

THEORETICAL INVESTIGATION
OF NEW MAGNETIC RECORDING MEDIA
USING AN ENERGY LANDSCAPE METHOD

by

RU ZHU

PIETER B. VISSCHER, COMMITTEE CHAIR
JAMES W. HARRELL
SUBHADRA GUPTA
PATRICK LECLAIR
TIM MEWES
RICHARD H. TIPPING

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Physics and Astronomy
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2011

Copyright Ru Zhu 2011
ALL RIGHTS RESERVED

ABSTRACT

Magnetic material has played an important role in the information storage technology. We have worked out a phase diagram showing spin torque switching for perpendicular anisotropy media at nonzero temperature. A new area is predicted where the system can vibrate between precessional and parallel states quickly and therefore creates “telegraph” noise region. The result was later confirmed by a nanopillar experiment. The density of magnetic recording is restricted by the superparamagnetic limit. To overcome this limit people have used perpendicular recording to replace the traditional longitudinal recording and are working on other approaches to get over the superparamagnetic limit including heat assisted recording, patterned media and exchange coupled media. Anisotropy graded media provides an alternative solution to achieve high areal density while keeping the writing field relatively low and good thermal stability. A two-dimensional energy landscape of the switching of anisotropy graded media is calculated to give an intuitive way of understanding the switching, as well as a way to quantitatively calculate the switching rate by mapping the problem onto a one-dimensional Langevin random walk. Then an iterative algorithm of for finding the saddle point of the energy landscape, which we call barbell algorithm, is used to locate the saddle point and the whole switching trajectory.

DEDICATION

This thesis is dedicated to my father Yuyang Zhu, and my mother Mingcui Liu.

LIST OF ABBREVIATIONS AND SYMBOLS

LLG	Landau-Lifshitz-Gilbert equation
MTJ	Magnetic Tunnel Junction
K	uniaxial anisotropy constant
\mathbf{H}	applied external magnetic field
\mathbf{M}	magnetization
M_s	saturation magnetization
B	Slonczewski's dimensionless torque asymmetry parameter
J	Slonczewski's current parameter
m	magnetic moment
m_i	i^{th} component of magnetic moment
m_x	x component of magnetic moment
m_s	saturation magnetic moment
m_x^b	x component of magnetic moment at the energy barrier
h	reduced magnetic field
H_K	anisotropy field
E	Energy
B	Slonczewski's dimensionless torque asymmetry parameter
θ	polar angle
ϕ	azimuthal angle

γ	gyromagnetic ratio
η	spin-polarization factor
H	Hessian matrix
λ	damping coefficient
ρ	population density
W	domain wall width factor
D	diffusivity
r	switching rate
r_0	switching rate prefactor
k_B	Boltzmann's constant
T	temperature
V	volume of each computational cell
L	grain length
V_{tot}	total volume of magnetic grain
E_b	magnetic energy barrier
H_c	coercivity
α	damping constant
γ	gyromagnetic ratio
$\langle \rangle$	time average
\mathfrak{M}	magnetization configuration of entire system

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my advisor, Dr. P. B. Visscher for his instructive guidance during all the period of my Ph.D study. I believe this kind of experience will be beneficial to my future career.

I would like to thank Dr. S. Gupta, Dr. J. W. Harrell, Dr. T. Mewes, Dr. P. LeClair, and Dr. R. Tipping for serving as the dissertation committee.

I would like to thank Dr. W. H. Butler for his great guidance in theoretical solid state physics.

I would like to thank Dr. G. Mankey for teaching me vacuum science and various experimental techniques.

I would like to thank to Dr. S. Sarker, for his kindly and supportive suggestions in my research.

I would like to thank to all faculty members of Department of Physics and Astronomy for their teaching and help.

I would like to thank the Department of Physics and the MINT center for providing financial assistance in the form of tuition and stipend through the years.

Thanks are also due to James Weston and John Hawkins, former and current facility manager, for providing technical support and training me on various instruments.

I would also like to thank Jason Foster for the computer support.

Thanks are also given to all the wonderful office staff, Casey, Jamie and Karrie, for streamlining travel arrangements and the tactful handling of administrative issues.

I would like to acknowledge several colleagues and former UA students, such as Dymtro Apalkov, Jin He, Chunsheng Liu, Michael Wofsey, Krishna Chetry, Sreenivasulu Vutukuri, Zhihong Lu, Shuxia Wang, Tianyi Xu and Lei Wen, for friendly help and discussion.

I would like to thank my parents for their constant love. Any part of my achievement is direct product of their support.

