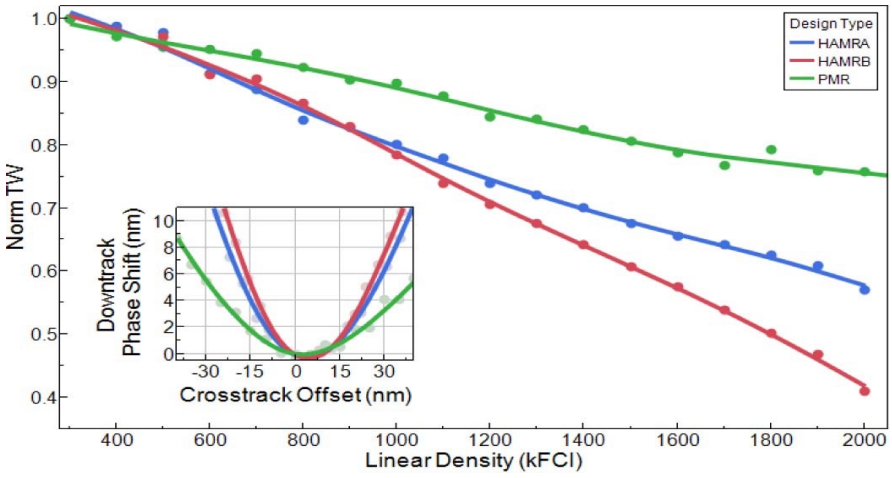


Figure 1.17: Track width normalized to that at 300 KFCI for a PMR head (green line), and two different HAMR designs. Inset: the measured curvature for all three designs, showing higher curvature for the HAMR heads, with HAMR B (red line) being worse than HAMR A. [74]

2 Computational and Modeling Methods

2.1 Simulation program overview

Basically, a complete magnetic recording system simulation can be divided into two parts (Figure 2.1): *Writing*, which means we change a binary file into a series of signals and then use these signals to vary the magnetic field generated by the write head. This field is then used to magnetize the media. As a result, some patterns, which will carry the information of the original binary file, will finally be recorded on the medium in a hard drive. *Reading*, is a process just opposite to writing. We use another head ("reader") to pick up the magnetization changes on the medium, getting the signal back, and then, this signal will be changed back to binary files. This step is called state detection. It is very necessary because there may be some distortion or overlaps in the analog signal, especially when the frequency is high (or BL is small). This is the same as transmission channel. Actually, we usually call the write process as write channel and read process as read channel [78].

In order to represent these processes successfully, our simulation program can be divided into four modules, which is show in Figure 2.2. The first two modules (media generation and recording simulation) are used to mimic the writing process, while the last two (playback and state detection) are corresponding to the reading process.

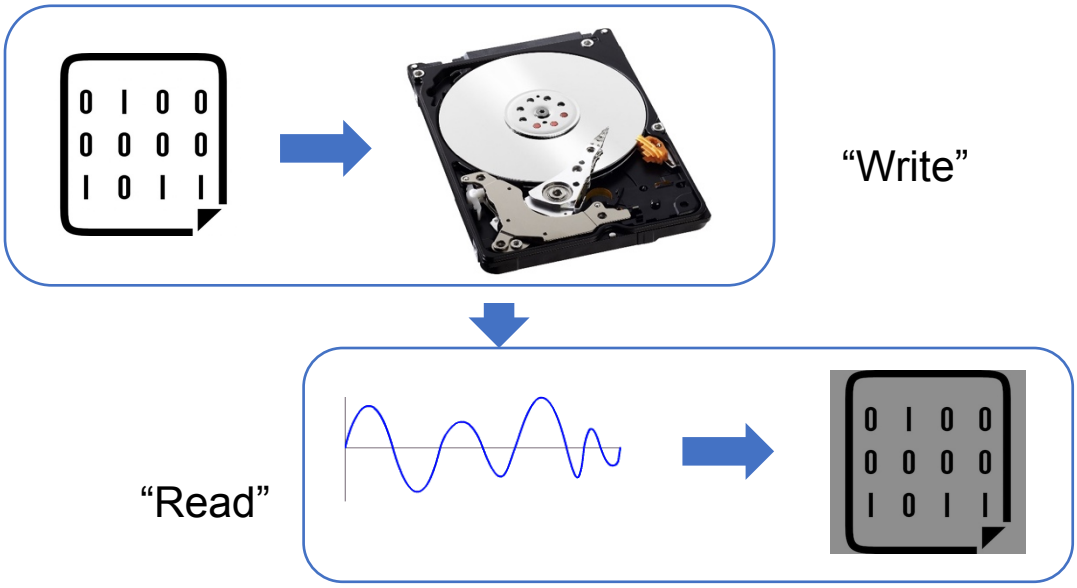


Figure 2.1: Working principles of an HDD interpreted with the "write/read" channel model

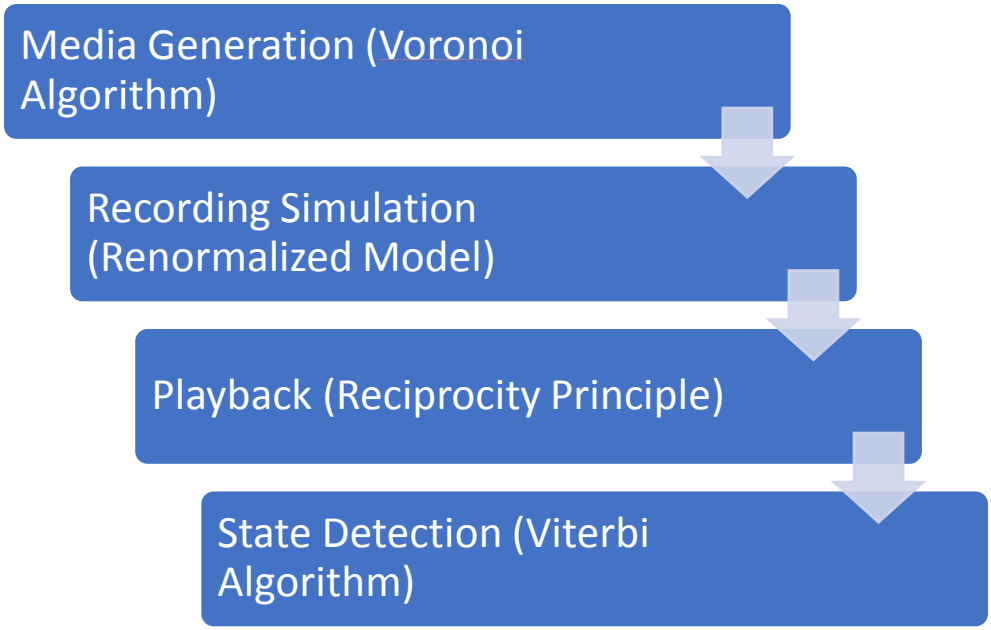


Figure 2.2: Basic modules in our HAMR simulation program.

2.2 Media generation

The first part is media generation module. Currently, commercial media consist of millions of randomly distributed grains and Voronoi diagram [79] is the best mathematical model to describe such a system [80].

Generation of a Voronoi diagram is an interesting but difficult work and there are lots of publications on this topic, which can be classified as three kinds of algorithms: naïve algorithm, incremental algorithm and sweeping line algorithm.

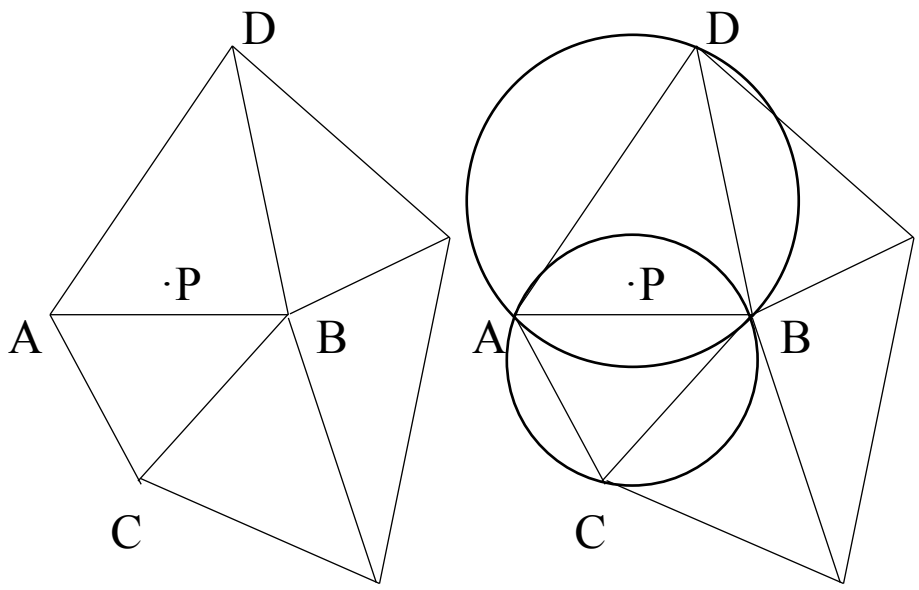
Naïve algorithm is the simplest way to generate Voronoi diagram. Based on its definition, to generate a Voronoi cell, we can: 1. Draw the perpendicular bisectors of every two seeds (also called site points); 2. Find the smallest area including the site point bounded by all these bisectors. That is,

$$Vor(p_i) = \bigcap\nolimits_{1 \le j \le n, j \ne i} h(p_i, p_j) \tag{2-1}$$

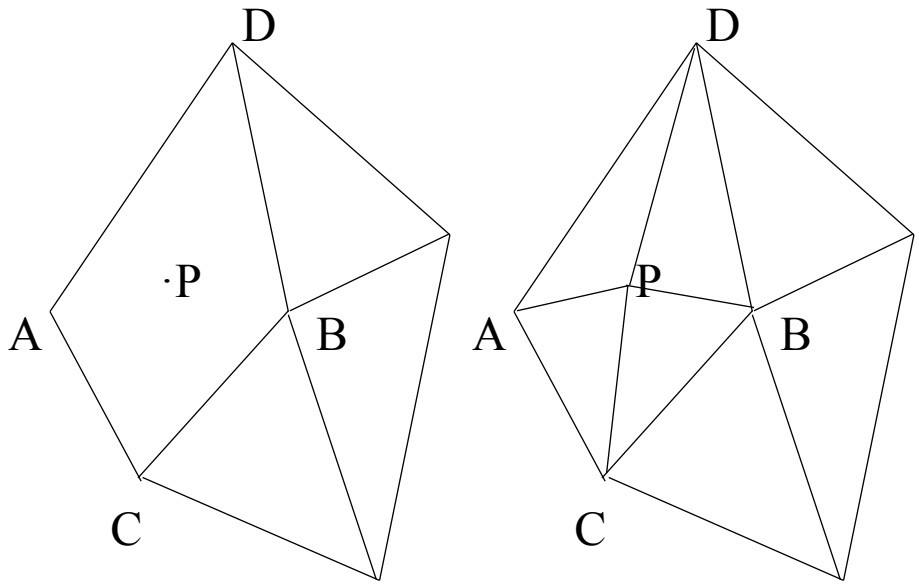
where p_i is the site point of the current Voronoi cell $Vor(p_i)$ and $h(p_i, p_j)$ is half-plane that contains p_i. In order to find all the half-planes, we need two for-loops whose time complexity is $O(n^2)$ or to use divide and conquer whose time complexity is $O(nlogn)$. As a result, the total time complexity for n grains in a medium for n cells is $O(n^3)$ or $O(n^2 logn)$.

The basic idea of incremental algorithms is to add the site points one by one into the diagram. When the inserting process is finished, the entire Voronoi diagram is generated. Bowyer–Watson algorithm [81] is one classical incremental method: 1. Find a super triangular which can completely contain all the site points. 2. Add the site points one by one into the super triangle. Whenever we add a new point, we should connect this point

to its closest three points and check whether the Delaunay triangulation [82] is satisfied, i.e., the circumcircle of any triangle inside the super triangular cannot contain this new site point. Otherwise, the triangles which violate the rule will be broken into points again and be used to regenerate new triangles with the current site point. Figure 2.3 shows this process. From this generation process, it is obvious that the circumcenters of Delaunay triangles are the vertices of Voronoi cells and we can get the final diagram by connecting all the adjacent vertices (Figure 2.4). In this method, it is only needed to check every triangle once when a new point is inserted, so the overall time complexity is $O(n^2)$.



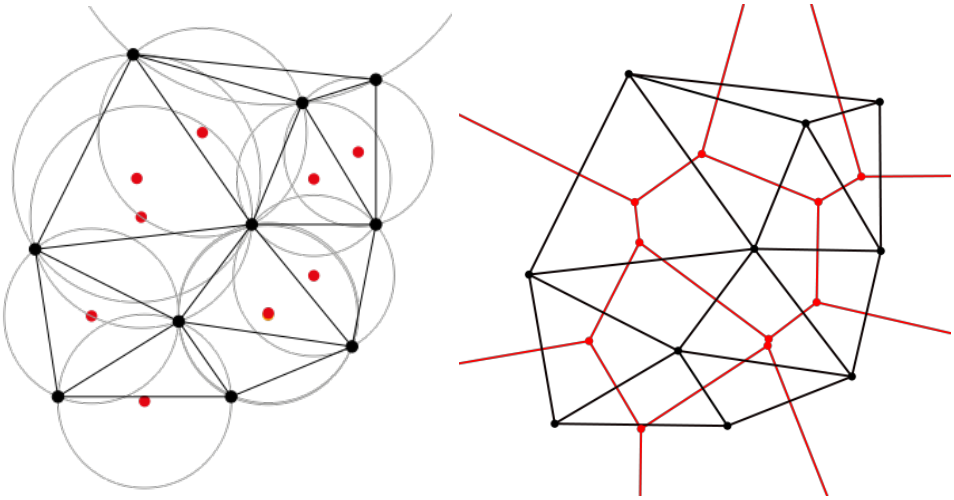
(a) (b)



(c) (d)

Figure 2.3: Bowyer–Watson algorithm: (a) a new point P is added into Delaunay mesh; (b) P is in the two circumcircles of $\odot ABD$ and $\odot ABC$; (c) $\triangle ABD$ and $\triangle ABC$ are broken; (d) regenerate Delaunay tringles

The third type is $O(nlogn)$ level method, which needs the support of some very complex data structures. For example, in Fortune's algorithm [83], [84], we maintain a combined data structure of a binary search tree, priority queue and doubly connected edge list (DCEL) as the sweeping line moves to generate Voronoi cells.



(a) (b)

Figure 2.4: Relationship between Delaunay triangles (black) and Voronoi cells (red): Voronoi seeds is the vertices of Delaunay triangles (left) while circumcenters of Delaunay triangles is the vertices of Voronoi polygons (right). [82]

Due to this problem, in our media generation module, we use the second method instead, though it has a little larger time complexity. Actually, a very large medium with too many grains is seldomly used because of its long running time even if we can use general-purpose computing on graphics processing units (GPGPUs) to accelerate the calculation [85]. Instead, we usually generate a section of the medium and then repeat this section for several times to make a larger one, which is shown in Figure 2.5, except when it will affect the results largely with a fully randomized Voronoi medium. Therefore, the number of Voronoi cells is always less than 2,000 in our simulation and in such a range, Bowyer–Watson algorithm and Fortune's algorithm show the similar running speed, but the former has a simpler data structure and the code will be more easily maintained.

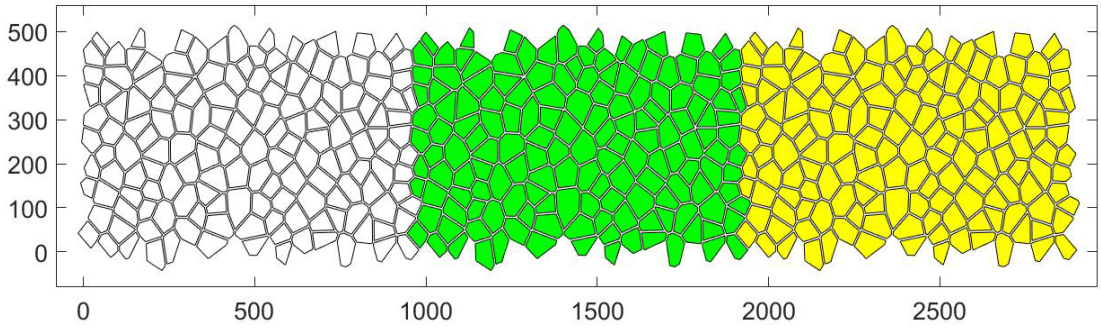


Figure 2.5: A medium sample using Voronoi cells with grain boundaries. The parts in green and yellow color are repetitions of the leftmost part in white. Therefore, the edges on both sides must match with each other. The units for both axes are angstroms.

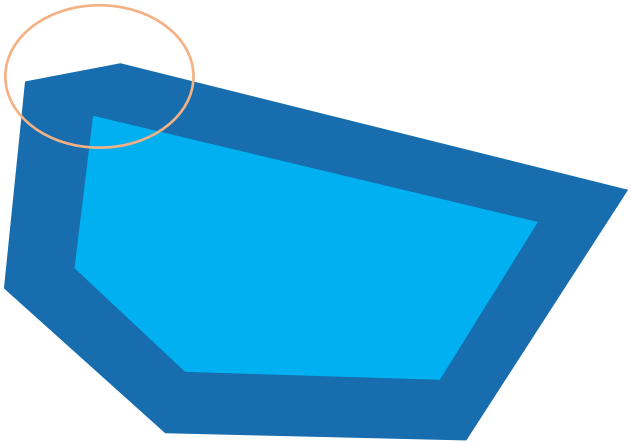


Figure 2.6: A short edge may disappear after the polygon is shrunk. The original polygon is filled with dark blue and the shrunk one is filled with light blue.

The grain boundary is another concern in the media generation module, which is classified as a polygon offsetting problem. This is difficult because after the polygon shrinks, some edges may disappear and our program needs to generate the new polygon correctly (Figure 2.6). [86] gives one possible solution to this problem. However, in this method we need to store all the possible polygons in memory in order to pick up the right one, which has a low memory efficiency. Actually, based on this special case, we can use

a simpler method: 1. Mark the edges of the original polygon in a direction, either clockwise or anti-clockwise, which is the same as the first step in [86]. 2. Find longest edge E_{max} in this polygon. 3. Starting from E_{max}, calculate the shrunk edges one by one. The calculation method for one edge is shown in the Figure 2.7 below. The reason for starting with E_{max} is because E_{max} will never disappear after offsetting so the calculation result is always valid. Assuming the polygons are stored in linked lists in memory, the worst runtime for shrinking one edge is $O(n)$, and the overall time complexity is $O(n^2)$. We can use binary search in the calculation of crossing points, whose overall time complexity is the same as the method in [86], $O(nlogn)$, where n is the number of edges in one polygon. Actually, the number of edges for Voronoi cells in our simulation is usually smaller than 10. Therefore, the overall runtime of shrinking one polygon is a constant, $O(1)$. And the total runtime of calculating the whole medium is almost $O(n)$. In addition, this method can only handle convex polygons while the first method is compatible with any cases include the concave, though this feature is useless for Voronoi polygons.

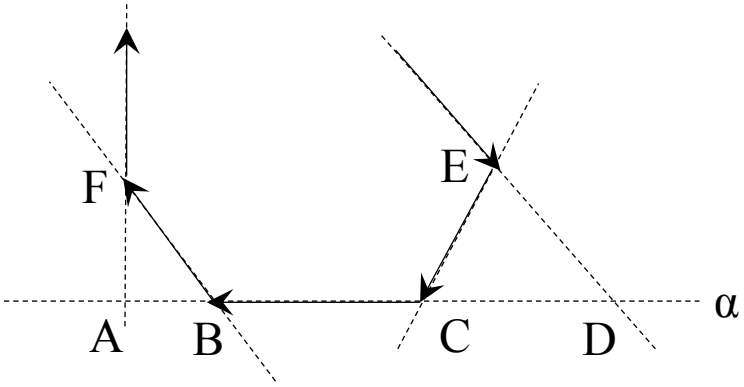


Figure 2.7: A case where several edges or their extension lines cross with the edge α at A, B, C and D after shrinking, whose directions are also marked. In order to pick up the two points after offsetting, we divided these crossing points into two groups, based on their cross products, i.e. $\overrightarrow{CB} \times \overrightarrow{FB}$, $\overrightarrow{CB} \times \overrightarrow{FA}$ fall into one group and $\overrightarrow{CB} \times \overrightarrow{EC}$, $\overrightarrow{CB} \times \overrightarrow{ED}$ make another one. Indeed, the first group means the edges will "leave" from edge α while the

second means they will "arriving" at edge α. Then we can pick up the point in group one which has the minimum distance to any points in group two and vice versa.

2.3 Recording Simulation: Micromagnetics and Renormalized LLG model

Although, there have been several Landau–Lifshitz–Gilbert (LLG) [87] or Landau–Lifshitz–Bloch (LLB) [88] based micromagnetic models for HAMR process, most of them utilized simplified assumptions or introduced empirical parameters, which are unlikely to be accurate, especially when the temperature of the media is very close to the Curie Temperature. Recently, Victora et al. [89] developed a new simulation method using renormalization group theory and LLG equation that provided a systematic technique for improving accuracy. In this section, we will introduce this basic renormalized-LLG model used in our simulation program and in next chapter, we will make some corrections to the basic model by adding one more parameter to our simulation program.

2.3.1 Landau-Lifshitz-Gilbert (LLG) equation

As we know, the Landau-Lifshitz-Gilbert (LLG) equation describes the dynamic motion of one single spin $\vec{M}$ under an equivalent field $\vec{H}_{eff}$:

$$\frac{d\vec{M}}{dt} = -\frac{\gamma}{1+\alpha^2}\vec{M}\times\vec{H}_{eff} - \frac{\gamma\alpha}{(1+\alpha^2)M_S}\vec{M}\times\left(\vec{M}\times\vec{H}_{eff}\right) \tag{2-2}$$

where M_S is the saturation magnetization, γ is the gyromagnetic ratio with a constant value of 1.76×10^7 $Oe^{-1}s^{-1}$, α is the Gilbert damping parameter, whose value is determined

empirically, and as a typical value, it is 0.01~0.05. $\vec{H}_{eff}$ is the effective magnetic field. The first term on the right-hand side is a torque exerted by the magnetic field $\vec{H}_{eff}$ and is responsible for the gyromagnetic precession of the magnetic moment $\vec{M}$. The frequency of the gyromotion is found to be $\left|\frac{\gamma}{1+\alpha^2}\vec{H}_{eff}\right|$. The second term describes the damping of the magnetization $\vec{M}$ towards the direction of the magnetic field $\vec{H}_{eff}$. The damping rate is proportional to $\frac{\gamma\alpha}{(1+\alpha^2)M_S}$. Due to the cross product $\vec{M} \times \vec{H}_{eff}$, both terms are orthogonal to the direction of the magnetization. Hence, the amplitude of the magnetization is preserved which is assumed in the derivation.

In the LLG equation (2-2), in order to calculate the effective field $\vec{H}_{eff}$, we must use magnetic Gibb's free energy,

$$\vec{H}_{eff} = \frac{\delta E}{\delta \vec{M}} \tag{2-3}$$

where the energy terms in the magnetic grains of the recording media are discussed based on the classical theories of magnetism [90], i.e.

$$E = E_{exchange} + E_{anisotropy} + E_{applied} + E_{magnetostatic} + E_{thermal} \tag{2-4}$$

where $E_{anisotropy}$ is the crystalline anisotropy energy, $E_{magnetostatic}$ is the magnetostatic interaction energy, $E_{applied}$ is Zeeman energy caused by external applied field, $E_{exchange}$ is the intergranular exchange coupling energy, and $E_{thermal}$ is the thermal agitation. We will discuss each term in detail below.