CONTENTS

ABSTRACT.....	ii
DEDICATION.....	iii
LIST OF ABBREVIATIONS AND SYMBOLS.....	iv
ACKNOWLEDGMENTS.....	vi
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
1. INTRODUCTION.....	1
2. SPIN TORQUE SWITCHING IN PERPENDICULAR FILMS AT FINITE TEMPERATURE.....	4
a. Spin Torque Magnetic Random Access Memory.....	4
b. Spin Torque Switching in Perpendicular Films at Finite Temperature.....	7
3. SWITCHING RATE CALCULATION OF ANISOTROPY GRADED MEDIA.....	15
a. The Bottleneck of High Density Magnetic Recording – Superparamagnetic Limit.....	15
b. Possible Solutions to Superparamagnetic Limit.....	18
c. One Dimension Switching Problem – Langevin Equation.....	21
d. Methods for Calculating Minimum Energy Path.....	28
e. Convergence of Micromagnetic System with Constraints – Generalized Newton-Raphson Method.....	30

f. One dimensional application of Newton-Raphson Algorithm: The 1D Energy Landscape	35
g. Two-dimensional Energy Landscape – a Semi-analytical Way to Describe Energy Barrier in Switching.....	42
h. Barbell Algorithm – a 3N Dimensional Algorithm to Find the Saddle Point in Switching.....	49
i. Switching Rate Calculation Results.....	54
4. MICROMAGNETIC MODELING OF PERMANENT MAGNET NANOSTRUCTURES	58
a. History and Modern application of Permanent Magnets	58
b. Modeling of MnAl-FeCo Two Phase Nanostructured Permanent Magnet	60
c. Energy Landscape Interpretation of Permanent Magnets’ Switching Mechanism	63
5. SUMMARY AND POSSIBLE FUTURE WORK	67
REFERENCES	68

LIST OF TABLES

1 Parameters of 2D energy landscape fitting	49
2 List of Parameters of Permanent Magnets	59

LIST OF FIGURES

1 Magnetic Random Access Memory (MRAM)	5
2 A GMR memory cell	5
3 Four states of the cell magnetization with different resistance	5
4 Spin Torque MRAM	7
5 Magnetic Tunnel Junction (MTJ)	8
6 Dynamics of magnetization in uniaxial symmetry	9
7 Phase Diagram at Zero Temperature	11
8 Phase diagram at nonzero temperature	13
9 Longitudinal Recording	16
10 Perpendicular recording	16
11 A contour plot of counterexample	29
12 Illustration of the Newton-Raphson method	34
13 Anisotropy graded media	35
14 Anisotropy as a function of x	36
15 Linear extrapolation of m_{x_seek}	40
16 Energy landscape generated by sweeping method	41
17 Energy paths at different sweeping speed	42
18 A demonstration of $Q_{xx} = 0$	43

19 A demonstration of $Q_{xx} = \text{maximum}$	43
20 System magnetizations at different stages of switching	47
21 System magnetizations at different stages of switching	48
22 Illustration of barbell algorithm	50
23 Illustration of M_1 , M_2 and r	51
24 “Barbell algorithm” to find the energy saddle point	53
25 The convergence of barbell algorithm to energy minimum	55
26 Energy landscape obtained by barbell algorithm	55
27 Energy landscape for different layered systems	56
28 Thermal switching rate prefactor as function of number of layers	58
29 The demagnetization curve in M-H loop and maximum $ BH _{\text{max}}$	60
30 MnAl-FeCo two-phase nanostructured permanent magnet	60
31 M-H loop of MnAl-FeCo permanent	62
32 BH curve of MnAl-FeCo permanent magnet	62
33 Two-layer soft-hard composite permanent magnet nanostructure	64
34 Coherent switching: Energy contours for a bilayer	65
35 Domain wall switching in two layered permanent magnet nanostructure with a moderate exchange constant	66

CHAPTER 1 INTRODUCTION

This dissertation describes several theoretical research projects related to different aspects of magnetic information storage.

In chapter 2, we discuss work we did on spin torque systems, with focus on the conditions for existence of precessional states and a new phase region called “telegraph noise region” observable at finite temperature, rather than exclusively on switching.

In Chapter 3, we discuss the main topic of the dissertation, the development of an energy landscape technique for understanding and visualizing switching in magnetic systems, particularly in exchange coupled magnetic grains. This not only allows the visualization of switching paths, but makes it possible to quantitatively calculate switching rates. This is very important in the case of very slow rates (thermal stability) for which it is not possible to calculate rates by direct simulation because of the long times involved. Direct simulations can only be done for artificially low barriers. There were reports around 2007 that the resulting rates were much faster than expected for coherent switching over similar energy barriers, casting doubt on the stability of exchange coupled or anisotropy graded media. We have succeeded in calculating rates for reasonable barriers and showing that this is not the case - rates are comparable to those of conventional media with the same barriers.

We introduce a new way to approach the problem of rate calculation in magnetic systems that has a number of advantages over traditional ones. Analytic or semi-analytic solution of the Fokker-Planck equation, such as was done by Brown[1] for the case of a uniaxial particle, is possible only for very simple systems. The general Hessian-matrix approach[2] is computationally demanding and gives little heuristic insight into what physical factors control the rate. The most physically transparent approach to thermal rates is that of Kramers[3], which is however only directly applicable to a one-dimensional problem, basically the one-dimensional Langevin equation.

So the approach we will take is to map our $2N$ -dimensional system (N is the number of computational cells, each of which has 2 degrees of freedom assuming the saturation magnetization is constant) into a 1D system governed by a Langevin equation, and compute the switching rate of that 1D system by the method of Kramers. The map describes motion along a one-dimensional path in $2N$ dimensional space, the switching path. This requires us to specify this path, and also to specify which points in $2N$ -dimensional space map onto a particular point P of the path. Once we have the path, we can define the set of points mapping to P as the $2N-1$ dimensional hyper-plane perpendicular to the path at P . This allows us to map an actual trajectory of the real $2N$ -dimensional system onto a 1D trajectory, and to derive a 1D Langevin equation that describes time-dependent motion along that trajectory. This is slightly more general than Langevin's original system, in that the mobility and diffusivity (which are constant in a textbook Langevin equation) are functions of position in our Langevin equation. They can be calculated from the configurations along the switching path – thus we need to define this path precisely, and calculate the configurations along it explicitly.

We will define the switching path as the path of steepest descent from the saddle point to the initial or final minimum-energy configuration. This is sometimes referred to as the minimum-energy path – various methods have been proposed to calculate it, such as the nudged-elastic band method[4]. We have found a computationally simpler method, which is simply to use the Landau-Lifshitz equation of motion (Eq. 2.1) with infinite damping (this is done by letting $\gamma \rightarrow 0$ while keeping $\alpha\gamma$ constant). Physically, this eliminates the precession term and causes the system to move along the energy gradient in $2N$ -dimensional configuration space. To do this, we must identify the saddle point – we discuss an efficient method for doing this in chapter 3.

Visualizing the switching path in the $2N$ -dimensional configuration space is beneficial to physical insight into the behavior of important quantities such as the easy-axis magnetic dipole moment m_x , by computing the minimum energy with one or more such moments constrained. We will discuss a one-dimensional version of such a constrained-minimum energy landscape and a two-dimensional one in chapter 3.

Finally, Chapter 4 discusses some preliminary applications of the energy landscape method to permanent magnet nanostructure.

CHAPTER 2 SPIN TORQUE SWITCHING IN PERPENDICULAR FILMS AT FINITE TEMPERATURE

a. Spin Torque Magnetic Random Access Memory

The fast advancing of nanotechnology has led to new applications of magnetic materials in spin electronic devices, magnetic sensors and functional materials in the recent years. One development in magnetic recording industry is magnetic random access memory (MRAM, see Fig. 1,2,3). Ordinary field switching MRAM uses tunneling magnetoresistance (TMR) effect, which was discovered in the 1970s[5]. An MRAM element is made up of two magnetic layers with different coercivities. They are separated by a non-magnetic interlayer. The resistance of the MRAM memory cell depends on the magnetization direction of the two magnetic layers. According to the magnetization orientations of the two ferromagnetic layers, this system has four states of magnetization as shown in Fig. 3, while only parallel state and anti-parallel state are needed. Thus, different magnetization direction of the sensing layers and storage layers can be used to represent bit information. Data writing can be done by applying a sufficiently large current, and use the field generated to switch the magnetization of the storage layer. Practically a selective writing operation is done by the vector sum of the H_{word} and H_{sense} . Only the memory cell at the cross-point of selected word and bit is applied by a sufficiently large switching field, while unselected cells are not switched. On the other hand, the data reading is done by applying a relatively small sensing field to flip the magnetization of the sensing layer, thus generating a rising or falling resistance change according to the stored data.

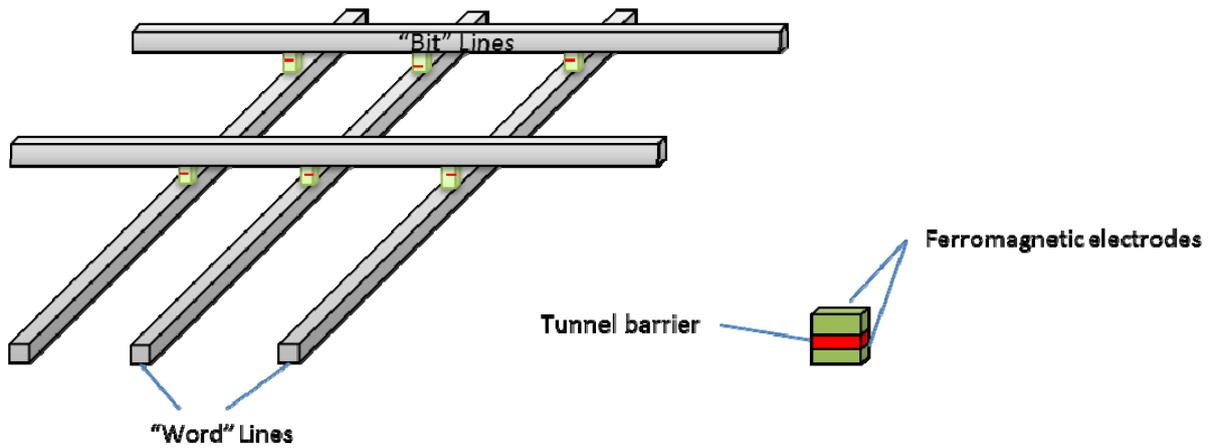


Figure 1 Magnetic Random Access Memory (MRAM). The tunnel barrier separates two ferromagnetic layers, and the relative magnetization change in these two layers causes change in resistance of the junction. Only the memory cell at the cross-point of selected word and bit will be switched.