Crystalline Anisotropy Energy

In the HAMR media, the materials of the magnetic grains which people usually use (such as FePt) have uniaxial crystalline anisotropy, and it could be expressed as [43]

$$E_{anisotropy} = K_u \, sin^2 \, \theta = K_u \left(1 - \left(\vec{k} \cdot \vec{m}\right)^2\right) \tag{2-5}$$

Hence

$$\vec{H}_{anisotropy} = \frac{\delta E_{anisotropy}}{\delta \vec{M}} = \frac{2K_u}{M_S}\left(\vec{k} \cdot \vec{m}\right)\vec{k} \tag{2-6}$$

where K_u is the anisotropy energy constant and θ is the angle between the anisotropy easy axis and the magnetization direction. $\vec{k}$ and $\vec{m}$ are the unit vectors of the easy axis orientation and the magnetization, respectively. M_S is the saturation magnetization, e.g. $\vec{M}=M_S\vec{m}$.

Exchange Coupling Energy

The dominant contribution to the magnetic Gibb's free energy in ferromagnetic materials comes from the internal or Weiss field, which attempts to align the atomic spin moments. The Weiss field in fact originates from the quantum mechanical exchange interaction, arising from the probability of an electron moving from one atomic site to another. The exchange interaction, as it is called, leads to very strong alignment of spin moments to their neighbors in ferromagnetic metals. The total exchange energy for each atom, i, is described by the sum over all neighboring atomic spin moments:

$$E_{exchange} = -\sum_{i \neq j} J_{ij} \cdot \vec{S}_i \cdot \vec{S}_j \tag{2-7}$$

where $\overrightarrow{S_i}$ and $\overrightarrow{S_j}$ are the magnetic spin moment of spin i and j, respectively, so $|\overrightarrow{S_i}| = |\overrightarrow{S_j}|$, and J_{ij} is the exchange coupling constant determined by material properties. Assuming isotropic nearest neighbor couplings, i.e. J_{ij} is a constant value J, and the exchange energy is

$$E_{exchange} = -J\sum_{i\neq j}\vec{S}_i\cdot\vec{S}_j = -J\sum_{i\neq j}\vec{S}_i\cdot\left(\vec{S}_i+\Delta\vec{S}_{ij}\right) = -J\sum_{i\neq j}\vec{S}_i\cdot\Delta\vec{S}_{ij} - J\sum_{i\neq j}\vec{S}_i^{\,2} \tag{2-8}$$

Since

$$\vec{S}_j^{\,2} = \left(\vec{S}_i+\Delta\vec{S}_{ij}\right)^2 = \vec{S}_i^{\,2} + 2\vec{S}_i\cdot\Delta\vec{S}_{ij} + \Delta\vec{S}_{ij}^{\,2} \tag{2-9}$$

$$\Rightarrow \vec{S}_i\cdot\Delta\vec{S}_{ij} = -\frac{\Delta\vec{S}_{ij}^{\,2}}{2} \tag{2-10}$$

we have

$$E_{exchange} = J\sum_{i\neq j}\frac{\Delta\vec{S}_{ij}^{\,2}}{2} - J\sum_{i\neq j}\vec{S}_i^{\,2} \tag{2-11}$$

Consider a simple cubic with lattice constant a and assume $A = J\vec{S}^2 = J\vec{S}_{ij}^{\,2}$, where A is the exchange stiffness constant of the material, using the continuum approximation,

$$E_{exchange} = \frac{2J}{a}\iiint_V \frac{1}{2}\left(\left|\frac{\partial\vec{S}_{ij}}{\partial x}\right|^2 + \left|\frac{\partial\vec{S}_{ij}}{\partial y}\right|^2 + \left|\frac{\partial\vec{S}_{ij}}{\partial z}\right|^2\right)dxdydz + const$$

(2-12)

$$E_{exchange} = \frac{A}{{M_S}^2}\left(\left|\nabla_x \vec{M}\right|^2 + \left|\nabla_y \vec{M}\right|^2 + \left|\nabla_z \vec{M}\right|^2\right)$$

(2-13)

Because

$$\vec{H}_{exchange} = \frac{\partial E_{exchange}}{\partial \vec{M}}$$

(2-14)

A final solution will be given as

$$\vec{H}_{exchange} = \frac{2A}{{M_S}^2}\left(\nabla_x^2 \vec{M} + \nabla_y^2 \vec{M} + \nabla_z^2 \vec{M}\right) = -\frac{2A}{M_S a^2}\sum_{nn} \vec{m}$$

(2-15)

a is the distance of two neighboring grains (usually, the lattice constant), $\vec{M}$ is the vector of magnetization whose magnitude is M_S, with $\vec{M} = M_S \vec{m}$ and "nn" means "the nearest neighbor". For a simple cubic packing structure, "nn" equals 6 which is used in a common FePt exchange model.

Zeeman Energy

The Zeeman energy is the magnetic energy of a magnetic moment under an applied field $\vec{\mathrm{H}}_{\mathrm{applied}}$. Obviously, this term is a known quantity and can be used directly.

Magnetostatic Energy

Magnetostatic energy is the interaction between the magnetization and the magnetic field. This reflects the long-range dipole-dipole interaction between spins at atomic lattice sites within magnetic bodies. Hence different from the exchange field, the magnetostatic

energy in one grain arises from the its interaction with all the other magnetic gains in the system. The calculation of the magnetostatic energy is the most time-consuming part of an LLG solving program. Basically, if we treat each grain as a dipole, we can get the magnetostatic field at position $\vec{r}$,

$$\vec{H}_{dipole}(\vec{r}) = -\frac{1}{4\pi}\left[\frac{3\vec{r}\left(\vec{M}\cdot\vec{r}\right)}{r^5} - \frac{\vec{M}}{r^3}\right] \tag{2-16}$$

Therefore, the total magnetostatic field, which is also called demagnetization field, at position $\vec{r}$ is,

$$\vec{H}_{\text{magnetostatic}}(\vec{r}) = -\frac{1}{4\pi}\int\left\{\frac{3\left(\vec{r}-\vec{r'}\right)\left[\vec{M}\left(\vec{r'}\right)\cdot\left(\vec{r}-\vec{r'}\right)\right]}{\left|\vec{r}-\vec{r'}\right|^5} - \frac{\vec{M}\left(\vec{r'}\right)}{\left|\vec{r}-\vec{r'}\right|^3}\right\}d\vec{r'} \tag{2-17}$$

which is a triple integral in space R^3. If we assume a medium is made up of $m \times n \times l$ dipoles, using its discrete form, the runtime of calculating magnetostatic field for one dipole is $O(mnl)$ and overall time complexity for the whole medium is $m \times n \times l \times O(mnl) = O[(mnl)^2]$. Obviously, this is a very large number and unacceptable in simulation. In 1988, M. Mansuripur and R. Giles introduced the FFT method into the calculation of magnetostatic field [91], improving the simulation speed significantly. Their method is based on this observation:

$$\vec{H}_{\text{magnetostatic}}(\vec{r}) = -\frac{1}{4\pi}\int \widetilde{N}\left(\vec{r}-\vec{r'}\right)\cdot\vec{M}\left(\vec{r'}\right)d\vec{r'} \tag{2-18}$$

This actually is a convention of $\widetilde{N}(\vec{r})$ and $\vec{M}(\vec{r})$, i.e.

$$\vec{H}_{\text{magnetostatic}}(\vec{r}) = \widetilde{N} * \vec{M}(\vec{r})$$

(2-19)

If all the dipoles are uniformly distributed in the space R^3, we can use FFT to solve this problem, which can reduce the total time complexity to $O(mnl \times log(mnl))$. In addition, $\widetilde{N}$ is the magnetostatic interaction matrix which only depends on the geometry of the mesh array and the mesh shape, and therefore, this matrix is only needed to be calculated once in the simulation no matter how frequently the magnetization of the medium is changed.

Thermal Agitation

By now, our LLG equation can only describe the dynamic properties of magnetization at 0K. In Chapter 1, it is mentioned that any magnetic particle will be subjected to thermal fluctuations with random directions and magnitudes above 0K, which can cause phase transitions from ferromagnetic to paramagnetic states, and we also used the energy barrier model to understand this phenomenon. However, this is a static model and cannot be used to investigate the dynamic process of the thermal fluctuations. In [92] a new approach was proposed in which the effective field is

$$\sigma = \sqrt{\frac{2k_B T\alpha}{\gamma V M_S \delta t}}$$

(2-20)

$$\vec{H}_{thermal} = \sigma \times \vec{h}$$

(2-21)

Here, V is the grain volume, and δt is the time step of the simulation, α is the damping constant. γ is the gyromagnetic constant, k_B is the Boltzmann constant and T is the temperature. $\vec{h} = h_1\vec{x} + h_2\vec{y} + h_3\vec{z}$, and $\vec{x}$, $\vec{y}$, $\vec{z}$ are unit vectors the Cartesian axes X, Y, Z, respectively. h_i is a white noise signal, i.e.

$$\langle h_i(t) \rangle = 0 \tag{2-22}$$

$$\langle h_i(t_1) h_j(t_2) \rangle = \delta_{ij}\delta(t_1 - t_2) \tag{2-23}$$

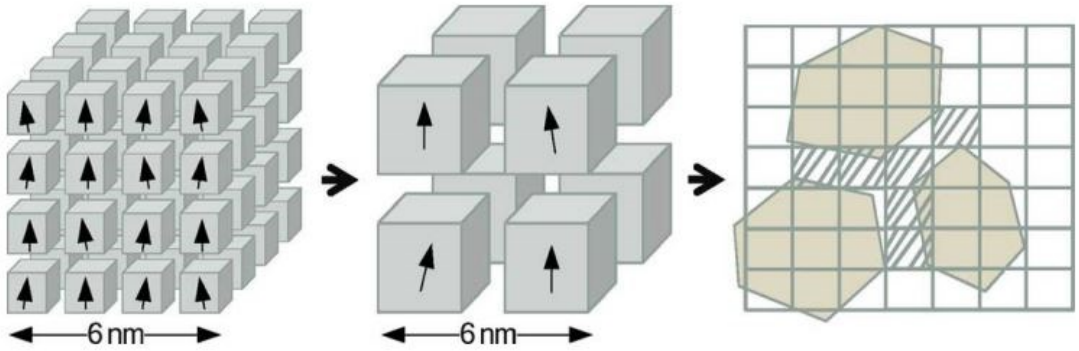
where the indices $i, j = 1,2,3$. δ_{ij} is Kronecker's delta which means that different directions in the fluctuation field are not correlated with each other and $\delta(t)$ is Dirac delta function with the assumption that the autocorrelation time of the fluctuation field is much shorter than the rotational-response time of the system, and therefore, the simulation time step δt must be chosen carefully.

2.3.2 Renormalized LLG model

The model in 2.3.1 is called a single spin model, which means it describes the dynamic process of one single spin under the external field with its magnetization magnitude constant. This is a good approximation for a single grain whose diameter is about 6~10nm in traditional PMR because the environment is far low below its Curie temperature. However, when temperature is elevated, the thermal agitation would make the spin vibrate significantly so the exchange interaction among the atomic spins can hardly hold them together any longer during the switching process and magnetization magnitude

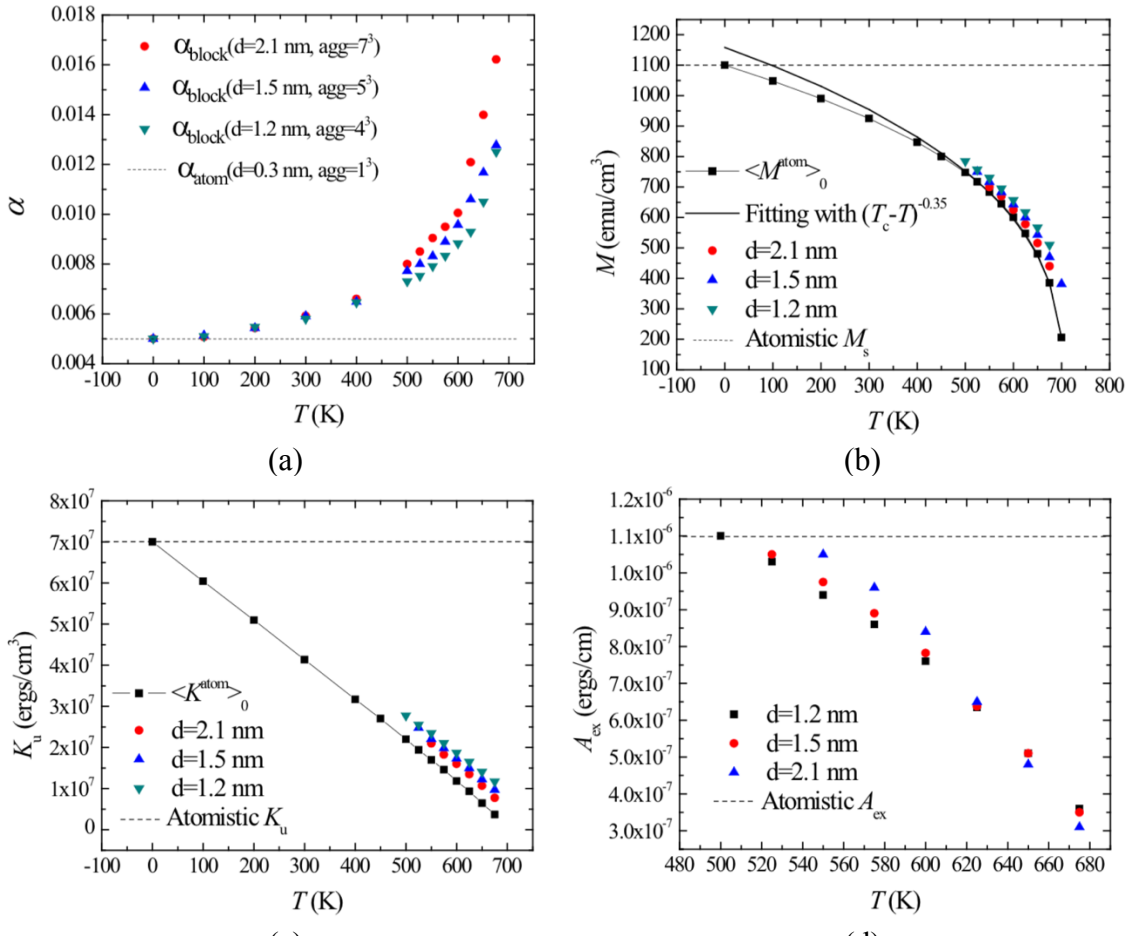
is not a constant [93]. As a result, the single spin model fails in HAMR. One solution for this problem is to introduce some correction terms to the single spin model, which is the basic idea of LLB-based model. Such method has an excellent simulation speed, like the single spin model, but still cannot predict the behavior of a grain (or a batch of spins) near Curie temperature very well. Another one is to increase the resolution and reduce the spin size in simulation. From [89], the size of one spin should be as small as $0.3\times0.3\times0.3$ nm^3 in order to provide a good prediction, which is comparable to the lattice size of FePt. As a result, this method is only suitable to simulate a small-scale system (such as a grain) otherwise it will be very time-consuming.

Renormalized LLG model is used to solve the conflicts between simulation speed and accuracy, which takes advantages of the fact that the correlation length of spin fluctuations becomes large at high temperatures near T_c from renormalization theory. The whole media can be divided into stacked cubes with a 1.5 nm or 2 nm edge length (Figure 2.8), which is smaller than the correction length of FePt. In order to mimic the behavior predicted by the atomistic model, the macroscopic parameters of "renormalized cubes" should be the same as the averaged corresponding values of the atomistic spins, such as anisotropy field and magnetization. Obviously, the microscopic parameters used in LLG model of "renormalized cubes" are functions of temperature (Figure 2.9).



(a) (b) (c)

Figure 2.8: The atomistic spins (a) are divided into renormalized spin block (b). Granular medium which is represented by Voronoi diagram are discretized using this spin block (c). [89]



(a) (b)

(c) (d)

Figure 2.9: Renormalized Parameters as functions of temperature and length scale. (a) Gilbert damping parameter; (b) Saturation magnetization (c) Anisotropy constant. (d)

Exchange stiffness parameter. d represents different length scales; “agg” is equal to the number of atomic spins in a renormalize spin. Dash lines indicate the corresponding values in atomistic (single spin) model, which are constants. [89]

2.3.3 Recording Simulation Program

Figure 2.10 shows the workflow of the recording simulation module in our program. After initialization of the program, we then calculate the equivalent field. “Hd”, “Ha”, “Hth” and “Hk” are the fields corresponding to magnetostatic energy, exchange interaction, thermal agitation and crystalline anisotropy, respectively. The LLG solver uses the 4th order Runge-Kutta method to solve the magnetization $\vec{M}$ of each renormalized cube. And “Head moving” block will change the boundary conditions of the system, i.e. move the recording head forward for a certain distance on the medium, and the program will return to the equivalent field calculation again for the next time step.

Parallel computing techniques can be used here to accelerate the simulation and each “renormalized cube” can be calculated in one thread/core of a GPGPU at the same time. If we have multiple GPUs, we can also execute the four (kernel) functions with dotted borders in Figure 2.10 parallelly [94], [95], because they have less dependence on each other. As a result, we will have a two-layer parallel program: the first layer is the parallelization of cubes as we mentioned above; the second layer is parallelization of these four different blocks. Although this method could increase simulation speed further, this two-layer framework is used only when we want to simulate a very large system in a short time once. Usually, we run one smaller simulation on each GPU: Even though these four blocks are executed sequentially, but finally we can get the several results at the same time. In addition, lookup-tables of micromagnetic parameters (M_s, K_u, A_{ex_xy} and A_{ex_z}) were

built to avoid branching statements in the GPGPU [94]. For each block, the most frequently accessed data will be copied to the shared memory first and then used by each thread (GPGPU), because the shared memory has a shorter access time but limited memory size which is insufficient for storage of these lookup-tables comparing with the global memory. The memory model of GPGPU is shown in Figure 2.11. With these methods, we can usually gain 3X~4X speed-up using GPGPU (GTX 580) compared to the same simulation on CPU (i7 6700K).

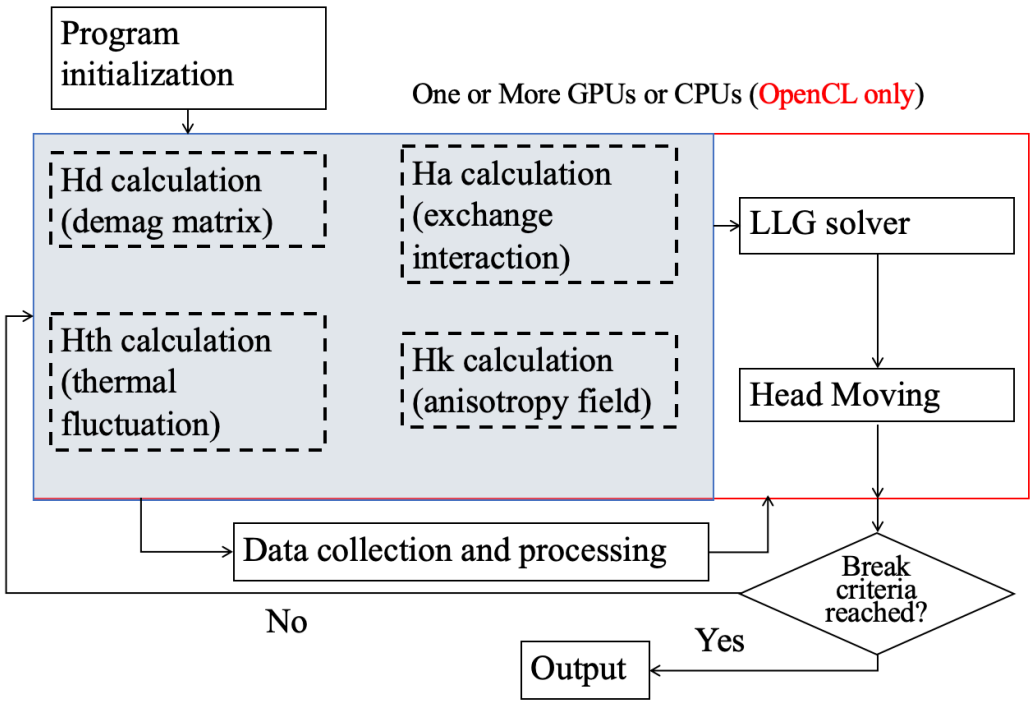


Figure 2.10: The workflow of the recording simulation module. All blocks/functions in the red box are executed on GPGPU which are also called kernel functions. And four of them in the blue box, marked with dotted borders and no arrows, can also be executed at the same time on different devices (GPGPUs or CPUs).

Another efficient way to accelerate the simulation is to only consider the "active" part and for HAMR simulation, the "active" part means the region currently being written. Although this is an approximation method, it still gives good results, i.e. only 0.2 dB difference in SNR calculation, when we use it in Chapter 6, and most importantly, we can save a lot of computing resources. More details will be shown in Chapter 6.

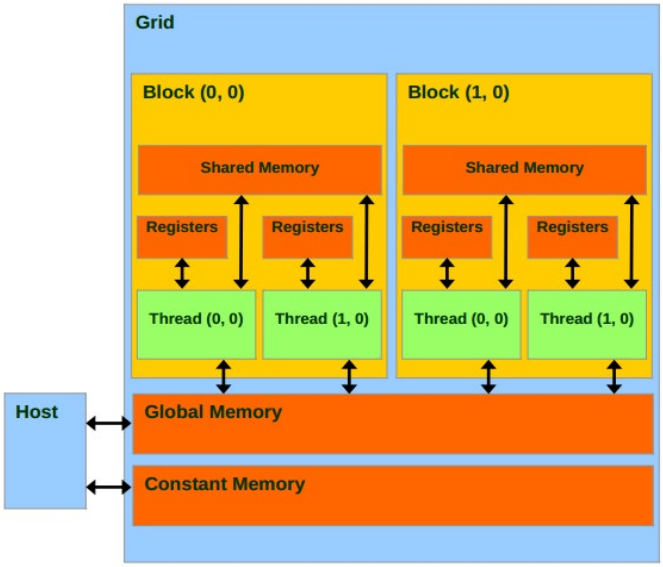


Figure 2.11: A general memory model of GPGPU. The access time for shared memory is 20 ~ 40 Cycles while for global memory it is 400 ~ 600 Cycles. The capacity of shared memory is usually tens of Kilobyte while it is several Gigabytes for global memory. [94]

2.4 Playback and State Detection

Figure 2.12 shows the schematic diagram of reader in this dissertation and the field profile. Based on the principles discussed in Chapter 1, the playback signals are an integral over all the magnetization regions when the reader is moving [25], i.e. a convolution of the writing field $\vec{h}$ and magnetization $\vec{M}$ on the media,

$$V(x) = c\int_{-\infty}^{+\infty} dx' \int_{-\delta/2}^{+\delta/2} dy' \vec{h}\,(x'+x, y') \cdot \frac{d\vec{M}(x', y')}{dx'}$$

(2-24)

where δ is the track width and c is a constant, relative to the number of wires and head moving speed for an inductive head while only relative to the physical properties in a GMR reader, such as MR ratio, largest current density etc.