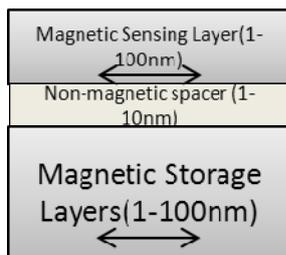


Figure 2 A GMR memory cell.

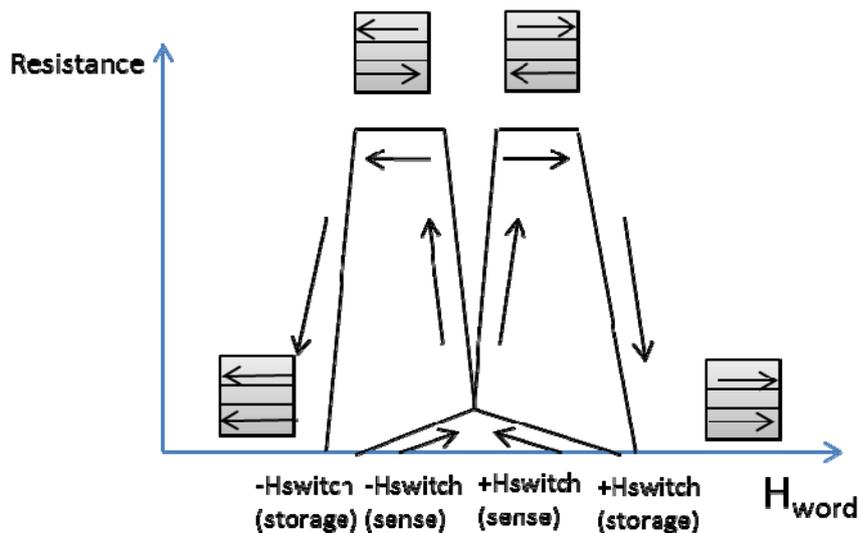


Figure 3 Four states of the cell magnetization with high and low resistances. A selective write operation is done by the vector sum of the H_{storage} and H_{sense} . The data read is done by applying a relatively small sensing field to flip the magnetization of the sensing layer and then giving a

rising or falling resistance change according to the storing data, which can be seen by the steep slope around the origin point.

Compared with traditional MRAM, spin-torque transfer MRAM (STT-RAM, Fig. 4, 5) has advantages due to the reduction of switching current and elimination of half-select problem.

The working mechanism of STT-RAM is described below. Electrons have charges and spins. In non-ferromagnetic materials the spins are randomly oriented and do not affect the behavior of devices significantly. However, if ferromagnetic components are incorporated into devices, transferring electrons are partially spin polarized and can change the behavior of the devices, due to spin-based interactions between electrons and ferromagnets. More specifically, the spin polarized electrons can change the orientations of the magnetizations. This effect is called spin transfer torque. Using this effect as a switching scheme in a new generation of STT-RAM has been proposed and studied. As mentioned previously, STT-RAM gives many advantages over field switching MRAM, including eliminating half-select problem, increasing the bit density and improving scalability[6].

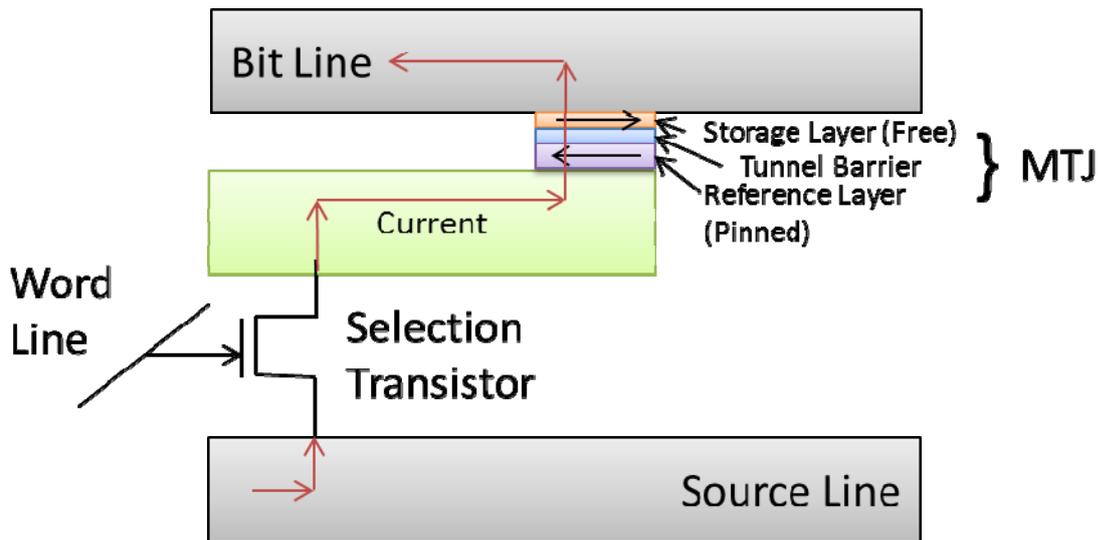


Figure 4 Spin Torque MRAM bit. Current flows perpendicularly through Magnetic Tunnel Junction (MTJ).

b. Spin Torque Switching in Perpendicular Films at Finite Temperature

The core component of STT-MRAM is a Magnetic Tunnel Junction (MTJ) as shown in Fig. 5. It comprises two ferromagnetic layers and an ultrathin non-ferromagnetic interlayer between them. One of the ferromagnetic layers is a reference layer which has a “pinned” magnetization direction due to its larger thickness or its contact with an antiferromagnetic layer. The role of the reference layer is to polarize the current passed through the junction. The other ferromagnetic layer is a free layer whose magnetization can be more easily switched. If one passes polarized current through the junction, depending on the magnitude of passing current and its polarity, a steady precession or even switching of magnetization in the free layer can occur[7].

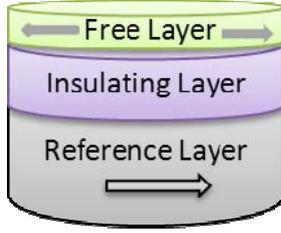


Figure 5 Magnetic Tunnel Junction (MTJ).

Some discrepancies have been shown in various experiments[8, 9], which may result from not taking into account the thermal effect.

The behavior of magnetization in the free layer is governed by Landau-Lishiftz-Gilbert (LLG) equation with spin torque term[10]:

$$\frac{d\mathbf{M}}{dt} = -\gamma \mathbf{M} \times \mathbf{H} - \frac{\alpha\gamma}{M_s} \mathbf{M} \times \mathbf{M} \times \mathbf{H} + \frac{J}{1+B \cos \theta} \mathbf{M} \times \mathbf{M} \times \mathbf{M}_p, \quad (2.1)$$

where γ is the gyromagnetic ratio (Fig. 6), M_s is the saturation magnetization and \mathbf{H} is the magnetic field. The second term (proportional to the LL damping parameter α) controls damping and pushes magnetization (\mathbf{M}) toward the easy axis, and the second (spin torque) term scales

with a parameter $J = -\frac{\gamma\eta I\hbar}{2eVM_s^2}$ proportional to the polarized current and spin torque transfer

efficiency (where η is the spin-polarization factor, I is the spin-polarized current, \hbar is the Planck constant, e is the electronic charge, and V is the volume of the magnetic grain), and B which is Slonczewski's dimensionless torque asymmetry parameter[7, 11].

In the following paragraphs we will first derive the spin torque switching diagram at zero temperature, and then investigate the change of phase diagram if the temperature is nonzero. The phase diagram for STT-MRAM has previously been studied theoretically at zero temperature[12]. We tried to investigate a single-domain model at nonzero temperature, which was an attempt to

explain the existence in the current-field phase diagram of regions in which precessional states are stable.