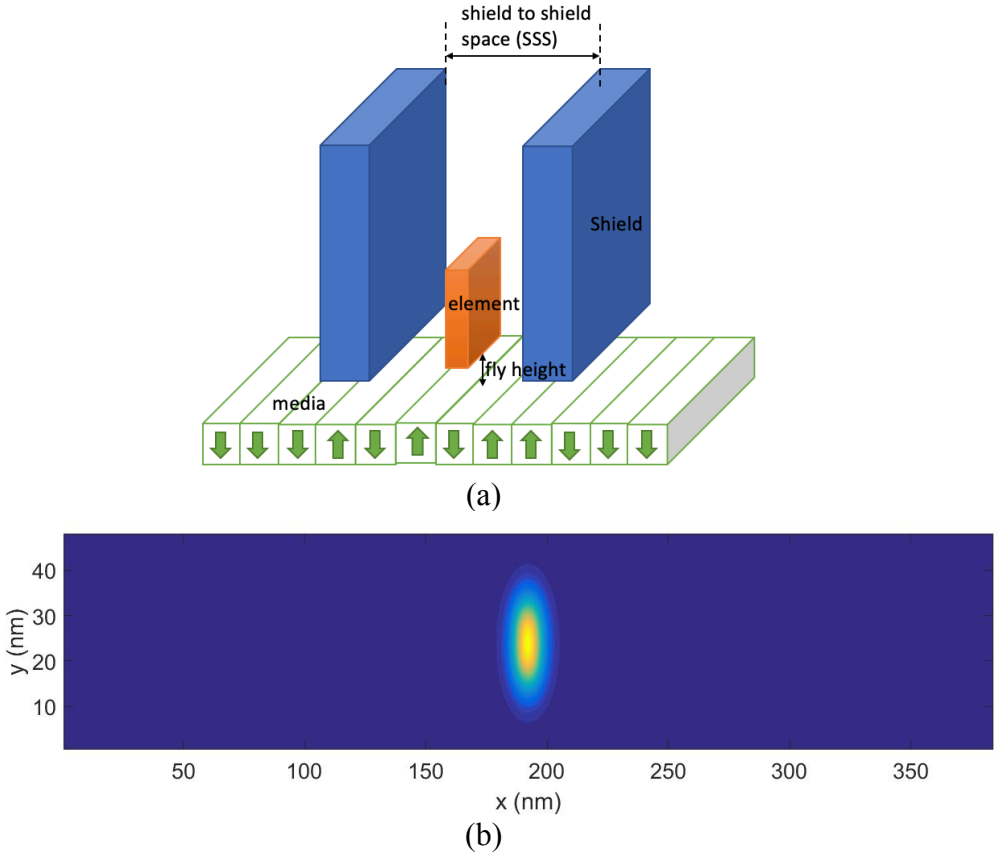


Figure 2.12: Layout of MR read head (a) and a sample the read field profile (b).

When areal density is low, each transition written on the magnetic medium results in a relatively isolated peak of voltage, and a peak detection method is used to recover

written information. However, in PMR and HAMR, the peak detection is not reliable for data detection due to the high areal density and ISI effect. So in order to solve this problem, people introduce the partial-response maximum likelihood (PRML) method, which is a combination of partial-response (PR) signaling and maximum-likelihood sequence detection (MLSD) [96]–[99]. The basic idea for partial response is to use some well-developed mathematical signal models to represent realistic playback signals (which can be referred to as a natural channel in telecommunications engineering) and these specific signals are called PR targets. In order to achieve our goal, two assumptions are used to mimic the ISI effect: 1. The shape of read-back signal from an isolated transition is exactly known and determined; 2. The superposition of signals from adjacent transitions is linear. At first, a PR Class IV (PR4) channel target [100], [101] is used,

$$y(D) = (1 - D)(1 + D)^N \tag{2-25}$$

where $N \geq 1$. For N = 1, 2 and 3, the channels are called PR4, EPR4 and E^2PR4 channels, respectively and $y(D)$ is D-transfer function,

$$D^m(a_n) = a_{n-m} \tag{2-26}$$

where a_n is the n-th bit in a data stream. So actually, PR target reflects how the current bit is affected by its neighboring bits. Around 1990s, a more general form of partial response target was proposed, using arbitrary real numbers instead of integer as indices in the PR target [99], which can be written as,

$$G = [g_0, g_1, \ \ldots, g_{L-1}]' = \sum_{l=0}^{L-1} g_l D^l$$

(2-27)

where g_l is a real number and the integer L is the length of the target response. At the same time, a finite impulse response (FIR) filter is responsible for shaping the channel to some desired, shorter partial-response target, which has a form,

$$F = [f_{-K}, f_{-K+1}, \ \ldots, f_K]' = \sum_{l=-K}^{K} f_l D^l$$

(2-28)

where f_l is the time domain coefficient. Then, the minimum mean-square error (MMSE) technique can be used to determine the parameters in G and F. A commonly used case is when $g_0 = 1$, we have

$$\varepsilon^2 = F'RF + G'AG - 2F'TG - \ 2\lambda(I'G - 1)$$

(2-29)

where λ is the Lagrange multiplier, I is an L-element column vector whose first element is 1 and all the other elements are zeros. $A = \{a_i a_j\}$, $R = \{s_i s_j\}$ and $R = \{s_i a_j\}$. s_i is the sampled signal sequence and a_i is the corresponding output sequence. Hence, G and F are actually pre-calculated in our signal processing module by a training process.

To minimize the error ε, using Lagrange method we can finally get,

$$\lambda = \frac{1}{I'(A - T'R^{-1}T)^{-1}I}$$

(2-30)

$$G = \lambda(A - T'R^{-1}T)^{-1}I$$

(2-31)

$$F = R^{-1}TG$$

(2-32)

Finally, the Viterbi algorithm is used to convert the analog signals from the equalizer to different states represented by binary bits, which is the method based on maximum likelihood estimation (MLE). At any time, this algorithm is trying to the find the most likely hidden states in a hidden Markov model (HMM) [102]. The binaries we'd like to have for output are the hidden states, the sampled signals are the observations and the state transition matrix can be known from the PR target G in last step. In addition, not every state transition is permitted due to properties of PR target. For example, if we have a 3-bit length PR target, there will be 4 possible states for this Viterbi decoder. However, because every two adjacent state will "share" one bit, there will be only 8 possible state transitions instead of 16. This is shown in Figure 2.13. Besides, soft decision is used in this decoder because the signals in HAMR are generated by a realistic physical system. In order to mimic a real-time decoder, the memory is assumed to be finite and a first-in-first-out (FIFO) queue is used, whose size is equal to the maximum path length of the Viterbi decoder. Hence, the decoding inside the queue uses a dynamic programming method, which has a time complexity of $O[(L-1)^P]$, where L is the length of target and P is the path length, which can be taken as a constant in one simulation. As the calculation window moves forward, we pick up the best values which are going to popped up from the queue. Indeed, this is a greedy and local optimization strategy. A complete example is shown in Figure 2.14.

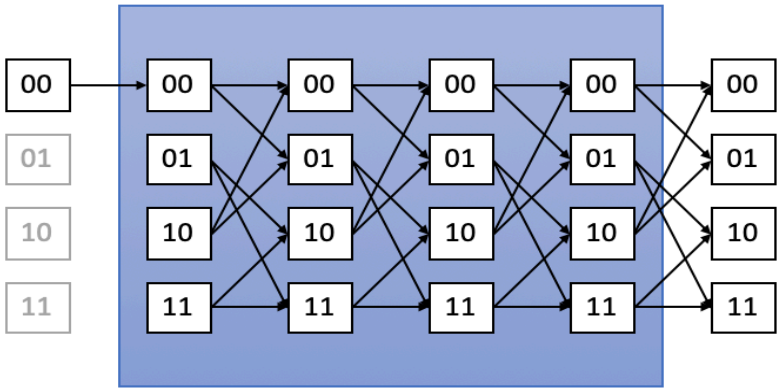
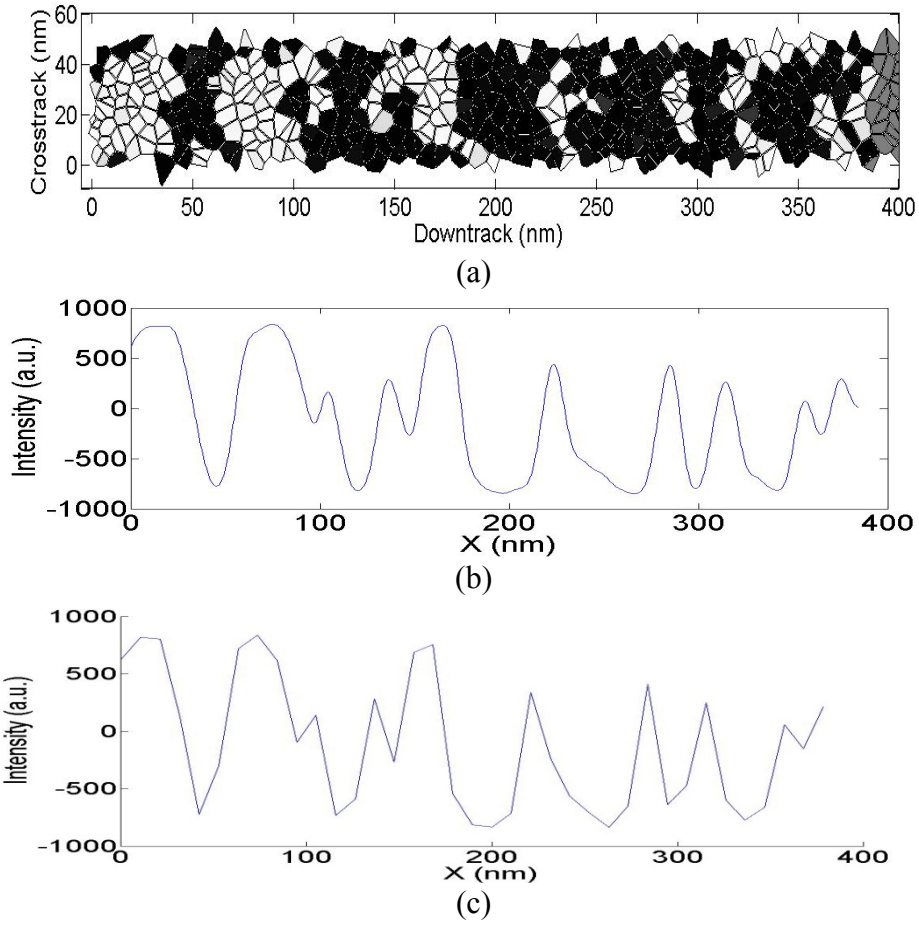


Figure 2.13: The Viterbi decoder used in our simulation, with four states and a path length of 4. Discarded states are shown in grey.



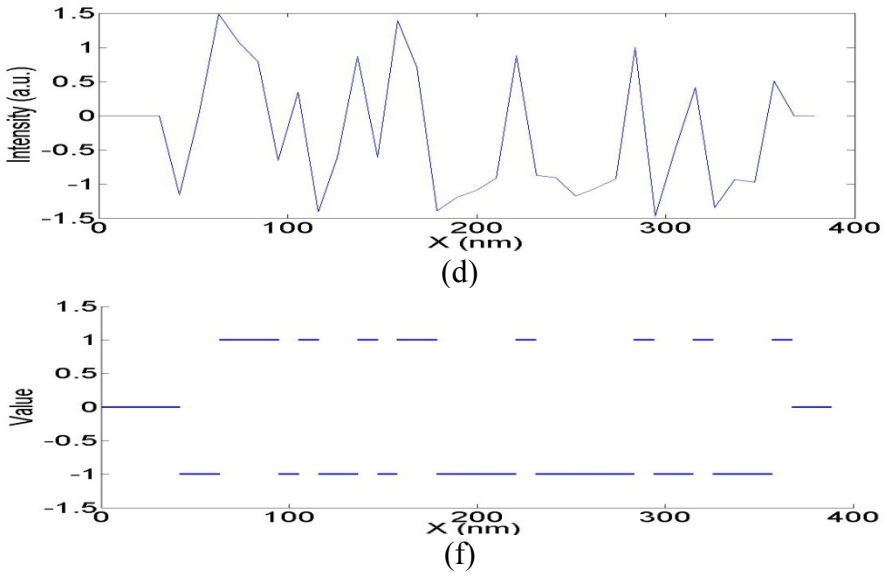


Figure 2.14: Outputs from each step. (a) recorded patterns on a Voronoi medium; (b) original playback signals with a reader; (c) sampled playback signals; (d) signals after equalization; (f) binary states $\{-1,1\}$ after decision, 0 is reserved for invalid value here.

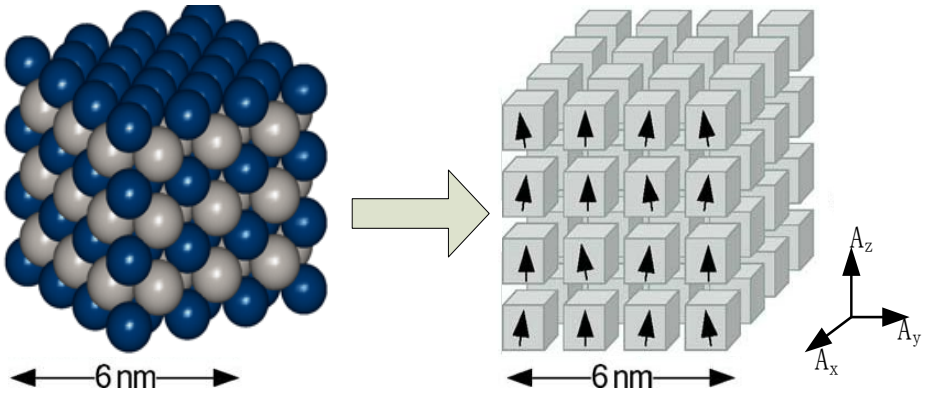
3 Renormalized Anisotropic Exchange for Representing Heat Assisted Magnetic Recording Media

3.1 Introduction

In a HAMR system, a laser guided by an NFT is used to heat the recording media. The media is heated to a temperature above the Curie Temperature and then cooled down to room temperature to finish the writing process. Obviously, this is much more complex than the current magnetic recording system and presents more problems such as thermal stability of the materials [103], thermal profile in the media [104], noise in the writing and reading process [87], etc. As a result, to solve these problems and achieve an optimal design, more simulations and theoretical development are needed.

In the last chapter, we introduced a basic renormalized LLG model for HAMR simulation. However, that model uses an isotropic exchange, but the most likely recording layer, FePt, is known to be anisotropic. In this chapter, we numerically obtain the renormalized parameters (Gilbert damping parameter, saturation magnetization, anisotropy constant and anisotropic exchange stiffness parameter (AEP)) from an anisotropic atomistic model for block dimension of 1.5nm×1.5nm×1.5nm. Then we compared this new model with the previous model at different block sizes under a given thermal pulse.

3.2 Method and Model



(a) (b)

Figure 3.1: Panel (a) shows the atomic locations in the FePt $L1_0$ that are replaced by renormalized blocks (b) in the simulation.

Figure 3.1(a) shows the $L1_0$ structure of FePt, which has Fe and Pt layers alternating along z the axis. Following the ab-initio calculations and discussion by Mryasov et al. [105], we use an effective Fe-Fe exchange constant $\tilde{J}_{ij}$ to express all exchange interactions at the atomic level. Thus, the exchange interactions can be classified in a cubic (although anisotropic) arrangement, which means that each Fe atom has 26 neighbors, and the exchange parameters between any two Fe atoms could be divided into five groups J_1, J_2, J_3, J_4 and J_5 according to the relative positions. The Hamiltonian is

$$\mathcal{H} = -\sum_{i \neq j} \tilde{J}_{ij} S_i \cdot S_j - d_i^z (S_i^z)^2 - \sum_{i \neq j} d_{ij}^z S_i^z \cdot S_j^z \tag{3-1}$$

d_i and d_{ij} are the effective single-ion and two-ion anisotropy constants, respectively. This is the object model we use to deduce our renormalized parameters.

In this model, the exchange interaction in the Fe layer (hard plane) is larger than that out of the plane (easy axis) [106], [107]. In our renormalized model, A_{xy} and A_z indicate the exchange stiffness parameters along these two directions respectively. Hence, in order to get the renormalized model shown in Figure 3.1 (b), that is, to find the exchange stiffness parameters in two directions (A_z and A_{xy}), the anisotropy energy (K_u) and the saturation magnetization (M_s), it is needed to establish four equations based on the idea that some macroscopic magnetic properties of the two systems (the renormalized system and the atomic-level one) should be maximally equivalent. In our calculation, we selected the averaged spontaneous magnetization <M>, anisotropy field <H_k> and their standard deviations δM and δH_k, respectively, because all these four macroscopic parameters could affect the magnetization process greatly. Therefore, the constraints become:

$$(\delta H_k)_{\text{atomistic}} = \left[\delta H_k\left(M_s, K_u, A_{xy}, A_z\right)\right]_{\text{renormalized}} \tag{3-2}$$

$$\langle H_k\rangle_{\text{atomistic}} = \langle H_k\left(M_s, K_u, A_{xy}, A_z\right)\rangle_{\text{renormalized}} \tag{3-3}$$

$$(\delta M)_{\text{atomistic}} = \left[\delta M\left(M_s, K_u, A_{xy}, A_z\right)\right]_{\text{renormalized}} \tag{3-4}$$

$$\langle M\rangle_{\text{atomistic}} = \langle M\left(M_s, K_u, A_{xy}, A_z\right)\rangle_{\text{renormalized}} \tag{3-5}$$

Because it is impossible to solve the four equations directly, we utilized a trial-and-error method, and constructed two cost functions that use different algorithms for the standard deviation terms δM and δH_k: they are usually much smaller than the averaged terms $\langle M \rangle$ and $\langle H_k \rangle$ except when the temperature is very to close the Curie Temperature, and a small change in δM and δH_k terms could affect the cost function more greatly than the averaged terms $\langle M \rangle$ and $\langle H_k \rangle$. Thus, in cost function 1, just the averaged terms are used as the denominators. The optimal solutions of A_z, A_{xy}, M_s and K_u are required to make cost function 1 less than 5%. However, in most instances, this does not yield one unique point but instead a connected area, so cost function 2, which is a more sensitive form to δM and δH_k, is used to find the minimum value in this area.

$$\text{costfunction1} = \left|\frac{\langle M\rangle_{\text{renorm}} - \langle M\rangle_{\text{atomistic}}}{\langle M\rangle_{\text{atomistic}}}\right| + \left|\frac{(\delta M)_{\text{renorm}} - (\delta M)_{\text{atomistic}}}{\langle M\rangle_{\text{atomistic}}}\right| + \left|\frac{\langle H_k\rangle_{\text{renorm}} - \langle H_k\rangle_{\text{atomistic}}}{\langle H_k\rangle_{\text{atomistic}}}\right| + \left|\frac{(\delta H_k)_{\text{renorm}} - (\delta H_k)_{\text{atomistic}}}{\langle H_k\rangle_{\text{atomistic}}}\right| \tag{3-6}$$

$$\text{costfunction2} = \left|\frac{\langle M\rangle_{\text{renorm}} - \langle M\rangle_{\text{atomistic}}}{\langle M\rangle_{\text{atomistic}}}\right| + \left|\frac{(\delta M)_{\text{renorm}} - (\delta M)_{\text{atomistic}}}{(\delta M)_{\text{atomistic}}}\right| + \left|\frac{\langle H_k\rangle_{\text{renorm}} - \langle H_k\rangle_{\text{atomistic}}}{\langle H_k\rangle_{\text{atomistic}}}\right| + \left|\frac{(\delta H_k)_{\text{renorm}} - (\delta H_k)_{\text{atomistic}}}{(\delta H_k)_{\text{atomistic}}}\right| \tag{3-7}$$

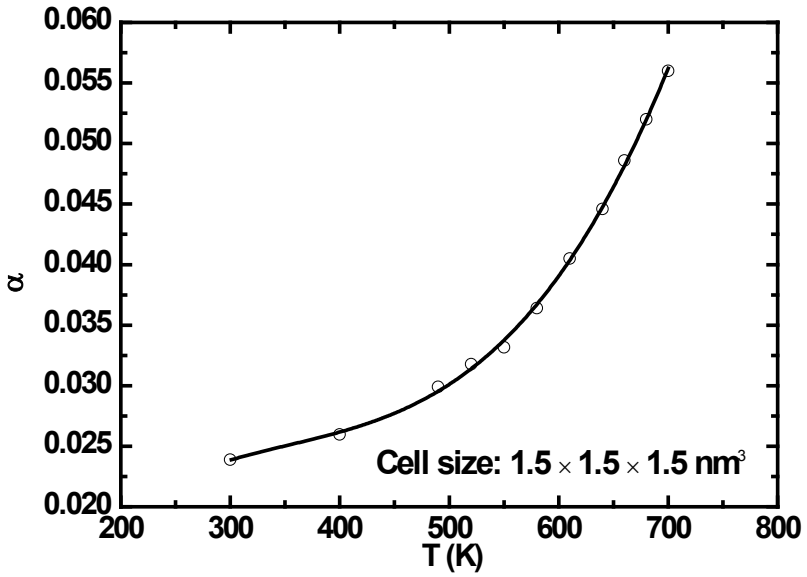
The damping constant α is calculated following [108]. Note that each block in the renormalized system is 1.5nm×1.5nm×1.5nm, which is about 5×5×4 = 100 atoms because the lattice constants for $L1_0$ FePt are a = 0.386 nm (x or y direction) and c = 0.374 nm (z direction).

Figure 3.2 (a)-(b) are the results calculated by this new method, where the dimension of the renormalized block is 1.5nm×1.5nm×1.5nm, and the results of ref. 6 are also plotted in Figure 3.2 (b)-(c) for comparison. In Figure 3.2 (a), the Gilbert Damping parameter increases rapidly with temperature owing to the increased magnon-magnon scattering [109]. Recent theoretical prediction [110], suggests that FePt damping at the atomic length scale should be 0.02, which is used for presenting both the anisotropic model and the isotropic model in the present article. In Figure 3.2 (b)(c), the results with this new method agrees well with the previous work, though there is a slight difference at the region close to the Curie Temperature. But in Figure 3.2 (d), the average AEP (the average AEP means the arithmetic average of the exchange parameters on one spin, which is defined as $^1/_3\,A_z + {}^2/_3\,A_{xy}$, and could be understood as the "equivalent" isotropic exchange parameters on this spin) changes with temperature as a power function, with exponent 0.55±0.05, when the temperature is larger than 520K. Below this temperature, the calculated average AEP deviates from the fitted curve. This is because the correlation length of the critical phenomena drops quickly when the material is cooled down from its Curie Temperature [111], or in other words, the renormalized approach become invalid at lower temperature. The ratio of AEP has a peak value at around 640-680K as shown in Figure 3.2 (d), and this could be understood as follows: because the thermal fluctuations are always trying to disturb the exchange interactions, and because the exchange interaction along the z direction (A_z) is smaller than that in the Fe plane (A_{xy}), A_z is reduced faster than A_{xy} as the temperature increases, and as a result, the ratio of A_{xy} over A_z becomes larger and larger when increasing the temperature. However, at T_c, not only the exchange

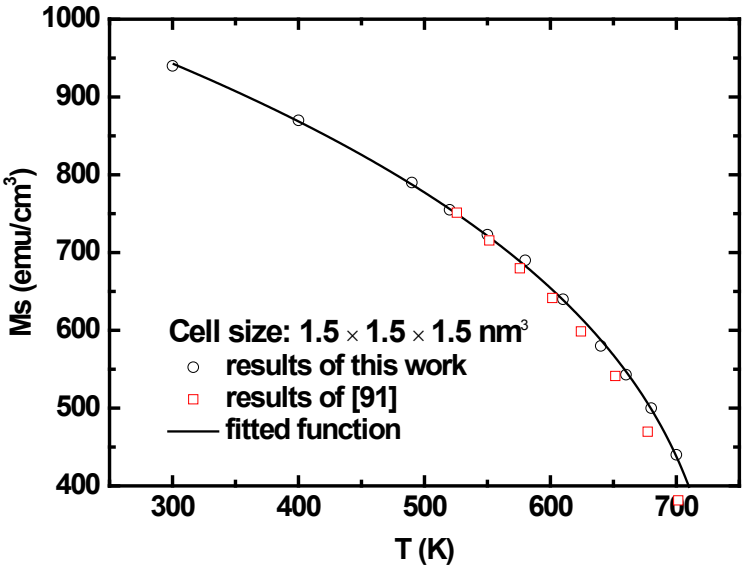
interactions but also the ratio will "disappear" because of thermal fluctuations, so there must be an abrupt drop near T_c. For low temperatures where mean field theory is valid and the grains have little chance to switch under the conditions used in this paper, we used the analytical values calculated from the continuum theory of micromagnetics for $<A_{ex}>$ and A_{xy}/A_z. The detailed fitting rules and parameters in our simulation are listed in Table 3.