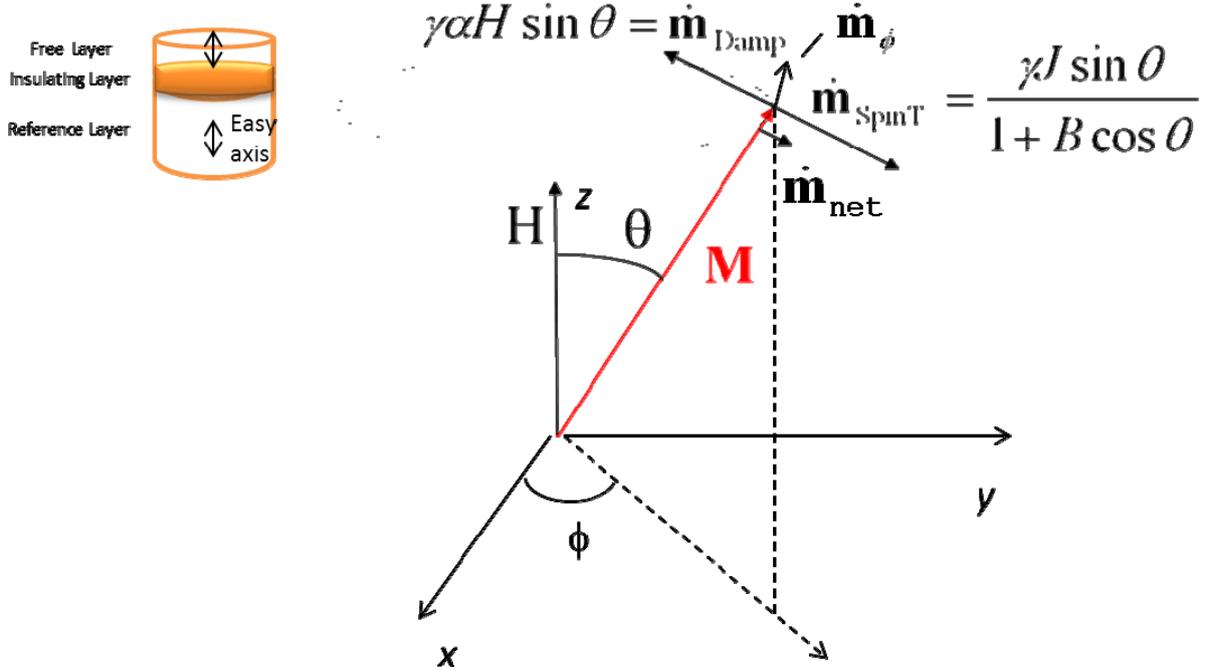


Figure 6 Dynamics of magnetization in uniaxial symmetry. Damping term and spin torque term are non-conservative. θ is the angle between applied field (in z direction) and magnetization. ϕ is the azimuthal angle from the x axis.

In uniaxial symmetry shown in Fig. 6, the LLG equation has a precession term which contributes only to the azimuthal component \dot{M}_ϕ and does not change the energy; the changes in energy are controlled by the θ component of time derivative of the magnetization. In the geometry illustrated in Fig. 6, the total field applied on magnetization is $H_K^{eff} \cos \theta + H_e$. According to LLG equation:

$$\dot{M}_\theta = M_s \dot{\theta} = -\gamma \sin \theta \left\{ \alpha M_s [H_K^{eff} \cos \theta + H_e] + \frac{JM_s}{1 + B \cos \theta} \right\}, \quad (2.2)$$

where $H_K^{eff} = H_K - M_s$ is the effective anisotropy and H_e is an external field along the easy axis.

Defining an effective energy in the presence of spin torque is nontrivial since only the precession term in the Landau–Lifshitz equation conserves energy; the damping term

$$\dot{\mathbf{m}}_{Damp} = -V_{tot} \frac{\alpha\gamma}{M_s} \mathbf{M} \times \mathbf{M} \times \mathbf{H} \text{ and spin torque term } \dot{\mathbf{m}}_{SpinT} = V_{tot} \frac{\gamma J}{1 + B \cos \theta} \mathbf{M} \times \mathbf{M} \times \mathbf{M}_p \text{ in Fig. 6}$$

are non-conservative, where V_{tot} is the total volume of the magnet. A rigorous way to derive an effective energy is by finding a steady state solution of the Fokker–Planck equation, which can be done using the method described by Visscher[13], giving a probability distribution of the form $\exp(-E_{eff} / k_B T)$, where

$$E_{eff} = M_s \left[\frac{1}{2} H_K^{eff} (1 - \cos^2 \theta) - H_e \cos \theta - \frac{J}{\alpha B} \ln(1 + B \cos \theta) \right], \quad (2.3)$$

It is worth noting, however, that Eq. 2.3 can be obtained heuristically; in this special case of uniaxial symmetry, if we compute the “work” done against the torque (Eq. 2.2), the result is proportional to the effective energy: $\int \dot{M}_\theta d\theta = \alpha\gamma E_{eff}$, or a non-dimensional

$$\text{form } e_{eff} = E_{eff} / \mu_0 H_K^{eff} M_s = \frac{1}{2} (1 - u^2) - h_e u - \frac{j}{B} \ln(1 + Bu),$$

where $u = \cos \theta$, $h = H_e / H_K^{eff}$, and $j = J / \alpha H_K^{eff}$.

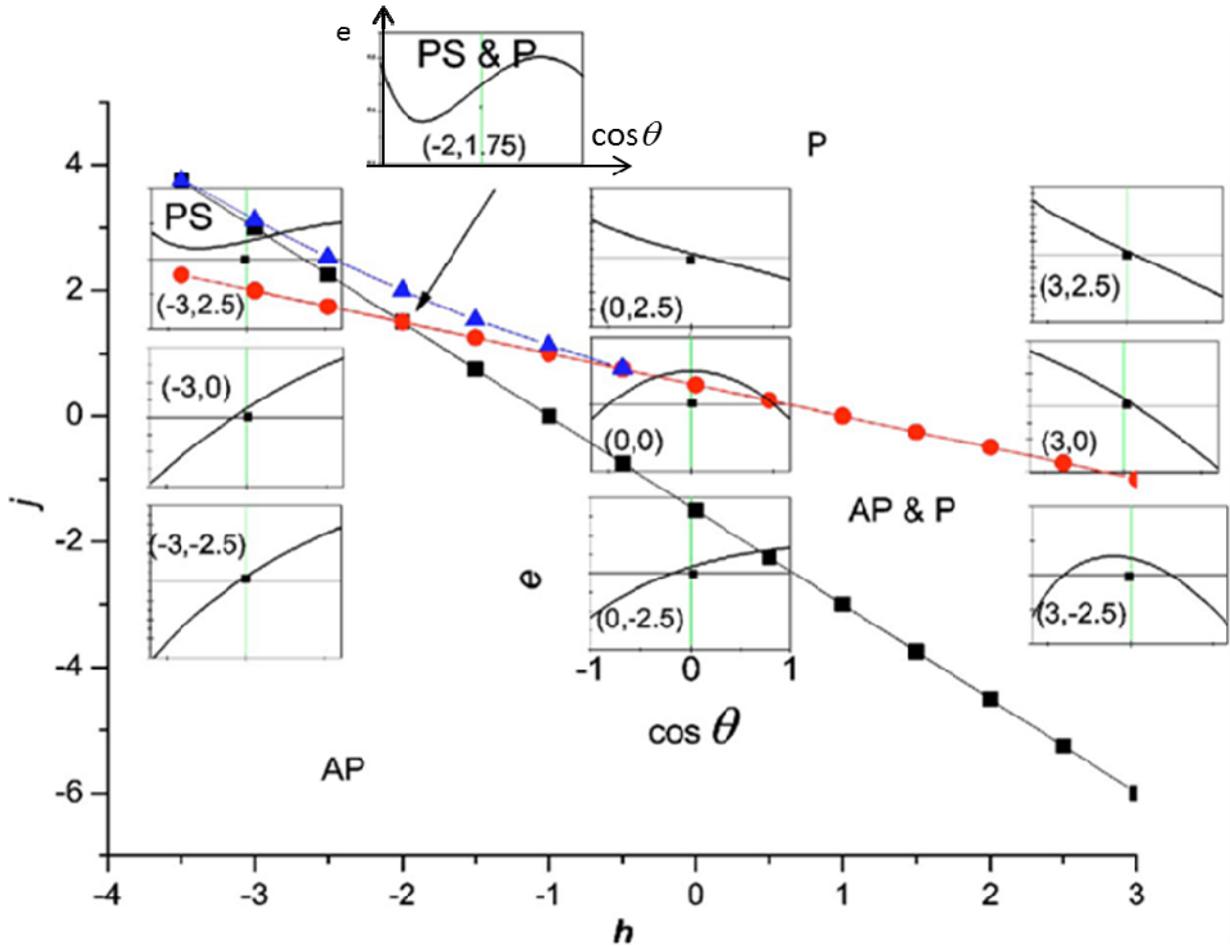


Figure 7 Phase Diagram at Zero Temperature. The y axes of insets are reduced energy (e in text), and the x axes are $\cos\theta$, as shown in the bottom center inset.

At zero temperature the phase diagram is shown in figure 7. The only free parameter is the Slonczewski parameter B , which we have taken to be 0.5 in the figures. The horizontal and vertical axes are the dimensionless magnetic field and current, h and j , and at each point of a 3 by 3 grid there is an inset graph showing the effective energy as a function of $u = \cos\theta$, at the field and current corresponding to the center of the inset. Lines with symbols are phase boundaries. The center graph, at zero field and current, shows only the anisotropy energy $-u^2/2$, a parabola. At this point, there are two stable states (minima of $e(u)$ at $u = \pm 1$ - note that $e(u)$ need not be flat at minima that lie at the boundaries of the physical region $-1 < u = \cos\theta < 1$. Thus, in this

region of the phase diagram both the parallel ($\cos \theta = 1$) and antiparallel ($\cos \theta = -1$) states are stable, so we have labeled it AP & P.

If we move to the right from the center of Fig. 7, we add the Zeeman term $-hu$, which simply shifts the parabola to the left so its maximum is outside the physical region, and there is only one (parallel) minimum at $u = 1$, and we have crossed a phase boundary (red circles) to the P region where only the parallel state is stable. At this boundary, the maximum is just leaving the physical region at $u = -1$, i.e., $de(u = -1) / du = 0$. On the left, at $h = -3.0$, the parabola shifts in the opposite direction (right) - the maximum passes out of the physical region at the black (square symbols) phase boundary, where $de(u = 1) / du = 0$, and only the AP state is stable.

Moving up from the center of Fig. 7 (increasing the scaled current to $j = 2.5$), the effective-energy inset graph includes the logarithmic spin-torque term $-\ln(1 + Bu)$. This has a divergence at $u = -1/B = -2$, outside the physical region, but starts to rise at the left as seen in the top center inset. This mimics the effect of a positive field (right inset) and brings us across the phase boundary to the upper-right single-phase P region.