Table 3 Temperature categories and corresponding fitting rules of renormalized material parameters

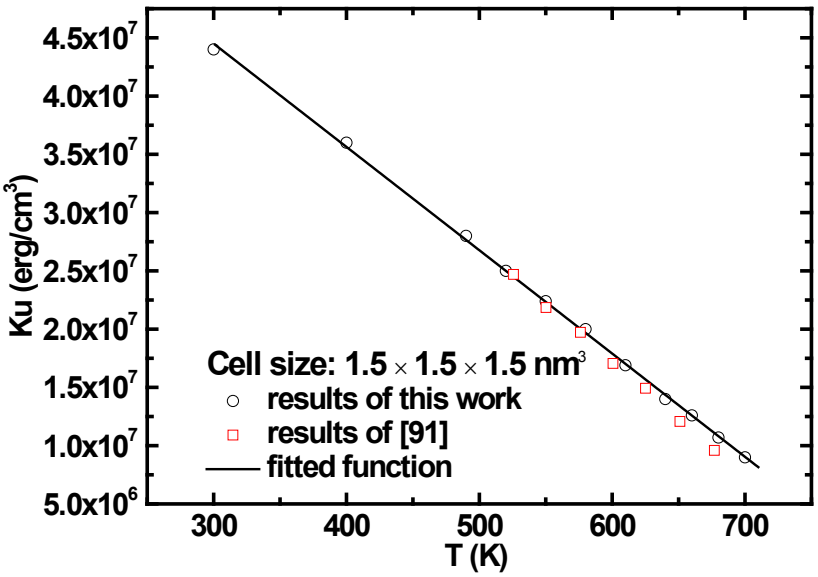
T (K)	α_{block}	M_{block}	K_{block}	$<A_{block}>$	Axy/Az
T<520	Fitted	Fitted	Fitted	$\langle A_{ex}\rangle \left(\frac{\langle M\rangle_{renorm}}{M_s}\right)^{1.5}$	1.3
520<T<705	Fitted	Fitted	Fitted	Fitted	Fitted
T≥705	0.1	30	0	0	1.0



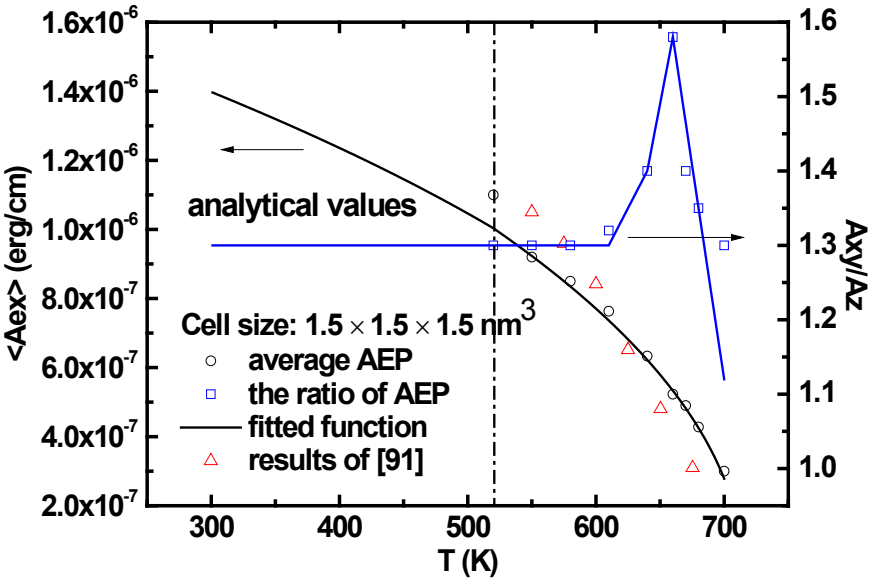
(a)


1000
900
800
700
600
500
400
Ms (emu/cm3)
300
400
500
600
700
T (K)
Cell size: 1.5 × 1.5 × 1.5 nm3
results of this work
results of [91]
fitted function

(b)


4.5x10^7
4.0x10^7
3.5x10^7
3.0x10^7
2.5x10^7
2.0x10^7
1.5x10^7
1.0x10^7
5.0x10^6
Ku (erg/cm3)
300
400
500
600
700
T (K)
Cell size: 1.5 × 1.5 × 1.5 nm3
results of this work
results of [91]
fitted function

(c)



(d)

Figure 3.2: Renormalized parameters as functions of temperature. Circles are the results using this new method, squares are the result from [89], and solid lines show the fitting functions.

3.3 Results and Discussions

A heat pulse was used to test the difference between different models. It is defined as

$$T = 300K + \Delta T * g \tag{3-8}$$

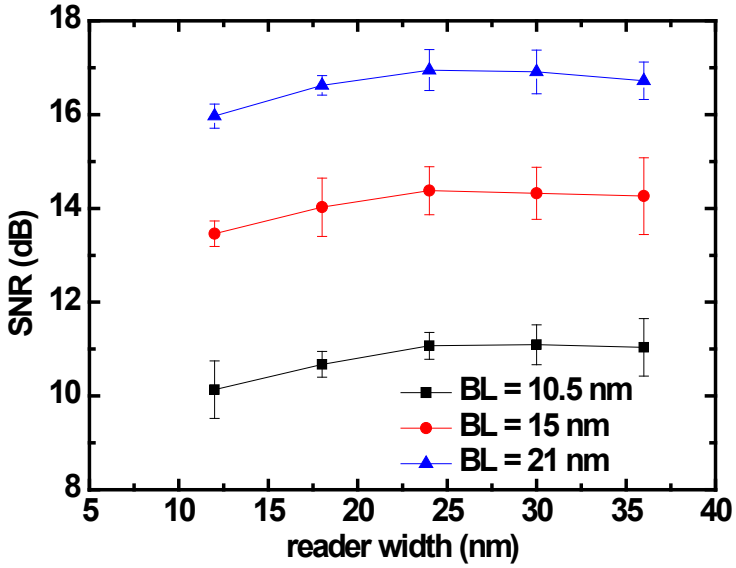
$$g = \begin{cases} 1 - e^{-t/\tau} (t < T_P) \\ e^{-(t-T_P)/\tau} (t \geq T_P) \end{cases} \tag{3-9}$$

T_P is 1.269ns and τ is 0.0551197ns. The initial and final temperature is 300K, and the peak value T_{max} is defined as 300K + ΔT. ΔT could equal 400K, 550K and 600K in the simulation.

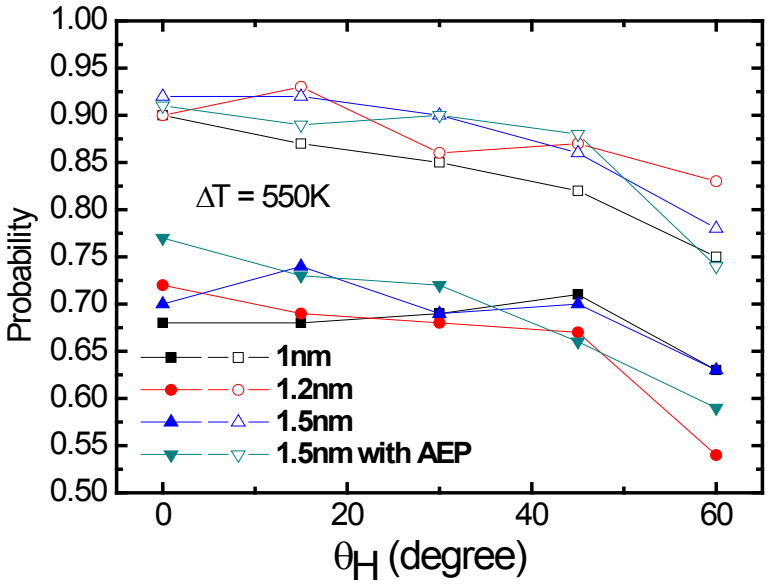
Figure 3.3 indicates the switching probability for a single grain under different conditions with different models——anisotropic exchange model (block size is 1.5nm) and isotropic exchange models (block size is 1.0nm 1.2nm and 1.5nm respectively). θ_H is the angle between the field and z-axis, and its value could be 0°,15°, 30°, 45° or 60°. Each case was repeated 100 times. After comparing these three figures, it is found that: 1) For ΔT = 400K, the switching probability using 1.5nm block size and isotropic exchange parameters is significantly smaller than the probabilities generated by the other three models, which may be a distortion due to larger block size. 2) For ΔT = 550K or 600K, there seems to be not much difference of the switching rate among different models: this is possibly because much less time is spent near the Curie temperature where results are known to be model sensitive.

To have a better understanding of point 1), we plotted the average M_z against time for each case when ΔT = 400K. Here "average" M_z includes only the successful switching processes. For example, for the anisotropic exchange model when H = 5000 Oe and θ_H is 0, the switching probability was 0.96 and the average M_z was calculated from the 96 cases which switched. Again, we find that results using 1.5 nm anisotropic exchange (Figure 3.4 (a)) are similar to the results using 1.0 nm isotropic exchange (Figure 3.4 (b)). In Figure 3.4 (c), we find that as the block size decreases, the average M_z curves are approaching the values calculated by the anisotropic exchange model. This means that the new model could

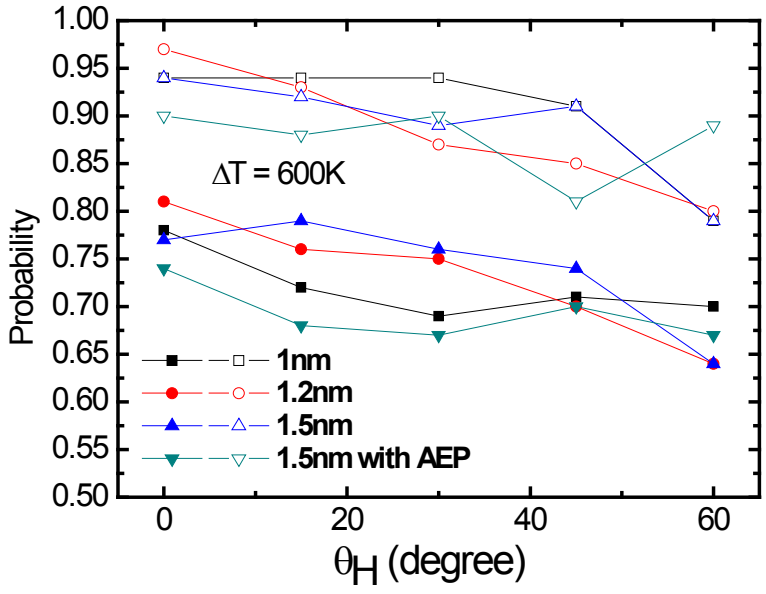
improve the simulation accuracy only at the cost of introducing one parameter, and therefore won't affect the simulation speed seriously.



(a)



(b)

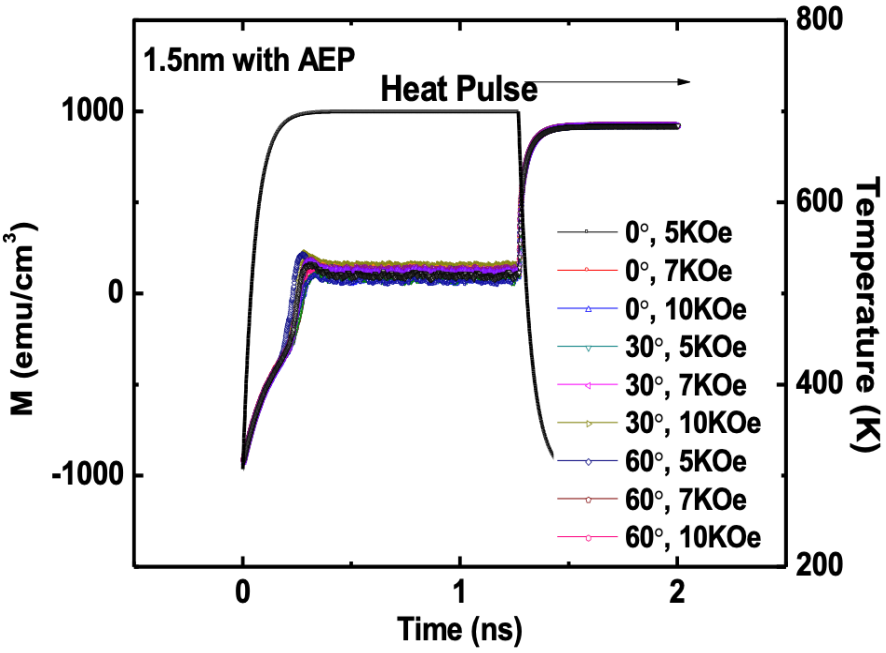


(c)

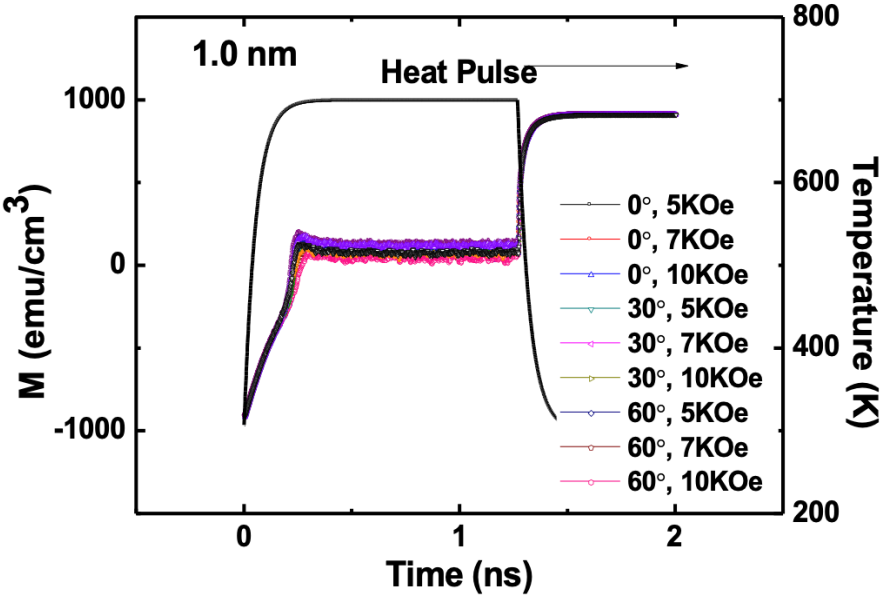
Figure 3.3: Switching probability for different discretization and exchange, solid symbols are the results with a field of 5 kOe and open ones are 10 kOe.

3.4 Conclusion

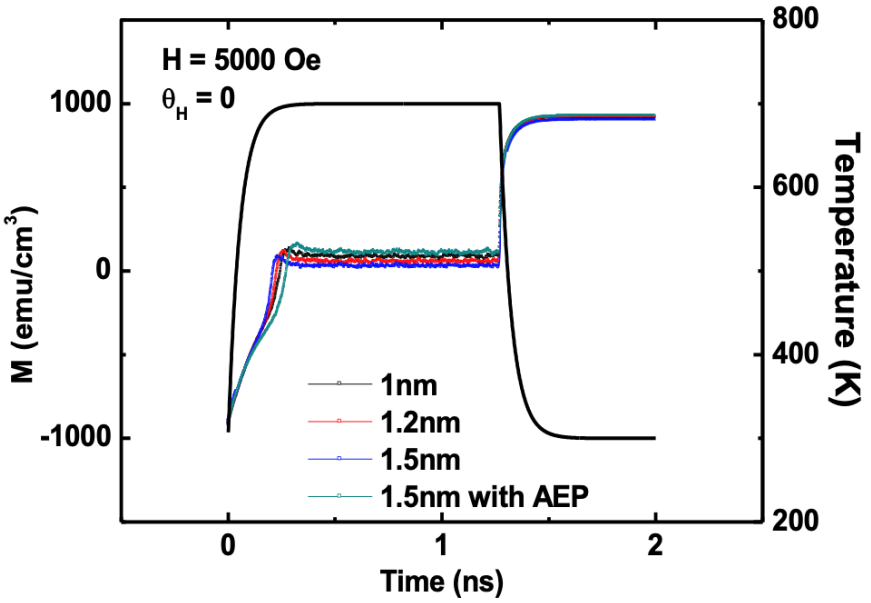
A new set of renormalized media parameters that includes anisotropic exchange was developed for a range of different temperatures using the renormalized block size 1.5nm×1.5nm×1.5nm. These parameters yielded the same switching behavior as smaller discretization using isotropic parameters, and thus make accurate calculation using renormalized media blocks more efficient.


1.5nm with AEP
Heat Pulse
1000
0
-1000
M (emu/cm3)
800
600
400
200
Temperature (K)
0
1
2
Time (ns)
0°, 5KOe
0°, 7KOe
0°, 10KOe
30°, 5KOe
30°, 7KOe
30°, 10KOe
60°, 5KOe
60°, 7KOe
60°, 10KOe

(a)


1.0 nm
Heat Pulse
1000
0
-1000
M (emu/cm3)
800
600
400
200
Temperature (K)
0
1
2
Time (ns)
0°, 5KOe
0°, 7KOe
0°, 10KOe
30°, 5KOe
30°, 7KOe
30°, 10KOe
60°, 5KOe
60°, 7KOe
60°, 10KOe

(b)



(c)

Figure 3.4: Average M_z (magnetization along z direction) variation with time under a given heat pulse. In 4(a) the AEP is considered with block size of 1.5×1.5×1.5 nm^3, and in (b) only isotropic exchange parameters are considered with block size of 1.0×1.0×1.0 nm^3. In (c) the different models are compared for an applied field H = 5 kOe and θ_H (the angle between the z axis and H) of 0°to show the differences in M_z during the switching process.

4 A Study of SNR and BER in Heat Assisted Magnetic Recording

4.1 Introduction

As the areal density increases in HAMR, it will bring many challenges to the recording system, such as increased non-linear bit shift and transition noise, and could lower the SNR. It is known that for non-patterned thin-film media, a small deviation of SNR may lead to a significant change of BER, and consequently the effective user bits per unit area or user areal density capability.

In order to have a full understanding of the HAMR system, it is important to calculate the SNR. Most published work is based on a uniform recording pattern ("1010101…"), but this pattern is not enough to characterize the actual recorded pattern, which is more like a pseudo-random binary sequence (PRBS). Furthermore, detection of a PRBS allows bit error rate to be calculated and, ultimately, the user density to be estimated via the Shannon Capacity Limit or actual codes. In this chapter, we focus on the factors that could affect the SNR and hence BER of the HAMR system: we try to determine which ones are most important for small bit length (around 10 nm). We then study SNR vs BER by processing the playback signal with a 1D MMSE equalizer and Viterbi detector implemented in Section 2.4. Finally, we predict the BER and then the recording capacity for a range of bit lengths.

4.2 Simulation Models and Methods

4.2.1 Recording Media and Writing Head

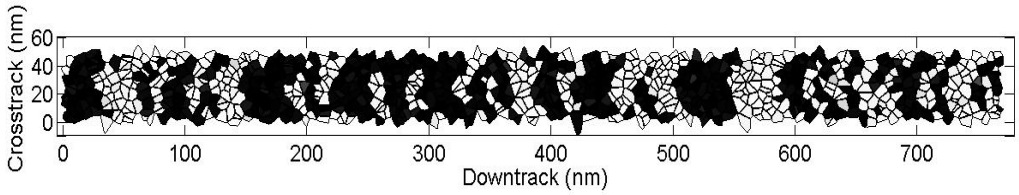


Figure 4.1: A recorded medium sample with PRBS pattern (11100010110100111101011101101110100110111001001111000011010101011010010 1). Bit length = 10.5 nm.

In this work, we used 6-nm-thick and 48-nm-wide recording media. The average and standard deviation of the grain pitch are 6.4 nm and 22% with 1 nm nonmagnetic grain boundaries. The standard deviation of anisotropy field and exchange are set as 15% and 0% respectively, if it is not specified in the simulation. Low temperature damping is taken to be 0.02. Additionally, the Curie temperature of the media is about 705K and the saturation magnetization M_s is 1100 emu/cm^3. We employ anisotropic exchange parameters that we have previous demonstrated to accurately describe the media even at the block length of 1.5 nm.

A quasi-Gaussian thermal profile was used for the writer, with a full-width half-maximum of 50 nm. Assuming the downtrack direction is along the x axis and crosstrack is along the y direction, the temperature distribution on the surface of the media reads,

$$T(x,y,t) = 300 + (T_{max} - 300)exp\left[\frac{-(x-vt)^2 - (y-W_0)^2}{2\sigma^2}\right]$$

(4-1)

where $\sigma = FWHM/2\sqrt{2ln2}$ and W_0 is the track center. T_{max} is the peak value of the heat spot and v is the velocity of the recording head, which is set as 15 m/s here. The canting angle of the head field with respect to the z-axis is 0 and we change the switching frequency of the field to get patterns with different bit lengths (BL). A sample recorded pattern is shown in Figure 4.1.

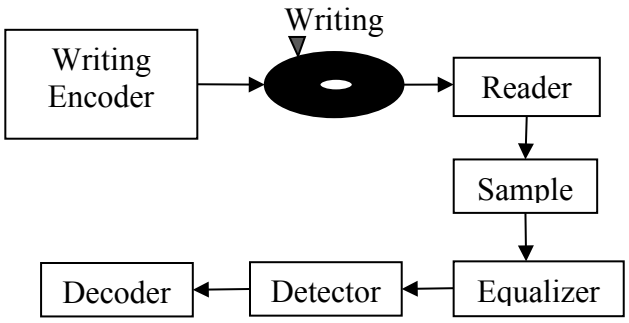


Figure 4.2: System diagram of our simulation program.

4.2.2 Playback and Detection System Architecture

A basic structure of the HAMR system is illustrated in Figure 4.2. It is known that the data storage system behaves similarly to a communication system, and the writing and reading are much like a process where we send and receive a message via a channel. Consequently, at least five steps are needed in order to obtain the correct information written on the media,

The first step is playback by a conventional reader with flying height of 6 nm based on the reciprocity principle. The simulations for each case were repeated at least 8 times with a pseudo-random binary sequence (PRBS) pattern, and the SNR is

$$SNR = \frac{\int \overline{V(x)}^2 dx}{\int \overline{\delta V(x)^2} dx} \tag{4-2}$$

where $V(x)$ is the playback signal voltage along x direction.

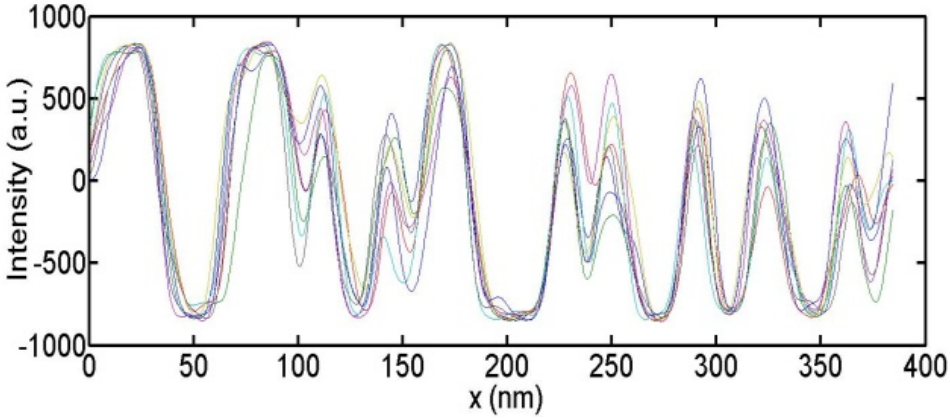
The playback signal is then sampled and processed with a 1-D MMSE equalizer and a Viterbi detector to obtain the BER. We select the sampling points carefully, with a sampling period equal to the minimum bit length, because there is some latency in the writing process owing to the switching behavior of the magnetic grains in HAMR: this allows the sampling points to better represent the original signal. The function of the equalizer is to make the output signal closer to the desired one, that is, to find the target function $G = [g_0, g_1, \dots, g_{L-1}]^T$ and the equalizer function $F = [f_{-K}, f_{-K+1}, \dots, f_0, \dots, f_K]^T$ by minimizing the difference between the equalized signal and the ideal signal, where L is the length of the target response and $2K + 1$ is the length of the equalizer.[3] In this chapter, we used $L = 2$, $g_0 = 1$ and $2K + 1 = 7$ bits for a media with a bit length (BL) of 10.5 nm or $2K + 1 = 3$ bits for that of 21 nm. The purpose of using these parameters is to keep the total equalizer length unchanged (around 60-70 nm) for different BL. The last step is to determine the states ("0" or "1") of the signals with the Viterbi detector and decoder [112], [113]. These predicted results are compared with the original written bits to calculate the BER.

In order to get an optimal target and equalizer function G and F, a training process is needed: the training dataset is the same one that we used to calculate the SNR. It is important to recognize that the optimization of the function G and F and hence the BER obtained by this method is limited by the small amounts of data available.

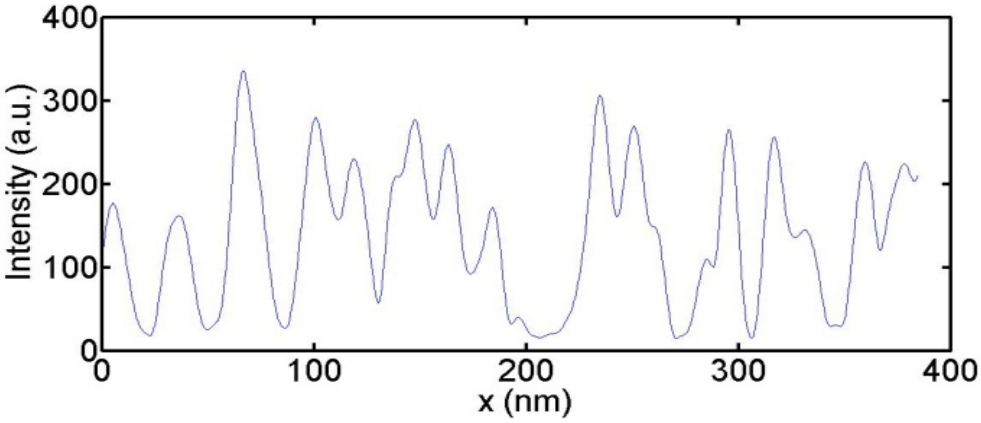
4.3 Results and Discussion

4.3.1 Noise

Figure 4.3 shows the playback signals for eight runs when BL =10.5nm. It is found that the noise mainly comes from the transition noise, that is, the noise becomes larger near the transition centers, and it is also clearly seen that the signal decreases for very small bits. For example, the SNR performance of "1010" is much worse than "1100". This is because, as the BL decreases, the jitter becomes very sensitive to the distribution of the grains; a very small deviation of the grain position near the transition center could cause relatively large magnetization fluctuation.



(a)

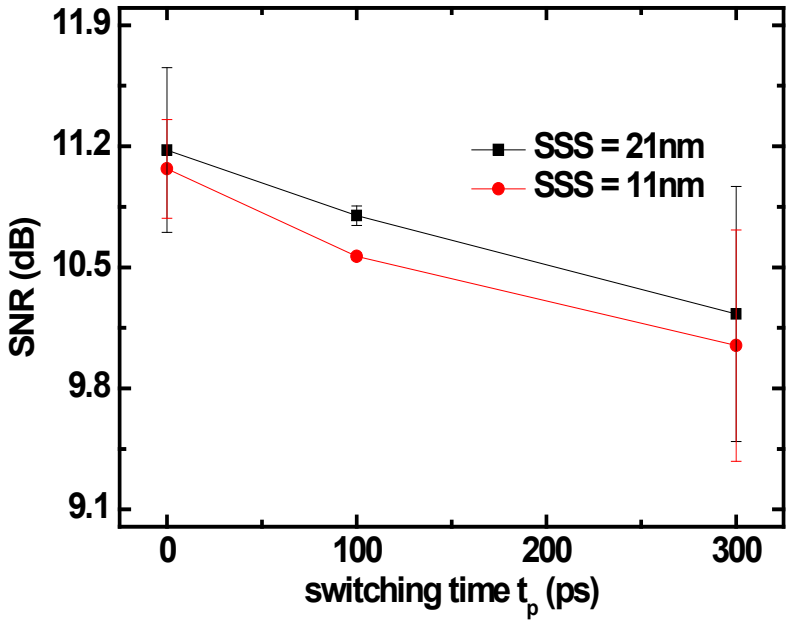


(b)

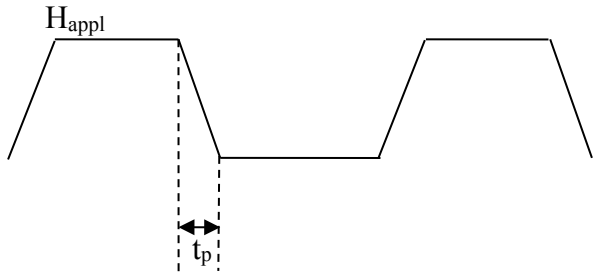
Figure 4.3: Readback signals (a) and noise (b) for one case with 8 runs. H = 8kOe, T_{max} = 850K, and PRBS is used.

4.3.2 Switching Time

Figure 4.4 indicates that the SNR drops significantly when we increase the switching time of the field produced by the head. This is very reasonable because increasing the rising or falling time of the field means increasing the transition area and the transition noise. As a result, the SNR performance becomes worse. We could also see in this figure that the read head with larger SSS performs better than that with smaller SSS.



(a)



(b)

Figure 4.4: SNR curves against the switching time t_p of the writing field (a). The definition of t_p is illustrated in (b). H = 8kOe, BL = 10.5nm, and T_{max} = 850.

4.3.3 Peak Temperature of the Heat Spot

In Figure 4.5, it is found that there exists a best SNR value for peak temperature of the heat spot: the optimum occurs around 850K. This result is consistent with [89] for jitter that also showed the optimal peak temperature to be 150 K above the Curie temperature. It is a consequence of optimizing the thermal gradient: cooling too fast (high gradient) freezes in superparamagnetic fluctuations, but insufficient gradient cannot overcome the switching field distribution. We again find that the header with larger SSS could produce better SNR.

4.3.4 Standard Deviation of Anisotropy Field and Exchange

Comparing Figure 4.6 (a) and (b), we find that when BL is larger, the SNR drops a little more quickly when the standard deviation of H_k or A_{ex} increases. (Curie temperature is largely proportional to exchange, but also depends on the effect of grain size that is intrinsic to our renormalization approach.) In the case of small BL, this effect is weakened, which means that BL is a limitation of SNR for the media recorded at higher areal density. SNR performance from smaller BL is less sensitive to SSS because the frequency content

of the noise and signal are more similar. The dependence on media parameters is qualitatively similar to Zhu et al. [114], but quantitative comparison is difficult owing to their use of uniform recording patterns, i.e., not a PRBS, and lack of a read head.

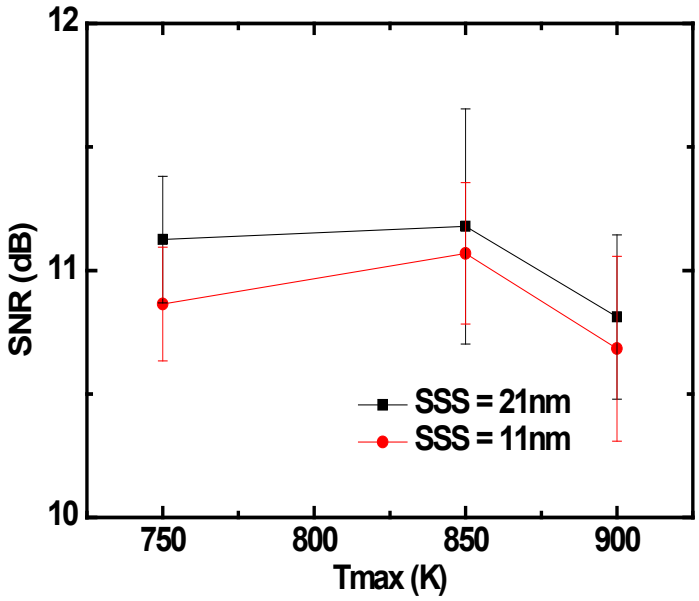
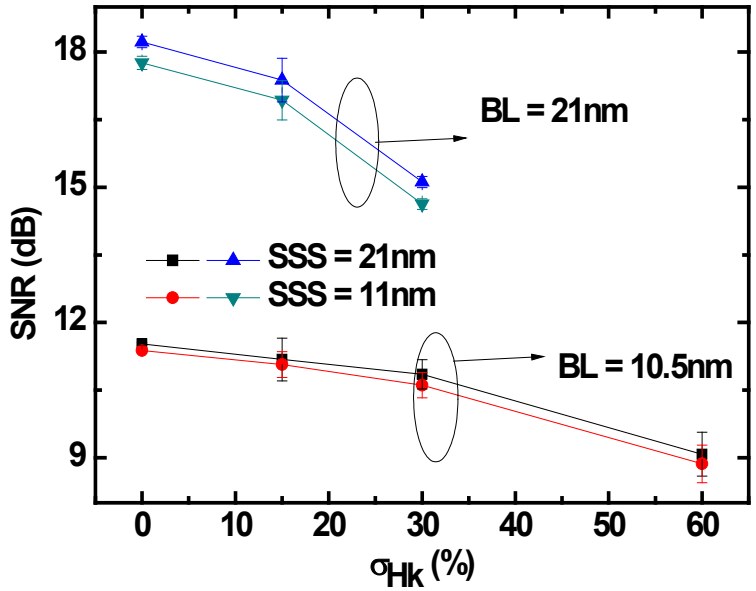
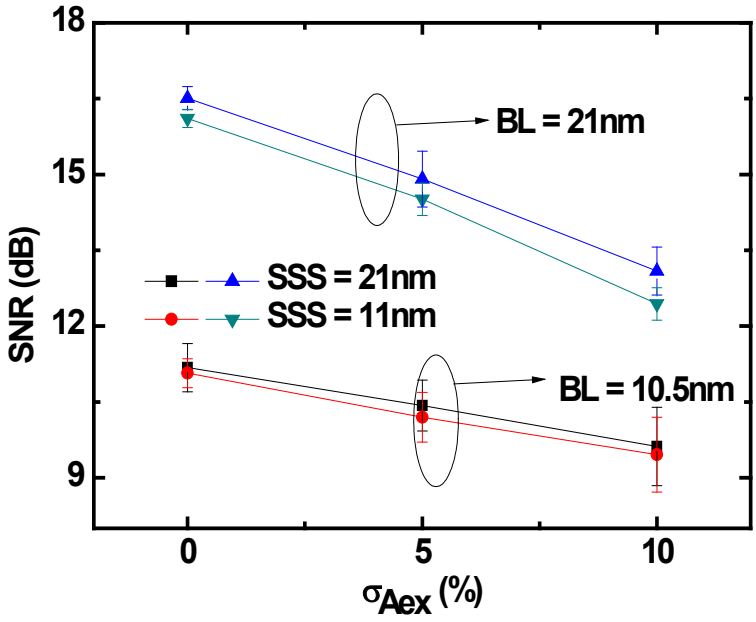


Figure 4.5: SNR curves against the maximum temperature of the heat spot T_{max}. H = 8kOe, switching time of field t_p = 0ps, and BL = 10.5nm.



(a)



(b)

Figure 4.6: SNR curves against the standard deviation of anisotropy field H_k (a) and average exchange parameter A_{ex} (b). H = 8kOe, $\theta_H = 0$, switching time of field 0ps, BL = 10.5nm, and T_{max} = 850K.

4.3.5 Reader Width

Figure 4.7 shows the relationship between the reader width and SNR with different BLs. The optimal reader width is around 24 nm. When the reader width is smaller than 24 nm, it drops very quickly as the reader width decreases. On the other hand, when reader width is larger than 24 nm, SNR is lowered only a little bit if we continue increasing the reader width.

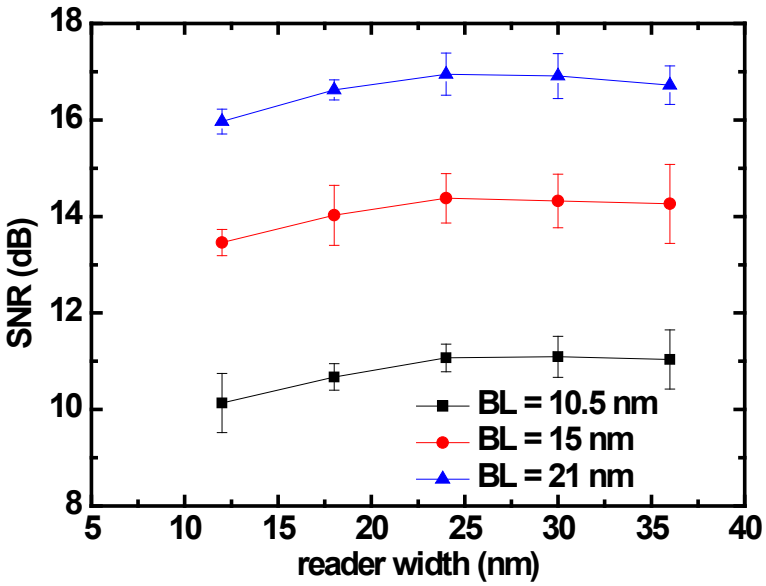


Figure 4.7: SNR curves obtained with different width of readers. H = 8kOe, switching time of field 0ps, BL = 10.5nm, and T_{max} = 850K.

We have computed the BER for a subset of the cases listed above that span the SNR range from 8 to 14 dB. (See Figure 4.8.) We fit this to the well-known BER-SNR relationship [115], [116] used in digital transmission theory,

$$BER = a \times erfc\left(b \times \sqrt{\frac{P_{signal}}{P_{noise}}}\right)$$

(4-3)

where P_{signal} and P_{noise} are the average power of a series of signals and their noise, a = 0.4325 and b = 0.4013. This relationship allows the computationally intensive BER to be obtained from the less demanding SNR. We can then easily obtain the effective user density or capacity [117] from the following equation,

$$C = 1 + BER \log_2(BER) + (1 - BER)\, log_2(1 - BER)$$

(4-4)

where C is the effective user bit ratio defined as the total user bits over the total bits. If this value is divided by BL, we could get the effective bit ratio as an inverse length. Several computed results with different BLs are listed in Table 4. For a HAMR system with the default values mentioned in Section II, the calculation shows that BL = 10.5 nm is the best choice to get the largest storage capacity. This result corresponds to the idea that the shortest length bit should be longer than the largest diameter grain likely to be encountered, i.e. one standard deviation larger than the average diameter of 6.5 nm.

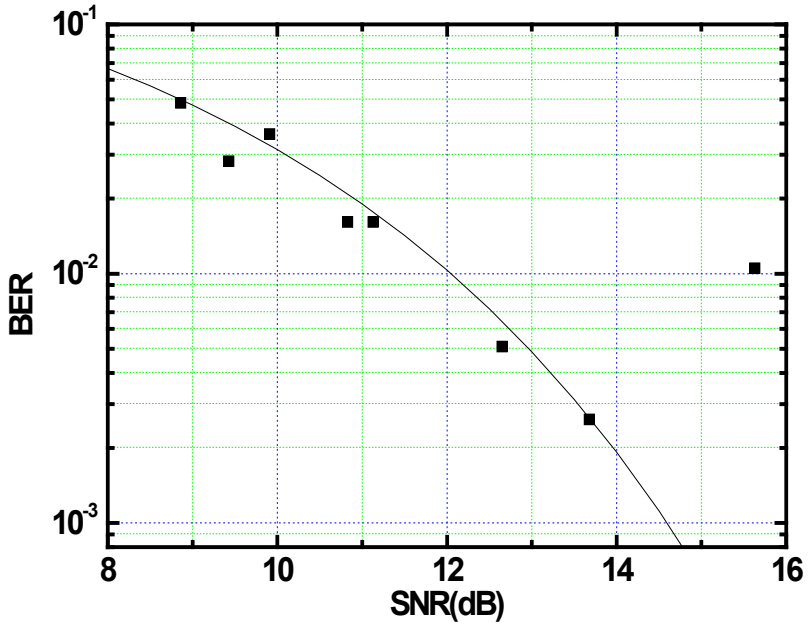


Figure 4.8: The relationship between SNR and BER. The black line is the fitting function.

4.4 Conclusion

In this chapter, we present an integrated HAMR simulation system that includes writing, playback and decoding modules. Although only a simplified reader and media design were employed, the results should be valuable for estimating general trends in optimal system design. Bit length, for small values, is found to be the single most important determinant of SNR owing to the importance of jitter and non-linear-transition-shift. Peak temperature of the heat spot is found to optimize to about 850 K and the best reader width is determined to be about ½ the Full-Width Half Maximum of the heat spot. An SNR vs BER curve is obtained that allows the reader to compute the BER conveniently using SNR calculated from a limited dataset. With this curve we also predicted the possible storage

capacity with different bit lengths and found that the optimal recording bit length is around 10 nm for grain pitch (GP) = 6.4 nm.

Table 4 Calculated results with equation (4-3) and (4-4)

BL (nm)	SNR (error)	Calculated BER	Effective bit ratio/nm
4.5	2.4 (±0.1)	0.20	0.062
6.0	4.7 (±0.3)	0.14	0.068
7.5	6.0 (±0.2)	0.11	0.067
10.5	11.1 (±0.3)	0.018	0.083
15	13.7(±0.3)	0.0026	0.065

5 Dependence of Predicted HAMR Areal Density on Common Optimization Strategies

5.1 Introduction

It is an open question how much user areal density (UAD) can be obtained in HAMR with current or future technology. It is also an interesting question how we could improve current technology to achieve this aim. Some results [118], [119]have already been reported on this topic but parts of their assumptions are too idealized to realize in the near future, such as very small grain size or narrow NFT. Therefore, in this work, we first performed micromagnetic simulation with the renormalized-LLG method to optimize the system based on achievable technologies. The BER is calculated using a Viterbi detector and UAD is estimated using the Shannon capacity limit following the procedure described in last chapter. We illustrate how to improve the performance of HAMR with low cost and high gain for products in the near future and, in order to determine the ultimate UAD, we simulate perfect recording[120].

5.2 HAMR Optimization with Micromagnetic Simulation

In this section, we discuss methods to increase the areal density of HAMR with micromagnetic simulation. Related issues, such as the distance between NFT and writing pole and switching time of the head, are also included in our analysis. Gaussian thermal profiles and uniform writing field are used if not specified. Some default simulation parameters are listed in Table 5.

Table 5 Important parameters in micromagnetic simulation

Parameters	Value
GP	6.4±20% nm
T_{peak}	850 K
Head velocity	10 m/s
Media thickness	6 nm
M_s (0 K, media)	1100 emu/cm^3
K_u (0 K, media)	7×10^7 erg/cm^3
T_c (media)	700 K
δT_c (media)	2%
δH_k (media)	10%
Discretization size	1.5 nm

5.2.1 Optimization of NFT and reader size

Given the specified media, the simplest idea (but maybe not the simplest way) to obtain larger UAD is to optimize the writer and the reader. [121] have already shown that

higher thermal gradient can result in larger media SNR. Hence, we shrink the FWHM of the thermal spot from 50nm to 30nm with different reader widths for playback. The simulation results are shown in Figure 5.1: we can see that, as we decrease FWHM, it is possible to get a better SNR with reader width equal about one half of the FWHM.

The thermal spot with a FWHM of 30 nm decreases optimal track width by about 60% but increases media SNR more than 2 dB relative to a 50 nm spot, so the capacity with 30nm thermal spot should improve by at least 60%.

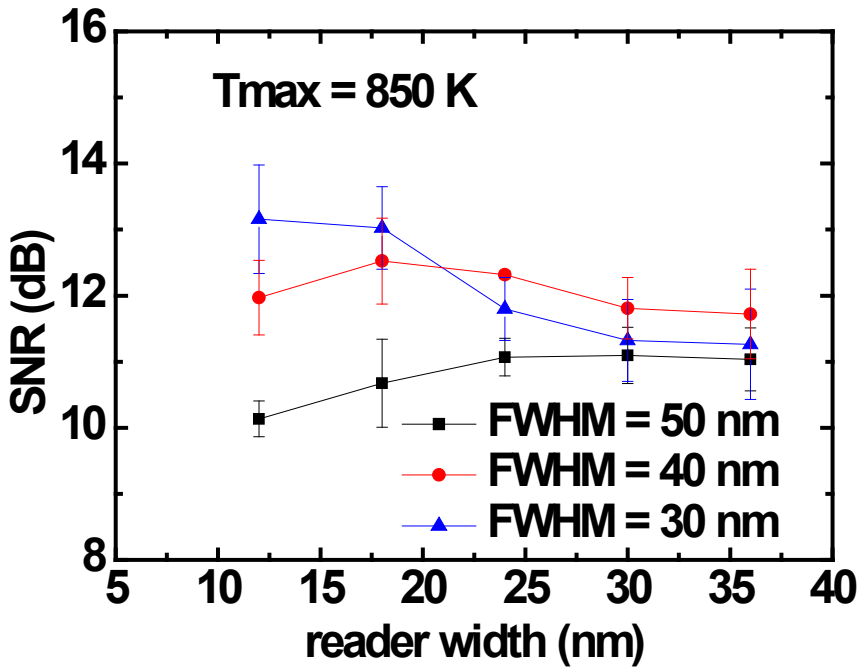


Figure 5.1: Media SNR with different reader width. 30 nm FWHM gives the best SNR performance.

5.2.2 Effects of Grain Size

In Figure 5.2, the effects of grain size are studied. When average GP is comparable to bit length (BL), both GP and its deviation have very significant effects on SNR (blue

and green curves in Figure 5.2) and hence the capacity. However, when BL is very large (10.5nm, about two times GP in this work), the "grain size effect" is weakened, as the red and black curves show in this figure. Although the calculated data always shows that smaller grain size generates larger media SNR, smaller grain size has lower thermal stability and that is why the slope of the curves become gentler for the smaller grain pitches. In this work, we only calculate the cases with relatively large BL, i.e., 6 nm, so that the larger grain pitches can still produce tolerable results.

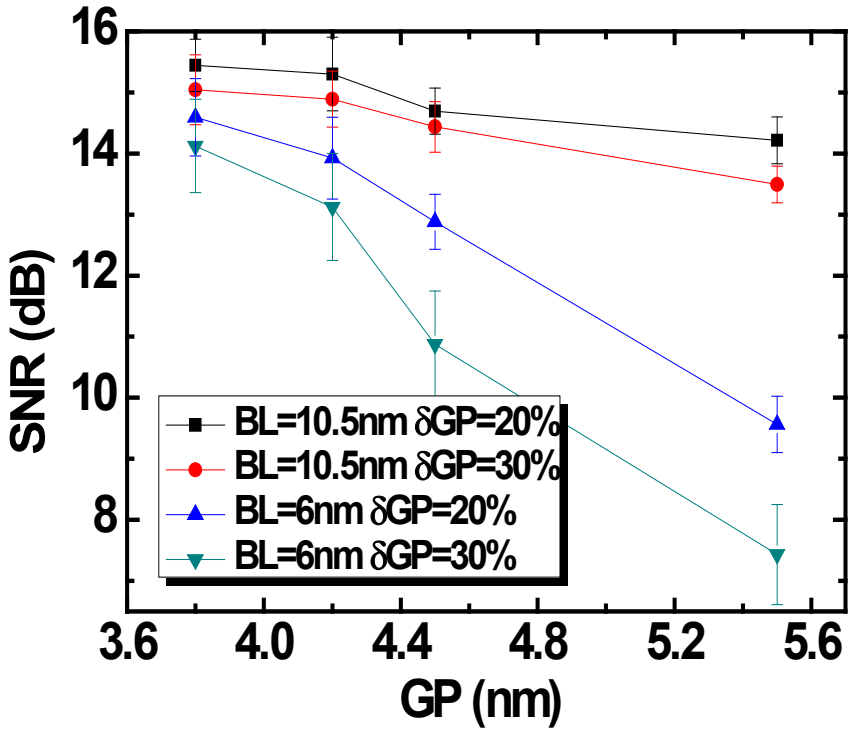


Figure 5.2: Media SNR with different BL, GP and its standard deviation. FWHM = 50nm.

5.2.3 Distance between NFT and Writing Pole

We tried different distances between NFT and a 50nm × 50nm write pole and found that there is an effective distance (the distance is measured between the centers of the NFT and the writing pole) within which we could have a very good performance. The effective distance for a 50nm FWHM thermal spot is about 30nm, but it is only 15nm for a 30nm thermal spot (Figure 5.3). This means a smaller thermal spot will lead to a smaller effective distance which might decrease the thermal reliability of system and increase the fabrication cost and difficulty.

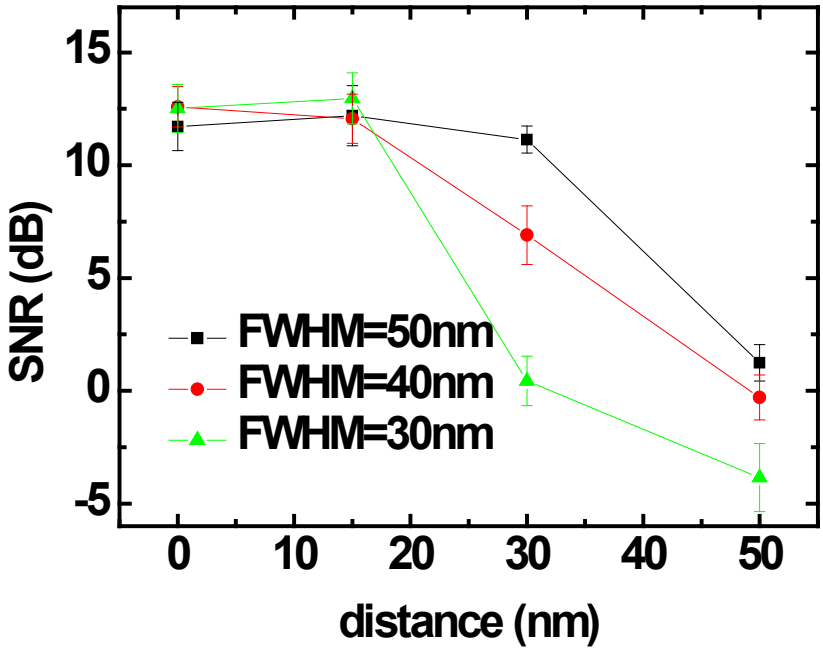
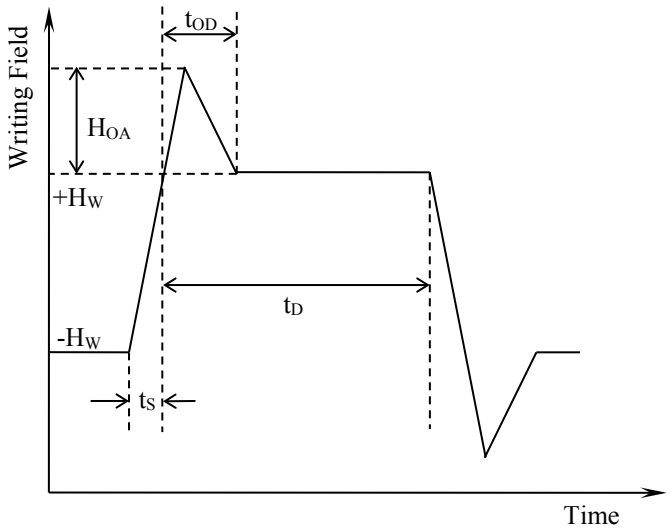


Figure 5.3: SNR vs distance between write head and NFT with different thermal spot size.

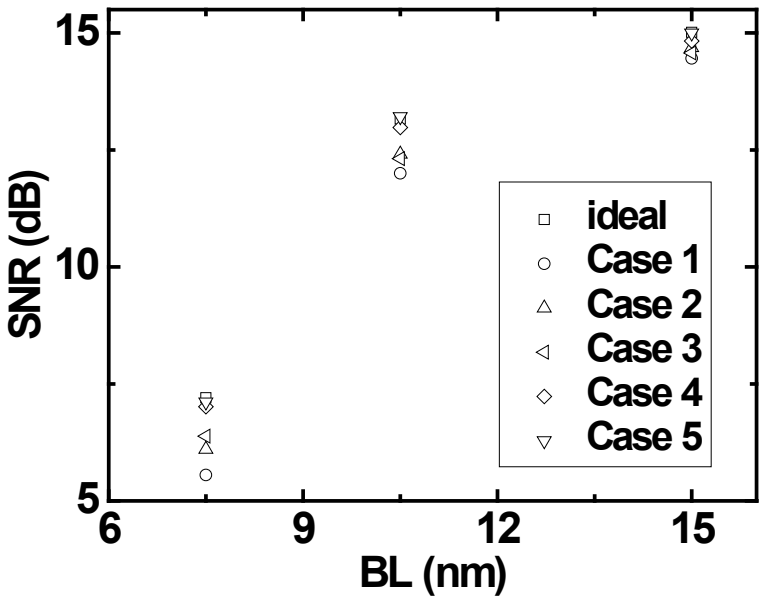
5.2.4 Write Field Compensation and Pulse Laser Recording

Transition noise becomes a leading problem, especially when BL is very small [Jiao 2015b], partly due to the switching time of the writing pole. One solution is to use write

field pre-compensation as shown in Figure 5.4. Compared with a constant field, we found that the SNR performance could be increased by at least 0.5dB using overshoot, especially when the BL is smaller than 10.5nm because the transition area could be reduced significantly. When we increase the overshoot field amplitude (H_{OA}), media SNR is improved further while it is not when we just increase the overshoot duration time (t_{OD}). Case 5 indicates that the transition noise due to switching time could also be eliminated by simply raising the writing field.



(a)



(b)

Figure 5.4: (a) Write field pulse including overshoot. (b) Media SNR performance with different overshoot fields. t_D = 1ns (BL = 10.5 nm), H_W = 10 kOe. "Ideal" means t_s = 0, H_{OA} = 0, t_{OD} = 0; case 1: t_s = 0.3 ns, H_{OA} = 0, t_{OD} = 0; case 2: t_s = 0.3 ns, H_{OA} = 2 kOe, t_{OD} = 0.2 ns; case 3: t_s = 0.3 ns, H_{OA} = 2 kOe, t_{OD} = 0.4 ns; case 4: t_s = 0.3 ns, H_{OA} = 5 kOe, t_{OD} = 0.2 ns; case 5: t_s = 0.3 ns, H_{OA} = 0, t_{OD} = 0, H_W = 15 kOe.

Unlike [122], we find media SNR improvement to be limited using a pulsed laser to heat the media, though it may reduce average laser power (Figure 5.5) [123]. When FWHM = 50nm, the slight increase of media SNR may be caused by the thermal gradient produced by pulsed laser heating but needs further study in the future.

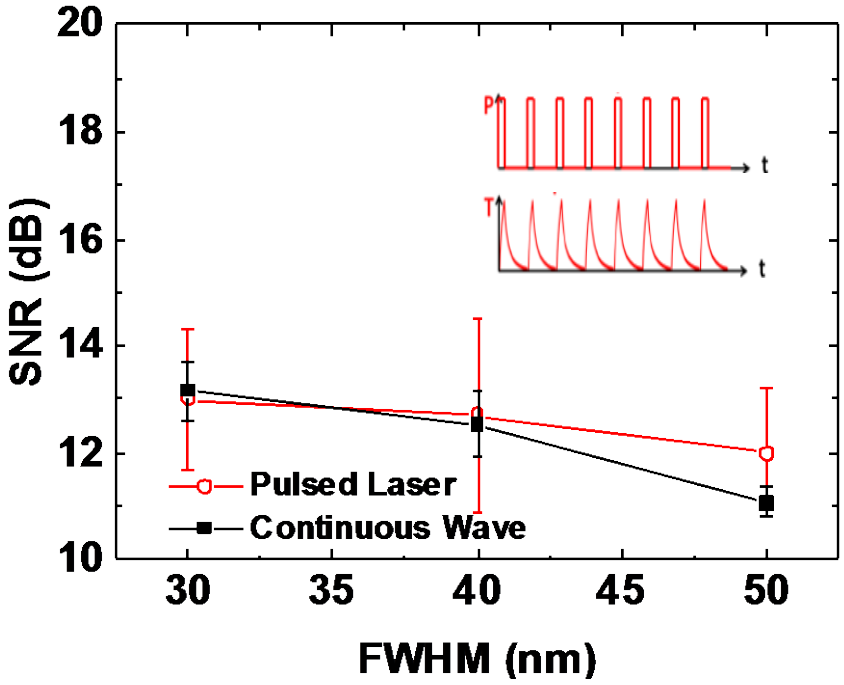


Figure 5.5: Media SNR performance w/ and w/o pulsed laser. The inset shows the wave shape of a pulsed laser, which was calculated by [122] (top) and its corresponding temperature changes on the media (bottom).

5.3 HAMR User Areal Density Prediction

Within the discussion above, the most important factors to achieve high UAD are still grain size and thermal spot size. In Table 6, the largest UAD is about 1.8 Tb/in^2 with the current system and optimization method. In order to determine the ultimate areal density, we simulated perfect recording, where we assume that if most of a grain is inside a round area with a diameter also defined as FWHM, it should be switched. At least 5000 transitions were written on the media of 5.5-nm grain pitch to obtain the SNRs and BERs under various conditions. Track pitch is assumed to be the same as FWHM when

calculating user areal density. The optimum areal density (Figure 5.6) is found for BL = 5nm and FWHM = 30 nm, which is around 3.1 Tb/in^2. This value agrees well with the channel areal density formula [124],

$$UAD = \frac{0.37Bits}{(D^{3/2} \times (track\ pitch)^{1/2})} = 3\,Tb/in^2 \tag{5-1}$$

In Figure 5.6, we can see that, as FWHM increases, SNR also increases: the trend is opposite to Figure 5.1 where we did micromagnetic simulation. This is because micromagnetic simulation reflects the smaller thermal gradient associated with a 50nm FWHM heat spot, thus causing the transition area to be the largest and SNR to be the worst. That is to say, if we could develop a non-Gaussian writer with a sharper thermal gradient, the SNR could be greatly improved (7dB when FWHM = 50nm comparing with 2dB when FWHM = 30nm).

Table 6 Storage capacity prediction with micromagnetic simulation

BL (nm)	7.5 (FWHM=50nm)	10.5 (FWHM=50nm)	15 (FWHM=50nm)
SNR (error)	6.0 (±0.2)	11.1 (±0.3)	13.7(±0.3)
Calculated BER	0.11	0.016	0.0020
UAD (Tb/in^2)	0.80	1.0	0.76

BL (nm)	7.5 (FWHM=40nm)	10.5 (FWHM=40nm)	15 (FWHM=40nm)
SNR (error)	7.0 (±0.2)	12.5 (±0.6)	14.8(±0.2)
Calculated BER	0.083	0.0060	0.0015

UAD (Tb/in^2)	1.1	1.3	1.1

BL (nm)	7.5 (FWHM=30nm)	10.5 (FWHM=30nm)	15 (FWHM=30nm)
SNR (error)	7.2 (±0.5)	***13.2 (±0.6)***	15.2(±0.3)
Calculated BER	0.079	***0.0033***	0.0015
UAD (Tb/in^2)	1.6	***1.8***	1.4

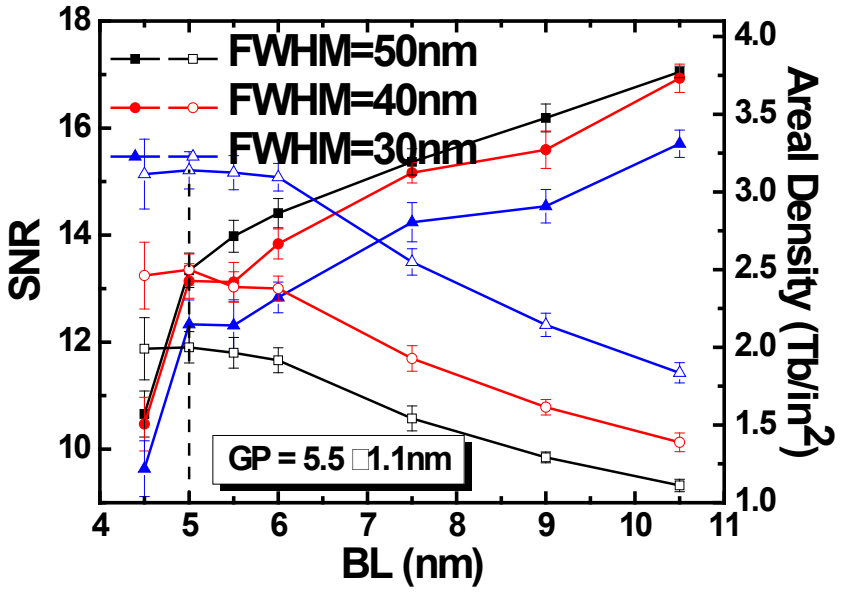


Figure 5.6: Media SNR (with solid symbols) and UAD (with open symbols) predication with different BL and FWHM with perfect recording simulation.

In order to increase the areal density even further, we shrink the track pitch as well as BL. For each FWHMs and BLs, we calculated their corresponding UADs with different track pitch values (0.6~1.0 FWHM) and the largest predictions are shown in Figure 5.7. The maximum UAD is about 5.1Tb/in^2 for FWHM = 20nm, BL = 6 nm and track_pitch =

0.8 FWHM. This prediction is very close to the data in [118] but with a larger grain pitch (5.5nm in this work and <5nm in [118]), a consequence of perfect recording and hence no switching field distribution in this work.

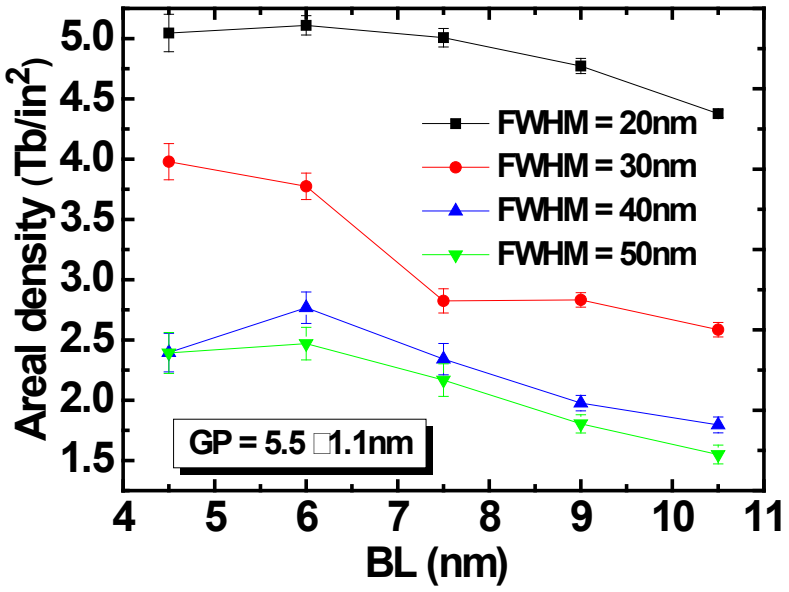


Figure 5.7: UAD with squeezed track, prediction using perfect recording simulation.

5.4 Conclusion

In order to optimize a HAMR system, we first performed micromagnetic simulation with the renormalized-LLG method under different conditions: we examined the effects of BL, thermal spot size, grain size and deviation, magnitude and shape of writing field, reader size, and pulsed recording. The demonstrated UAD could be as high as 1.8Tb/in^2 for FWHM and track pitch equal to 30nm and GP = 6.4nm, which are technologically feasible.

Perfect recording (ideal thermal gradient) could greatly increase the UAD, to more than 3Tb/in^2 and 5 Tb/in^2 w/ and w/o squeezed track (GP = 5.5nm), respectively. Finally, we note that TDMR would likely improve these results.

6 Understanding Transition and Remanence Noise in HAMR

6.1 Introduction

In Seagate's recent presentation [125], remanence noise appears to be a very important factor contributing to jitter, unlike the case of PMR. Some earlier results [126]–[129] also examined the effects of different types of noise in HAMR. These results suggest the importance of a systematic study.

This chapter will focus on three parts: first, we will introduce two methods for noise splitting. Then we will try to understand the different contributions to noise in HAMR, specifically the remanence noise and transition noise. We will discuss how they affect the performance of HAMR. Finally, we will examine the relationship between jitter and remanence SNR using our simulation model and method.

6.2 Simulation Parameters and Method

Our micromagnetic simulation program is based on a renormalized Landau-Lifshitz-Gilbert (LLG) model, where we assume the media consists of thousands of spin cubes which represent a cluster of atoms. The typical size of cubes is $1.5\times1.5\times1.5\text{nm}^3$, so

if the dimension of a media is 384×96×12nm^3 there will be 256×64×8=13,1072 cubes (grids) in our simulation. For each time step, the LLG equation is solved for each cube independently and in a general simulation, the total time steps are of the order of 10^5-10^6. The computations are executed using GPGPU that can execute thousands of computations concurrently, e.g., the GTX 580 and 680 have 512 and 1536 CUDA cores, respectively. These cores can be used efficiently by employing pipeline and scheduling techniques [11]. As a result, a 1×10^5-step simulation will take about 2 days on our computer with a GTX 580, which is about 4X faster than with a CPU i7 6700K.

However, this is still slow compared with LLB (Landau-Lifshitz-Bloch) [12] or single-spin based models. We find the computation of demagnetizing field (using FFT) takes about O(nlogn), while the complexity of LLG solver is only about O(n). Obviously, FFT calculation takes most of the resources especially when n is very large so we should limit the number of grids involved in FFT calculation to a region around the thermal spot (demagnetizing field window). The size of this window should be carefully optimized because demagnetizing field plays an important role in SNR calculation: we find that limiting the size to 96×96nm^2 for BL (bit length) = 25nm and 50nm and 96×192 nm^2 for BL = 100nm does not impact the recorded patterns and the playback SNRs too much. With this technique, our simulation can be accelerated by 2~3X further.

The thermal spot generated by the NFT is described using a 2D Gaussian distribution on the media,

$$T(x,y,t) = 300 + (T_{max} - 300)exp\left[\frac{-(x-vt)^2 - (y-y_0)^2}{2\sigma^2}\right]$$

(6-1)

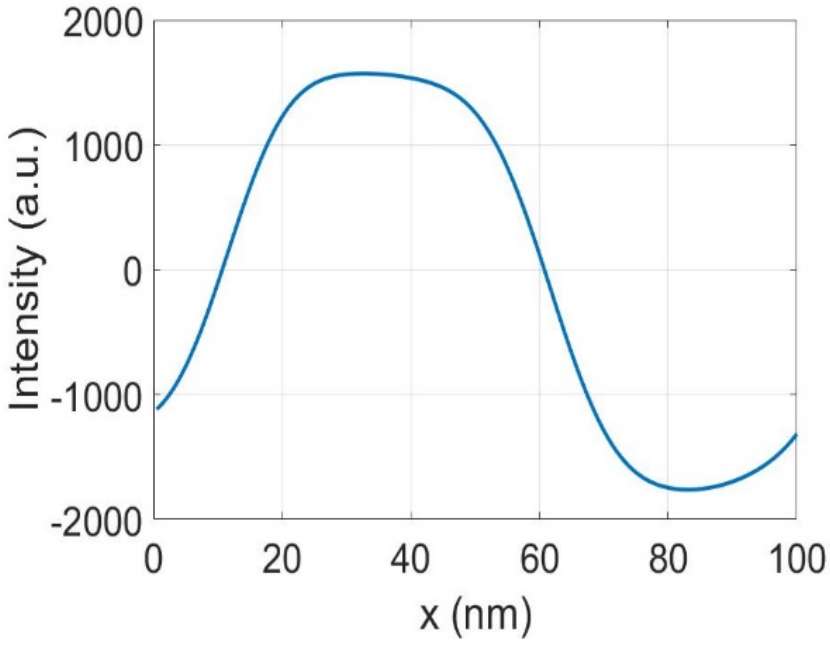
where $\sigma = FWHM/2\sqrt{2ln2}$ and y_0 is the track center. T_{max} is the peak value of the heat spot and v is the velocity of the recording head. To investigate the writing field effect, we need a constant track width, which is 70nm, so we vary the FWHM values between 80nm ~100nm. Writing field is considered as uniform and the angle between easy axis of media and writing field is 22°. There is no rise time. The other parameters in the writing process are listed in Table 7.

Table 7 Important parameters in micro-magnetic simulation

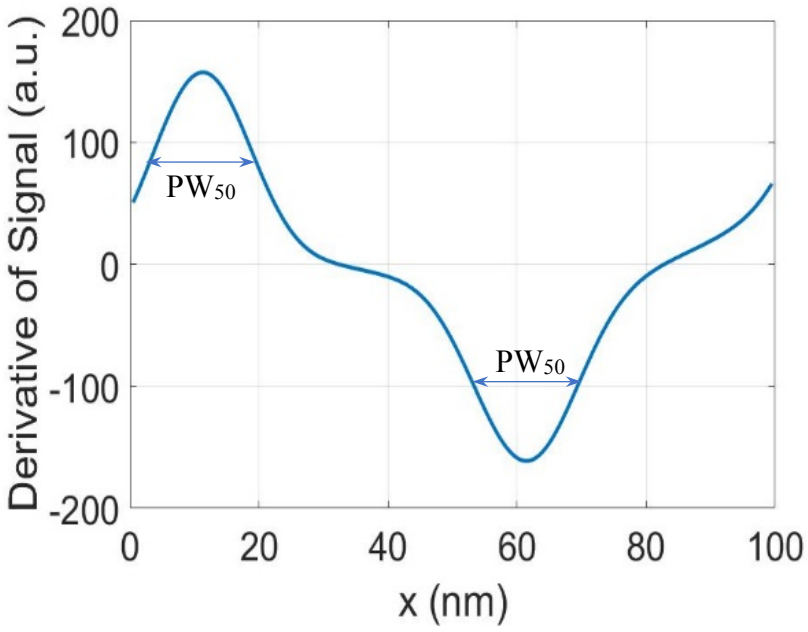
Parameters	Values
GP	5.5±20%, 6.5±20%, 8.5±20% nm
Grain boundary	1 nm
Head velocity	20 m/s
Media thickness	12 nm
M_s (0C, media)	1100 emu/cm^3
K_u (0C, media)	7×10^7 erg/ cm^3
T_c (media)	700 K
δT_c (media)	0%
δH_k (media)	5%

6.3 Transition Noise and Remanence Noise

There are mainly two types of media noise in magnetic recording: transition noise and remanence noise. Separation of the noise from the playback signal can be a difficult and arbitrary task. One reason is that neither noise is well-defined in HAMR; the other reason is that it is unclear whether they could interact with each other in the transition regions. Besides, the reader also plays an important role in the playback process, which makes this problem even more problematic. The simplest case is large bit length (BL) and uniform patterns (tone), where a simple division based on distance to the transition (spatial splitting) can suffice if we assume PW_{50} is independent of BL and remanence noise is constant at all points in the media [129]. PW_{50} is defined as the half-amplitude pulse width of the derivative of the playback signal (Figure 6.1) [25].



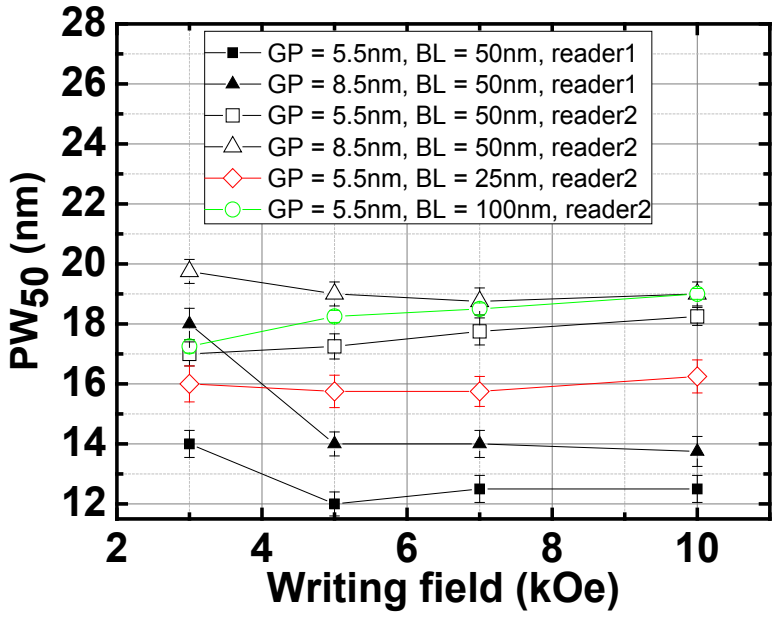
(a)



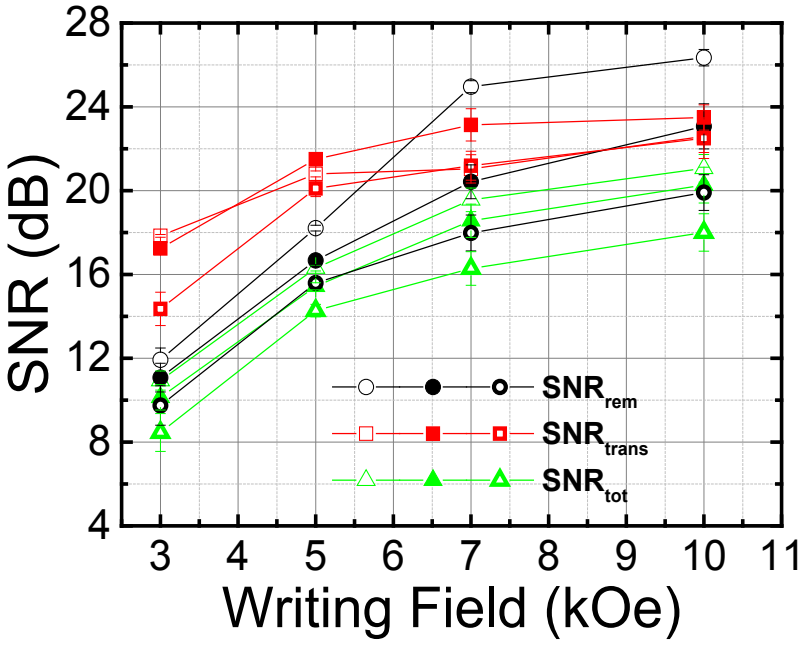
(b)

Figure 6.1: (a) Averaged playback signal in one period (dibit) with reader 2 when H = 5kOe, GP = 5.5nm. (b) The derivative of (a). PW_{50} is also shown.

Figure 6.2(a) shows PW_{50} values under different simulation conditions. We find that PW_{50} changes with applied field only for excessively low values. However, it is very sensitive to GP and reader parameters. BL (above 50 nm) has little effect on PW_{50} which indicates that a simple division between transition and remanence noise may be valid: the remanence noise can be calculated by subtracting the total noise with a smaller bit length (e.g. 50nm) from a larger one (e.g. 100nm); then we can get the transition noise based on the total noise and remanence noise.



(a)



(b)

Figure 6.2: (a) PW_{50} vs writing field. Parameters of Reader 1: SSS (shield-to-shield spacing) = 11nm, reader width = 42nm and fly height = 6nm; Parameters of Reader 2: SSS = 22nm, reader width = 36nm and fly height = 10nm. (b) Calculated SNR using the method in [8]

with BL = 50nm. Open symbols are the results at GP = 5.5nm, solid ones are at GP = 6.5nm and half open ones are at GP = 8.5nm.

Figure 6.2 (b) shows SNRs vs writing field. It is found that applied field has a larger effect on remanence SNR than transition SNR, especially when the field is larger than 5kOe. In addition, remanence SNR for small GP can be more greatly improved with increased writing field than is found for large GP. This is a consequence of two effects: small grains have more potential for low noise, but insufficient Zeeman energy, i.e., insufficient field, is also more likely for small grains.

However, the assumption for this method only holds when BL is about 1.5~2 times PW_{50}, which means the transition region is $(1.5\sim2) \times PW_{50}$ in our calculation. For cases where BL is smaller than this value, we can use a more sensitive reader (reader 1 in Figure 6.2 (a)) to calculate playback signals, which decreases PW_{50} for 5.5-nm GP and 8.5-nm GP by 4~5nm, respectively. However, a more sensitive reader may also reduce SNR and also misrepresent the performance found experimentally, where the SSS is not easily decreased.

Another method to address this problem is to use principle components analysis (PCA) [130], [131], which is a statistical technique based on signal variances. The assumption is that the playback signals consist of a set of values which are linearly uncorrelated variables called principal components (PC) in a certain coordinate system. The task of PCA is to find this coordinate system. This may allow a more rigorous splitting of noise into remanence and transition contributions.

Let matrix X be the centered data with n×p size, n is the length of discretized signals (n = 100 when BL = 25nm and n = 200 when BL = 50nm) and p is the number of runs we

repeated (p =48). Then we could use singular value decomposition (SVD) to extract the PCs,

$$X = USV^T \tag{6-2}$$

where V is a matrix of eigenvectors (each column is an eigenvector or principal axis), and S is the diagonal matrix of singular values s_i. Based on the definition of covariance, we obtain

$$C = \frac{VSU^T USV^T}{n-1} = V\frac{S^2}{n-1}V^T \tag{6-3}$$

meaning that the singular values are related to the eigenvalues of the covariance matrix by

$$\lambda_i = \frac{s_i^2}{n-1} \tag{6-4}$$

Principal components are given by

$$XV = USV^T V = US \tag{6-5}$$

To reconstruct the signals, we can use

$$X_R = U_R S_R V_R^T \tag{6-6}$$

where U_R and V_R are the first R columns of U and V, respectively, and S_R means the $R \times R$ upper-left part of S. X_R is the reconstructed playback signal. Figure 6.3 shows the reconstructed playback signals in one period with the first R PCs with 5.5-nm grain pitch.

When R = 1 (Figure 6.3 (a)), we show only signal components with little noise. In (b) (c) when R is larger, we can see larger fluctuations among the reproduced signals, especially in the vicinity of peaks and valleys. These fluctuations are considered to be remanence noise. When R = 10 (Figure 6.3 (d)), more components were counted to reproduce the playback signal and these signals are very close to our original playback signals. Actually, if we compare the different PCs in Figure 6.4, we can find: (1) the first PC (blue curve) always corresponds to the "pure" signals; (2) lower rank PCs (r = 2~4) are remanence-noise-like (amplitudes don't change with position x); (2) higher rank PCs (r = 6~8) are transition-noise-like (amplitudes become larger around transition regions).

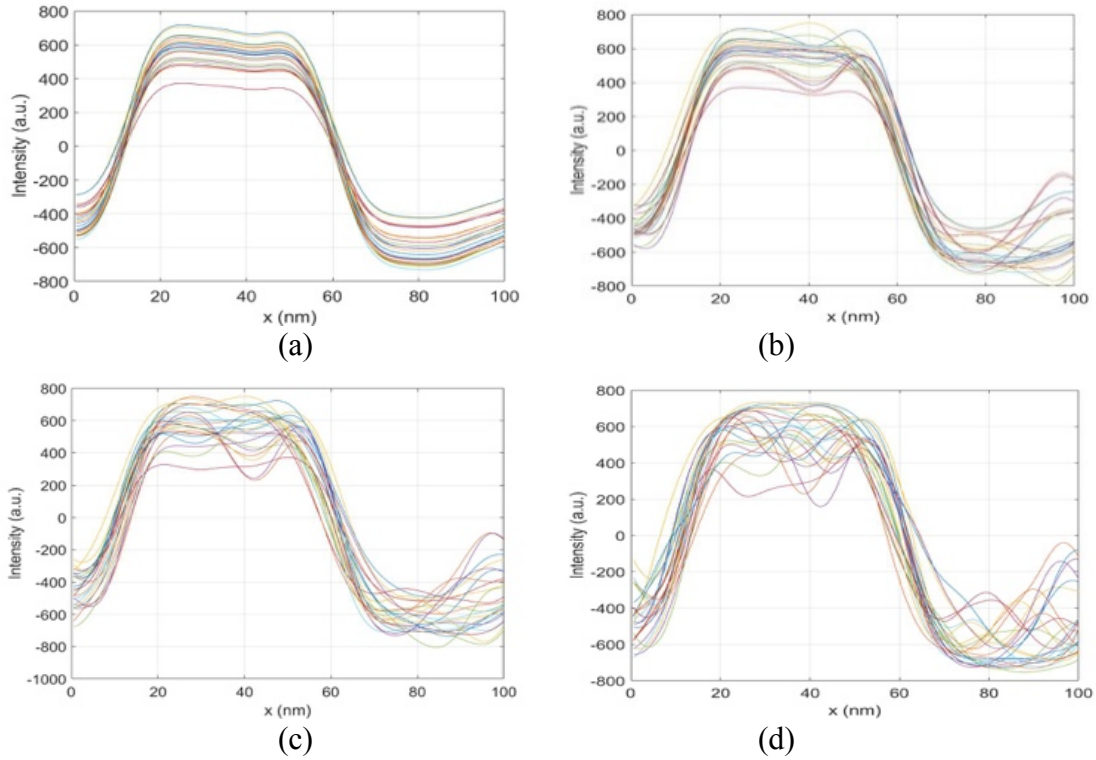
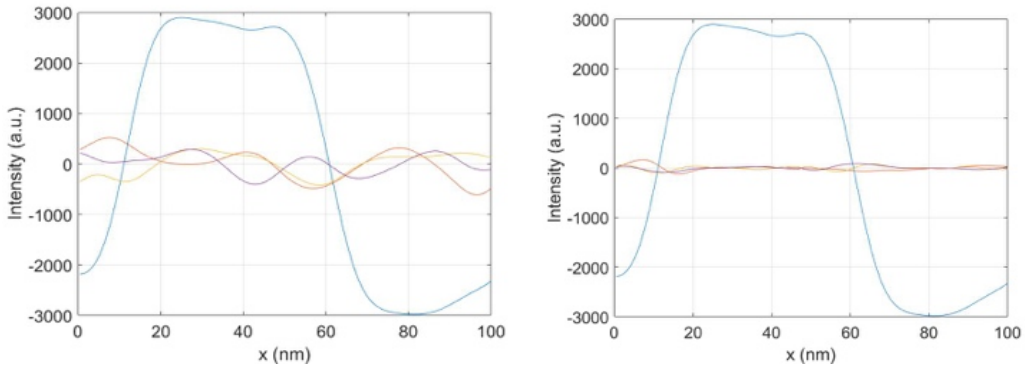


Figure 6.3: Reconstructed playback signals with different number of PCs. (a) R = 1; (b) R = 2; (c) R = 3; (d) R = 10.



(a) (b)

Figure 6.4: Different PCs of Figure 6.3 in one period. The largest blue curve is the first PC (r = 1). (a) r = 1, 2, 3, 4; (b) r = 1, 6, 7, 8.

We applied this technique for BL = 25nm and again obtained transition and remanence noise for 3kOe~10kOe writing field (Figure 6.5). These curves are very similar to Figure 6.2 (b) except that the saturation field for overall SNR is only 7kOe, much smaller than the 50-nm BL case. It can be seen that the transition noise continues to improve with write field above 7kOe (unlike Figure 6.2 (b)): Presumably this means that the PCA approach includes slightly more remanence region than does the traditional approach. In contrast, the remanence SNR saturates sooner in the PCA approach than in the traditional method.

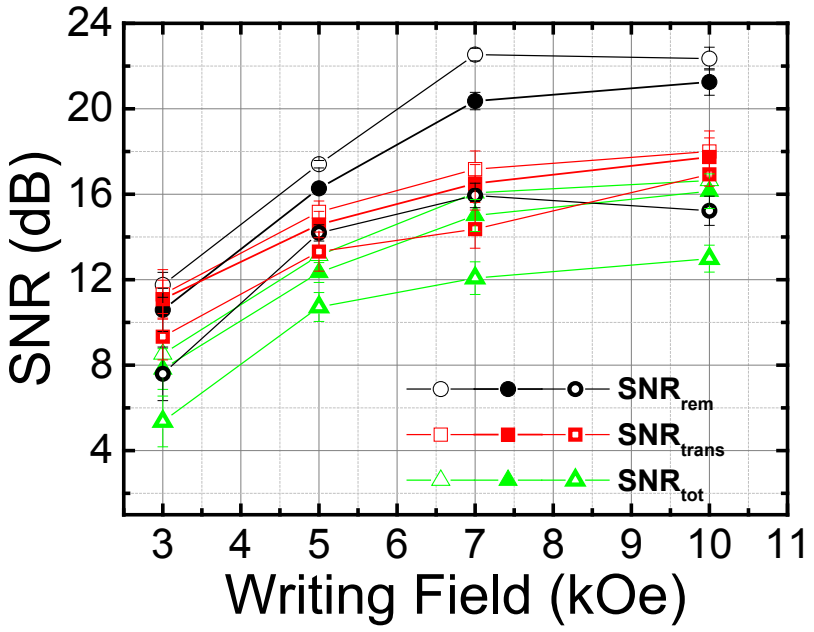


Figure 6.5: Calculated SNR using PCA with BL = 25nm. Open symbols are the results at GP = 5.5nm, solid ones are at GP = 6.5nm and half open ones are at GP = 8.5nm.

6.4 Relationship between SNR and Jitter

Jitter is an important experimentally measured quantity used to evaluate the quality of recorded patterns. So it is important to know the relationship between jitter and SNR (which is more commonly used in signal processing). [125] has introduced an equation to calculate jitter in HAMR:

$$jitter = \sqrt{\left(\sqrt{\frac{D^3}{12 \times RW}}\right)^2 + \left(\sqrt{\frac{D}{min(WW, RW)}} \times \frac{\delta T}{\nabla T}\right)^2 + (c \times PW_{50} \times NSR_{rem})^2} \quad (6\text{-}7)$$

where D stands for the grain pitch. RW and WW are reader width and write width, respectively. The first term $\sqrt{\frac{D^3}{12\times RW}}$ is called the grain-size theoretical limit, the second term $\sqrt{\frac{D}{min(WW,RW)}}\times\frac{\delta T}{\nabla T}$ is jitter caused by switching field distribution (SFD) [72], [127] and the last term introduces the remanence noise-to-signal ratio (NSR_{rem}), which depends strongly on the strength of the writing field, as well as PW_{50}, which is highly reader dependent (Figure 6.2 (a)).

We define remanence jitter,

$$jitter_{rem} = c \times PW_{50} \times NSR_{rem} \tag{6-8}$$

PW_{50} and NSR_{rem} are known, thus the prefactor c could be determined from $jitter_{rem}$. Two methods could be used to calculate it: (1) using equation (6-7), we can get

$$jitter' = \sqrt{jitter^2 - jitter_{gs}^2} \tag{6-9}$$

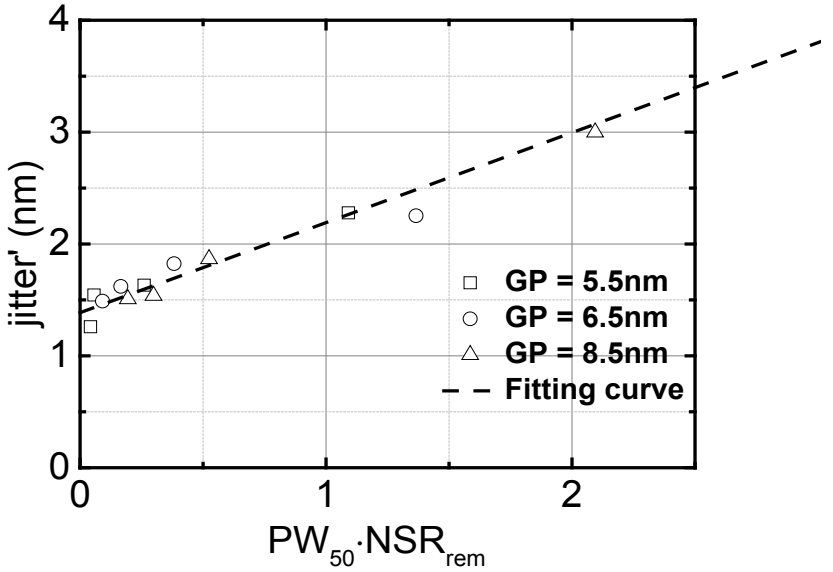
where $jitter_{gs} = \sqrt{\frac{D^3}{12\times RW}}$.

The second term can be eliminated if we use δT = 0%. Method (2) uses the remanence jitter directly extracted from the reconstructed signal using PCA. Noting that principal axes are orthogonal to each other, it is reasonable to get the "sum of square" relationship if we consider each term of the total jitter in equation (6-7) as a projection along one principal axes.

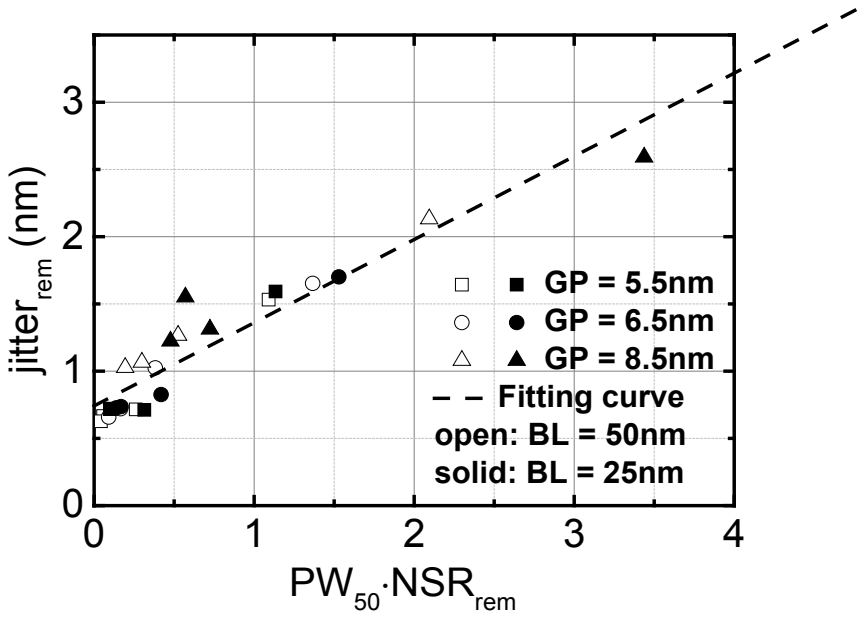
Figure 6.6 shows the results using the two methods. First, the points in both figures follow the linear relationship which is also demonstrated in Seagate's experimental results. Second, the slopes of the fitting lines are 0.80 and 0.62, respectively, which agrees with the experimental result in [125]. Further work will focus on HAMR with more realistic simulation conditions, such as finite switching time, non-uniform writing field distribution and composite media.

6.5 Conclusion

Simulation results show that transition noise and remanence noise exhibit different behaviors, so it is meaningful to distinguish them. Principal components analysis can be used to extract both remanence SNR and jitter and appears to be a powerful tool to study noise in HAMR. Finally, the linear relationship between jitter and $PW_{50} \cdot NSR_{rem}$ is verified by our simulation with different BLs and GPs.



(a)



(b)

Figure 6.6: Relationship between remanence jitter (or jitter' in method (1)) and the product of PW_{50} and remanence NSR using the (a) first and (b) second method, respectively. Dash lines are the fitting functions.

7 Summary

We will make a summary of this dissertation in this chapter:

- Chapter 1 introduced the current status of the magnetic recording industry and then HAMR, which is considered to be the most promising candidate to push areal density further beyond 1 Tb/in^2. Due to the complexity of its noise sources, it is necessary to conduct a numerical optimization at the system level.
- Our simulation program used for this numerical optimization is described in Chapter 2 and it can be divided into four modules: media generation, which uses Bowyer–Watson algorithm and an optimized polygon offsetting method; recording simulation, which is based on our renormalized LLG model; playback module, using the reciprocity principle; and state detection with PRML method. The simulation accuracy can be improved further without increasing system load by using anisotropic exchange parameters instead of isotropic values (Chapter 3).
- When using a Gaussian distribution as the thermal profile, it is found that the optimal peak temperature is about 850 K and the reader width should be about half of the FWHM in order to get the largest SNR. With the error function relationship between SNR and BER we obtained in Chapter 4 and the Shannon capacity limit equation, FWHM should be about 30 nm in order to get the best UAD though it is not the optimum for SNR. For a 6.4 nm-GP medium, the best UAD is found to be

about 1.8 Tb/in^2, which is obtained at BL = 10.5 nm. The target value predicted by perfect recording is 3 (w/o squeezed track) ~ 5 (w/ squeezed track) Tb/in^2. The grain size scaling is less important than thermal gradient optimization for the current level of technology. In addition, increasing the amplitude of writing field overshoot/bump is found to be very useful to compensate for the SNR loss caused by the writing field switching time.

- In Chapter 6, we introduced two methods (spatial splitting and principal components analysis) to distinguish remanence and transition noise. The remanence noise is more sensitive to the writing field than the transition noise and the former is also considered as the main contributing factor to the SNR/jitter saturation effect in HAMR. Finally, the linear relationship between $\text{jitter}_{\text{rem}}$ and $\text{PW}_{50} \cdot \text{NSR}_{\text{rem}}$ was verified.

References

[1] "Research Development Consultants: Ron Dennison--- Services." [Online]. Available: http://rondennison.com/services.htm. [Accessed: 10-Nov-2018].

[2] E. D. Daniel, C. D. Mee, and M. H. Clark, *Magnetic recording: the first 100 years*. John Wiley & Sons, 1999.

[3] C. D. Mee and E. D. Daniel, *Magnetic recording handbook: technology and applications*. McGraw-Hill New York, 1990.

[4] D. L. Curtis and R. B. Larsen, "Magnetic drum memory system," Apr-1960.

[5] D. A. Thompson and J. S. Best, "The future of magnetic data storage technology," *IBM J. Res. Dev.*, vol. 44, no. 3, p. 311, 2000.

[6] "Hard disk drive," *Wikipedia*. 05-Nov-2018.

[7] "Moore's law," *Wikipedia*. 04-Nov-2018.

[8] "Solid-state drive," *Wikipedia*. 10-Nov-2018.

[9] F. Chen, D. A. Koufaty, and X. Zhang, "Understanding intrinsic characteristics and system implications of flash memory based solid state drives," in *ACM SIGMETRICS Performance Evaluation Review*, 2009, vol. 37, pp. 181–192.

[10] R. Waser and M. Aono, "Nanoionics-based resistive switching memories," *Nat. Mater.*, vol. 6, no. 11, p. 833, 2007.

[11] H.-S. P. Wong *et al.*, "Phase change memory," *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, 2010.

[12] S. Raoux *et al.*, "Phase-change random access memory: A scalable technology," *IBM J. Res. Dev.*, vol. 52, no. 4.5, pp. 465–479, 2008.

[13] N. J. Knall and M. Johnson, "Three-dimensional memory array and method of fabrication," Jul-2002.

[14] "Data recovery," *Wikipedia*. 05-Nov-2018.

[15] "HGST Ships 6TB Ultrastar® He6 Helium-filled Drives for High-density, Massive Scale-out Data Center Environments | HGST." [Online]. Available: https://www.hgst.com/zh/company/media-room/press-releases/hgst-ships-6tb-ultrastar-he6-helium-filled. [Accessed: 10-Nov-2018].

[16] "Breaking Capacity Barriers With Seagate Shingled Magnetic Recording | Seagate US." [Online]. Available: https://www.seagate.com/tech-insights/breaking-areal-density-barriers-with-seagate-smr-master-ti/. [Accessed: 10-Nov-2018].

[17] "Seagate MACH.2 Multi Actuator Technology Breaks Throughput Record; HAMR Reliability Tests Exceed Industry Standards | Seagate Blog." [Online]. Available: https://blog.seagate.com/enterprises/mach2-and-hamr-breakthrough-ocp/. [Accessed: 10-Nov-2018].

[18] "IBM Archives: IBM 3340 disk unit." [Online]. Available: https://www-03.ibm.com/ibm/history/exhibits/storage/storage_PH11-52.html. [Accessed: 11-Nov-2018].
[19] "How to Recover a Dead Hard Disk," *wikiHow*. [Online]. Available: https://www.wikihow.com/Recover-a-Dead-Hard-Disk. [Accessed: 10-Nov-2018].
[20] M. Bohl, *Introduction to IBM direct access storage devices*. Sra, 1981.
[21] E. B. Svedberg, S. Khizroev, and D. Litvinov, "Magnetic force microscopy study of perpendicular media: Signal-to-noise determination and transition noise analysis," *J. Appl. Phys.*, vol. 91, no. 8, pp. 5365–5370, 2002.
[22] T. Chen and T. Yamashita, "Physical origin of limits in the performance of thin-film longitudinal recording media," *IEEE Trans. Magn.*, vol. 24, no. 6, pp. 2700–2705, 1988.
[23] K. E. Johnson, P. R. Ivett, D. R. Timmons, M. Mirzamaani, S. E. Lambert, and T. Yogi, "The effect of Cr underlayer thickness on magnetic and structural properties of CoPtCr thin films," *J. Appl. Phys.*, vol. 67, no. 9, pp. 4686–4688, 1990.
[24] S. N. Piramanayagam, H. B. Zhao, J. Z. Shi, and C. S. Mah, "Advanced perpendicular recording media structure with a magnetic intermediate layer," *Appl. Phys. Lett.*, vol. 88, no. 9, p. 092501, 2006.
[25] H. N. Bertram, *Theory of magnetic recording*. Cambridge University Press, 1994.
[26] B. R. Acharya, J. N. Zhou, M. Zheng, G. Choe, E. N. Abarra, and K. E. Johnson, "Anti-parallel coupled soft under layers for high-density perpendicular recording," *IEEE Trans. Magn.*, vol. 40, no. 4, pp. 2383–2385, 2004.
[27] M. Mallary, A. Torabi, and M. Benakli, "One terabit per square inch perpendicular recording conceptual design," *IEEE Trans. Magn.*, vol. 38, no. 4, pp. 1719–1724, 2002.
[28] R. H. Victora and X. Shen, "Exchange coupled composite media for perpendicular magnetic recording," *IEEE Trans. Magn.*, vol. 41, no. 10, pp. 2828–2833, 2005.
[29] J.-P. Wang, W. Shen, and J. Bai, "Exchange coupled composite media for perpendicular magnetic recording," *IEEE Trans. Magn.*, vol. 41, no. 10, pp. 3181–3186, 2005.
[30] H. J. Richter, "The transition from longitudinal to perpendicular recording," *J. Phys. Appl. Phys.*, vol. 40, no. 9, p. R149, 2007.
[31] R. Wood, "Shingled magnetic recording and two-dimensional magnetic recording," *IEEE Magn. Soc. St. Clara Val.*, 2010.
[32] "perpendicular recording Definition from PC Magazine Encyclopedia." [Online]. Available: https://www.pcmag.com/encyclopedia/term/49119/perpendicular-recording. [Accessed: 11-Nov-2018].
[33] M. N. Baibich *et al.*, "Giant magnetoresistance of (001) Fe/(001) Cr magnetic superlattices," *Phys. Rev. Lett.*, vol. 61, no. 21, p. 2472, 1988.
[34] G. Binasch, P. Grünberg, F. Saurenbach, and W. Zinn, "Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange," *Phys. Rev. B*, vol. 39, no. 7, p. 4828, 1989.

[35] M. Takagishi, K. Koi, M. Yoshikawa, T. Funayama, H. Iwasaki, and M. Sahashi, "The applicability of CPP-GMR heads for magnetic recording," *IEEE Trans. Magn.*, vol. 38, no. 5, pp. 2277–2282, 2002.
[36] J. R. Childress and R. E. Fontana Jr, "Magnetic recording read head sensor technology," *Comptes Rendus Phys.*, vol. 6, no. 9, pp. 997–1012, 2005.
[37] M. A. Gijs, S. K. Lenczowski, and J. B. Giesbers, "Perpendicular giant magnetoresistance of microstructured Fe/Cr magnetic multilayers from 4.2 to 300 K," *Phys. Rev. Lett.*, vol. 70, no. 21, p. 3343, 1993.
[38] P. M. Levy, S. Zhang, T. Ono, and T. Shinjo, "Electrical transport in corrugated multilayered structures," *Phys. Rev. B*, vol. 52, no. 22, p. 16049, 1995.
[39] A. Vedyayev, M. Chshiev, N. Ryzhanova, B. Dieny, C. Cowache, and F. Brouers, "A unified theory of CIP and CPP giant magnetoresistance in magnetic sandwiches," *J. Magn. Magn. Mater.*, vol. 172, no. 1–2, pp. 53–60, 1997.
[40] M. Saito *et al.*, "Narrow track current-perpendicular-to-plane spin valve GMR heads," *IEEE Trans. Magn.*, vol. 40, no. 1, pp. 207–212, Jan. 2004.
[41] T. Nakatani, S. Li, Y. Sakuraba, T. Furubayashi, and K. Hono, "Advanced CPP-GMR Spin-Valve Sensors for Narrow Reader Applications," *IEEE Trans. Magn.*, vol. 54, no. 2, pp. 1–11, Feb. 2018.
[42] "Giant magnetoresistance," *Wikipedia*. 07-Nov-2018.
[43] B. D. Cullity and C. D. Graham, *Introduction to magnetic materials*. John Wiley & Sons, 2011.
[44] M. H. Kryder and R. W. Gustafson, "High-density perpendicular recording—advances, issues, and extensibility," *J. Magn. Magn. Mater.*, vol. 287, pp. 449–458, 2005.
[45] B. Marchon and T. Olson, "Magnetic spacing trends: from LMR to PMR and beyond," *IEEE Trans. Magn.*, vol. 45, no. 10, pp. 3608–3611, 2009.
[46] M. H. Kryder *et al.*, "Heat assisted magnetic recording," *Proc. IEEE*, vol. 96, no. 11, pp. 1810–1835, 2008.
[47] J.-G. Zhu, X. Zhu, and Y. Tang, "Microwave assisted magnetic recording," *IEEE Trans. Magn.*, vol. 44, no. 1, pp. 125–131, 2008.
[48] W. Li, "Acoustically Assisted Magnetic Recording," 2015.
[49] "Seagate Demos HAMR Drives at InterMag Conference | Seagate Blog." [Online]. Available: https://blog.seagate.com/business/seagate-demos-hamr-drives-at-intermag-conference/. [Accessed: 11-Nov-2018].
[50] admin, "Seagate Continues to Lead as HAMR Technology Advances," *Seagate Blog*, 08-Apr-2016. .
[51] W. A. Challener *et al.*, "Heat-assisted magnetic recording by a near-field transducer with efficient optical energy transfer," *Nat. Photonics*, vol. 3, no. 4, p. 220, 2009.
[52] B. C. Stipe *et al.*, "Magnetic recording at 1.5 Pb m- 2 using an integrated plasmonic antenna," *Nat. Photonics*, vol. 4, no. 7, p. 484, 2010.
[53] T. Matsumoto, F. Akagi, M. Mochizuki, H. Miyamoto, and B. Stipe, "Integrated head design using a nanobeak antenna for thermally assisted magnetic recording," *Opt. Express*, vol. 20, no. 17, pp. 18946–18954, 2012.

[54] N. Zhou *et al.*, "Plasmonic near-field transducer for heat-assisted magnetic recording," *Nanophotonics*, vol. 3, no. 3, pp. 141–155, 2014.
[55] R. Ikkawi *et al.*, "Near-field optical transducer for heat-assisted magnetic recording for beyond-10-Tbit/in2 densities," *J. Nanoelectron. Optoelectron.*, vol. 3, no. 1, pp. 44–54, 2008.
[56] P. F. Liao and A. Wokaun, "Lightning rod effect in surface enhanced Raman scattering," *J. Chem. Phys.*, vol. 76, no. 1, pp. 751–752, 1982.
[57] 1PK Aravind, A. Nitzan, and H. Metiu, "The interaction between electromagnetic resonances and its role in spectroscopic studies of molecules adsorbed on colloidal particles or metal spheres," *Surf. Sci.*, vol. 110, no. 1, pp. 189–204, 1981.
[58] M. Moskovits, "Surface-enhanced spectroscopy," *Rev. Mod. Phys.*, vol. 57, no. 3, p. 783, 1985.
[59] D. Weller *et al.*, "High K/sub u/materials approach to 100 Gbits/in/sup 2," *IEEE Trans. Magn.*, vol. 36, no. 1, pp. 10–15, 2000.
[60] K. Barmak, J. Kim, S. Shell, E. B. Svedberg, and J. K. Howard, "Calorimetric studies of the A1 to L1 0 transformation in FePt and CoPt thin films," *Appl. Phys. Lett.*, vol. 80, no. 22, pp. 4268–4270, 2002.
[61] S. N. Piramanayagam, "Perpendicular recording media for hard disk drives," *J Appl Phys*, p. 23.
[62] S. D. Granz, K. Barmak, and M. H. Kryder, "Granular L10 FePt-B and FePt-B-Ag (001) thin films for heat assisted magnetic recording," *J. Appl. Phys.*, vol. 111, no. 7, p. 07B709, 2012.
[63] L. Zhang, Y. K. Takahashi, K. Hono, B. C. Stipe, J.-Y. Juang, and M. Grobis, "FePtAg-C nanogranular film as thermally assisted magnetic recording (TAR) media," *IEEE Trans. Magn.*, vol. 47, no. 10, pp. 4062–4065, 2011.
[64] O. Mosendz, S. Pisana, J. W. Reiner, B. Stipe, and D. Weller, "Ultra-high coercivity small-grain FePt media for thermally assisted recording," *J. Appl. Phys.*, vol. 111, no. 7, p. 07B729, 2012.
[65] K. F. Dong, H. H. Li, Y. G. Peng, G. Ju, G. M. Chow, and J. S. Chen, "Well-isolated L 10 FePt–SiN x–C nanocomposite films with large coercivity and small grain size," *J. Appl. Phys.*, vol. 111, no. 7, p. 07A308, 2012.
[66] J. Lyubina, B. Rellinghaus, O. Gutfleisch, and M. Albrecht, "Structure and magnetic properties of L10-ordered Fe–Pt alloys and nanoparticles," in *Handbook of magnetic materials*, vol. 19, Elsevier, 2011, pp. 291–407.
[67] C. –. Rong *et al.*, "Size-dependent chemical and magnetic ordering in L10-FePt nanoparticles," *Adv. Mater.*, vol. 18, no. 22, pp. 2984–2988, 2006.
[68] R. H. Victora and A. Ghoreyshi, "Optical Analysis of HAMR Media," *IEEE Trans. Magn.*, 2018.
[69] Y. Jiao, J. Hohlfield, and R. H. Victora, "Understanding Transition and Remanence Noise in HAMR," *IEEE Trans. Magn.*, no. 99, pp. 1–5, 2018.
[70] S. Wang and R. H. Victora, "Temperature distribution of granular media for heat assisted magnetic recording," *J. Appl. Phys.*, vol. 117, no. 17, p. 17D147, 2015.
[71] Z. Liu, Y. Jiao, and R. H. Victora, "Composite media for high density heat assisted magnetic recording," *Appl. Phys. Lett.*, vol. 108, no. 23, p. 232402, 2016.

[72] Z. Liu, P.-W. Huang, G. Ju, and R. H. Victora, “Thermal switching probability distribution of L10 FePt for heat assisted magnetic recording,” *Appl. Phys. Lett.*, vol. 110, no. 18, p. 182405, 2017.
[73] R. F. L. Evans, R. W. Chantrell, U. Nowak, A. Lyberatos, and H.-J. Richter, “Thermally induced error: Density limit for magnetic data storage,” *Appl. Phys. Lett.*, vol. 100, no. 10, p. 102402, 2012.
[74] C. Rea *et al.*, “Areal-density limits for heat-assisted magnetic recording and perpendicular magnetic recording,” *IEEE Trans Magn*, vol. 52, no. 7, pp. 1–4, 2016.
[75] C. Rea *et al.*, “Writer and reader head-to-media spacing sensitivity assessment in HAMR,” *IEEE Trans. Magn.*, vol. 52, no. 2, pp. 1–6, 2016.
[76] Z. Liu *et al.*, “Curvature and Skew in Heat-Assisted Magnetic Recording,” *IEEE Trans. Magn.*, 2018.
[77] C. Rea *et al.*, “High track pitch capability for HAMR recording,” *IEEE Trans. Magn.*, vol. 53, no. 2, pp. 1–7, 2017.
[78] Y. Wang, “System study of two dimensional magnetic recording system,” 2014.
[79] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational geometry: algorithms and applications*. Springer-Verlag TELOS, 2008.
[80] E. N. Gilbert, “Random subdivisions of space into crystals,” *Ann. Math. Stat.*, vol. 33, no. 3, pp. 958–972, 1962.
[81] S. W. Sloan and G. T. Houlsby, “An implementation of Watson’s algorithm for computing 2-dimensional Delaunay triangulations,” *Adv. Eng. Softw. 1978*, vol. 6, no. 4, pp. 192–197, 1984.
[82] “Delaunay triangulation,” *Wikipedia*. 29-Oct-2018.
[83] “Fortune’s algorithm,” *Wikipedia*. 19-Sep-2018.
[84] S. Fortune, “A sweepline algorithm for Voronoi diagrams,” *Algorithmica*, vol. 2, no. 1, p. 153, Nov. 1987.
[85] I. Majdandzic, C. Trefftz, and G. Wolffe, “Computation of Voronoi diagrams using a graphics processing unit,” in *2008 IEEE International Conference on Electro/Information Technology*, 2008, pp. 437–441.
[86] X. Chen and S. McMains, “Polygon offsetting by computing winding numbers,” in *ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2005, pp. 565–575.
[87] A. F. Torabi, J. V. Ek, E. Champion, and J. Wang, “Micromagnetic Modeling Study of Thermal Gradient Effect in Heat-Assisted Magnetic Recording (HAMR),” *IEEE Trans. Magn.*, vol. 45, no. 10, pp. 3848–3850, Oct. 2009.
[88] T. W. McDaniel, “Application of Landau-Lifshitz-Bloch dynamics to grain switching in heat-assisted magnetic recording,” *J. Appl. Phys.*, vol. 112, no. 1, p. 013914, Jul. 2012.
[89] R. H. Victora and P. Huang, “Simulation of Heat-Assisted Magnetic Recording Using Renormalized Media Cells,” *IEEE Trans. Magn.*, vol. 49, no. 2, pp. 751–757, Feb. 2013.
[90] W. F. Brown, *Micromagnetics*. Interscience Publishers, 1963.

[91] M. Mansuripur and R. Giles, “Demagnetizing field computation for dynamic simulation of the magnetization reversal process,” *IEEE Trans. Magn.*, vol. 24, no. 6, pp. 2326–2328, Nov. 1988.
[92] W. F. Brown Jr, “Thermal fluctuations of a single-domain particle,” *Phys. Rev.*, vol. 130, no. 5, p. 1677, 1963.
[93] J. Barker, R. F. L. Evans, R. W. Chantrell, D. Hinzke, and U. Nowak, “Atomistic spin model simulation of magnetic reversal modes near the Curie point,” *Appl. Phys. Lett.*, vol. 97, no. 19, p. 192504, Nov. 2010.
[94] “CUDA C Programming Guide.” [Online]. Available: http://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html. [Accessed: 12-Nov-2018].
[95] A. Munshi, B. Gaster, T. G. Mattson, and D. Ginsburg, *OpenCL programming guide*. Pearson Education, 2011.
[96] R. D. Cideciyan, F. Dolivo, R. Hermann, W. Hirt, and W. Schott, “A PRML system for digital magnetic recording,” *IEEE J. Sel. Areas Commun.*, vol. 10, no. 1, pp. 38–56, 1992.
[97] K. S. Immink, P. H. Siegel, and J. K. Wolf, “Codes for digital recorders,” *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2260–2299, 1998.
[98] R. Potter, “Digital magnetic recording theory,” *IEEE Trans. Magn.*, vol. 10, no. 3, pp. 502–508, 1974.
[99] J. Moon and W. Zeng, “Equalization for maximum likelihood detectors,” *IEEE Trans. Magn.*, vol. 31, no. 2, pp. 1083–1088, 1995.
[100] H. K. Thapar and A. Patel, “A class of partial response systems for increasing storage density in magnetic recording,” *IEEE Trans. Magn.*, vol. 23, no. 5, pp. 3666–3668, 1987.
[101] H. Harashima and H. Miyakawa, “Matched-transmission technique for channels with intersymbol interference,” *IEEE Trans. Commun.*, vol. 20, no. 4, pp. 774–780, 1972.
[102] R. D. Yates and D. J. Goodman, *Probability and stochastic processes: a friendly introduction for electrical and computer engineers*. John Wiley & Sons, 2014.
[103] T. W. McDaniel, “Ultimate limits to thermally assisted magnetic recording,” *J. Phys. Condens. Matter*, vol. 17, no. 7, p. R315, 2005.
[104] K. Eason, H. T. Wang, M. R. Elidrissi, B. Xu, Z. Yuan, and K. S. Chan, “Recording Performance and Comparison of Graded-$T_c$and -K_uHAMR Systems,” *IEEE Trans. Magn.*, vol. 50, no. 3, pp. 107–113, Mar. 2014.
[105] O. N. Mryasov, U. Nowak, K. Y. Guslienko, and R. W. Chantrell, “Temperature-dependent magnetic properties of FePt: Effective spin Hamiltonian model,” *EPL Europhys. Lett.*, vol. 69, no. 5, p. 805, Feb. 2005.
[106] D. Hinzke, U. Nowak, R. W. Chantrell, and O. N. Mryasov, “Orientation and temperature dependence of domain wall properties in FePt,” *Appl. Phys. Lett.*, vol. 90, no. 8, p. 082507, Feb. 2007.
[107] Y. Jiao, Z. Liu, and R. H. Victora, “Renormalized anisotropic exchange for representing heat assisted magnetic recording media,” *J. Appl. Phys.*, vol. 117, no. 17, p. 17E317, 2015.

[108] X. Feng and P. B. Visscher, “Coarse-graining Landau–Lifshitz damping,” *J. Appl. Phys.*, vol. 89, no. 11, pp. 6988–6990, Jun. 2001.
[109] A. Y. Dobin and R. H. Victora, “Intrinsic Nonlinear Ferromagnetic Relaxation in Thin Metallic Films,” *Phys. Rev. Lett.*, vol. 90, no. 16, p. 167203, Apr. 2003.
[110] T. Qu and R. H. Victora, “Dependence of Kambersky damping on Fermi level and spin orientation,” *J. Appl. Phys.*, vol. 115, no. 17, p. 17C506, Jan. 2014.
[111] S.-K. Ma, *Modern theory of critical phenomena*. Routledge, 2018.
[112] J. Caroselli, S. A. Altekar, P. McEwen, and J. K. Wolf, “Improved detection for magnetic recording systems with media noise,” *IEEE Trans. Magn.*, vol. 33, no. 5, pp. 2779–2781, 1997.
[113] Y. Jiao, Y. Wang, and R. H. Victora, “A study of SNR and BER in heat-assisted magnetic recording,” *IEEE Trans. Magn.*, vol. 51, no. 11, pp. 1–4, 2015.
[114] H. Li and J.-G. Zhu, “Understanding the impact of Tc and Hk variation on signal-to-noise ratio in heat-assisted magnetic recording,” *J. Appl. Phys.*, vol. 115, no. 17, p. 17B744, 2014.
[115] K. J. Hole, H. Holm, and G. E. Oien, “Adaptive multidimensional coded modulation over flat fading channels,” *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1153–1158, 2000.
[116] R. Lynch, E. M. Kurtas, A. Kuznetsov, E. Yeo, and B. Nikolic, “The search for a practical iterative detector for magnetic recording,” *IEEE Trans. Magn.*, vol. 40, no. 1, pp. 213–218, 2004.
[117] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
[118] X. Wang, K. Gao, H. Zhou, A. Itagi, M. Seigler, and E. Gage, “HAMR Recording Limitations and Extendibility,” *IEEE Trans. Magn.*, vol. 49, no. 2, pp. 686–692, Feb. 2013.
[119] S. J. Greaves, Y. Kanai, and H. Muraoka, “Thermally Assisted Magnetic Recording at 4 Tbit/in^{2}},” *IEEE Trans. Magn.*, vol. 49, no. 6, pp. 2665–2670, Jun. 2013.
[120] Y. Jiao and R. H. Victora, “Dependence of Predicted HAMR Areal Density on Common Optimization Strategies,” *IEEE Magn. Lett.*, 2017.
[121] S. J. Greaves, H. Muraoka, and Y. Kanai, “Modelling of heat assisted magnetic recording with the Landau-Lifshitz-Bloch equation and Brillouin functions,” *J. Appl. Phys.*, vol. 117, no. 17, p. 17C505, Feb. 2015.
[122] B. Xu, H. Wang, Z. Cen, and Z. Liu, “4–5 Tb/in^{2}Heat-Assisted Magnetic Recording by Short-Pulse Laser Heating,” *IEEE Trans. Magn.*, vol. 51, no. 6, pp. 1–5, Jun. 2015.
[123] T. Rausch *et al.*, “Recording Performance of a Pulsed HAMR Architecture,” *IEEE Trans. Magn.*, vol. 51, no. 4, pp. 1–5, Apr. 2015.
[124] R. H. Victora and S. Wang, “Simulation of Expected Areal Density Gain for Heat-Assisted Magnetic Recording Relative to Other Advanced Recording Schemes,” *IEEE Trans. Magn.*, vol. 51, no. 11, pp. 1–7, Nov. 2015.
[125] J. Hohlfeld, P. Subedi, P. J. Czoschke, and M. Benakli, “Jitter in heat assisted magnetic recording,” *Proc TMRC*, pp. 1–2, 2017.

[126] K. Kim and J. Moon, “Experimental characterization of transition noise in HAMR,” *IEEE Trans. Magn.*, vol. 49, no. 7, pp. 3675–3678, 2013.
[127] N.-H. Yeh, X. Wu, and T. Roscamp, “Jitter analysis and cross-track correlation length extraction through nonlinear noise scaling with reader width,” *IEEE Trans. Magn.*, vol. 50, no. 11, pp. 1–4, 2014.
[128] B. F. Valcu, H. Gee, M. Alex, and E. Champion, “Relationship between equalized SNR and jitter—Theory and application to PMR and HAMR,” *IEEE Trans. Magn.*, vol. 50, no. 11, pp. 1–4, 2014.
[129] S. Hernández *et al.*, “Using ensemble waveform analysis to compare heat assisted magnetic recording characteristics of modeled and measured signals,” *IEEE Trans. Magn.*, vol. 53, no. 2, pp. 1–6, 2017.
[130] M. E. Wall, A. Rechtsteiner, and L. M. Rocha, “Singular value decomposition and principal component analysis,” in *A practical approach to microarray data analysis*, Springer, 2003, pp. 91–109.
[131] I. Jolliffe, “Principal component analysis,” in *International encyclopedia of statistical science*, Springer, 2011, pp. 1094–1096.