At the upper left inset (negative field $h = -3.0$), the effect of the field (lowering e at the left) and the spin torque (raising e at the left) oppose each other, and the spin torque can raise the energy at the left enough to create a minimum away from the $u = \pm 1$ boundaries, physically corresponding to a precessional state - thus this region of the phase diagram is labeled "PS". The effective energy in the remaining small sliver of the phase diagram, between the curved line (blue triangles) and the straight phase boundaries, is shown as a tenth inset—both the

precessional and parallel states are stable, so this region is labeled “PS & P”. Making the current negative (bottom center inset) causes a decrease in energy at the left side of the inset, mimicking a negative field. The only minimum is now the AP (left) minimum, and the system is below the black (squares) phase boundary in the AP region. At the lower left, adding a real negative field further stabilizes the AP state. Moving to the lower right (positive field) restores the $u = 1$ (parallel) minimum and returns us above the black phase boundary to the AP & P region. But some discrepancies have been shown in various experiments[8, 9], which may result from not taking into account the thermal effect.

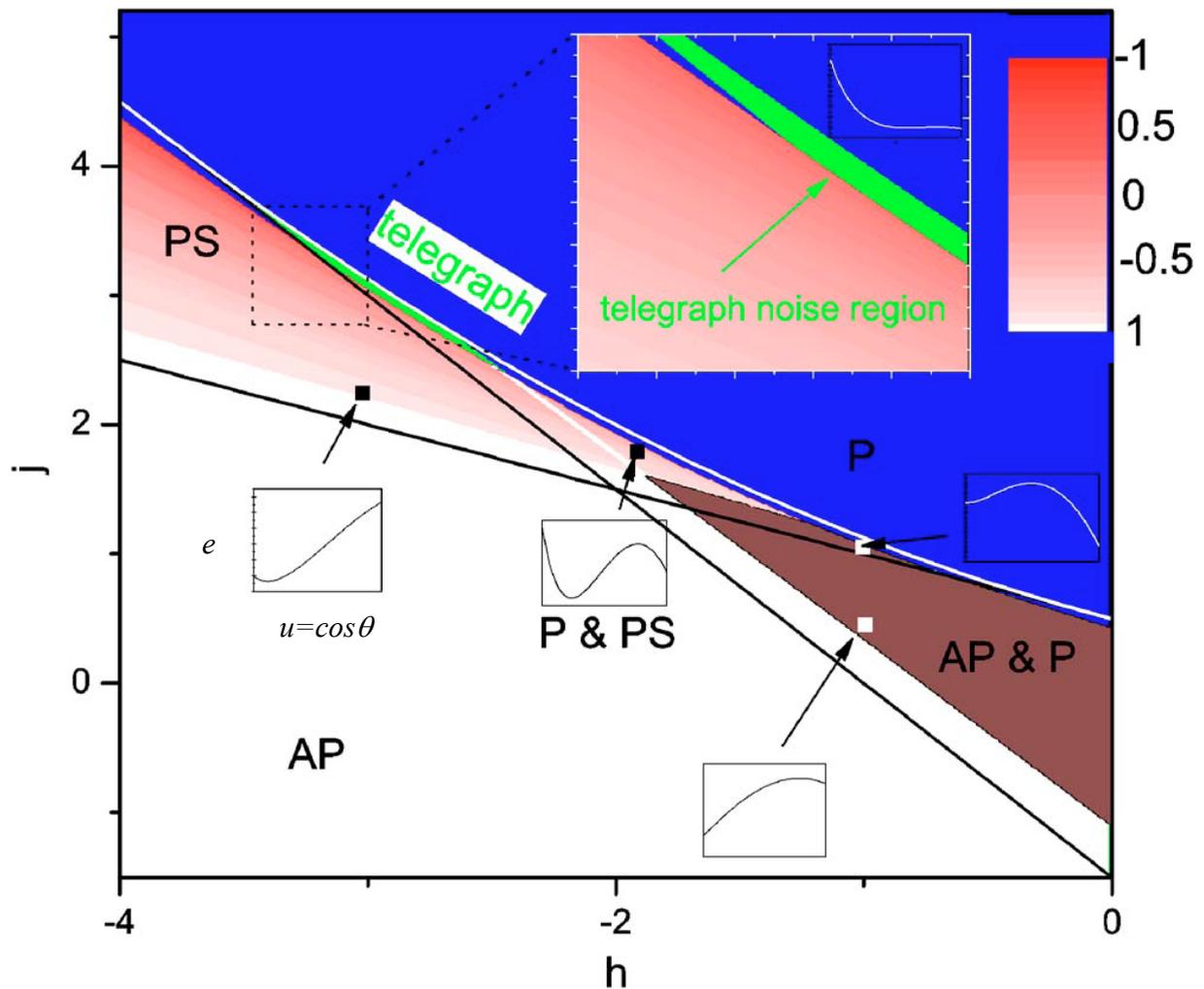


Figure 8 Phase diagram at nonzero temperature.

At nonzero temperature, the system will not remain in a well with a very low barrier. Fig. 8 shows the effect of adding nonzero temperature in the phase diagram. Zero-temperature boundaries are indicated by solid black or white lines. Color scale from white to red in the PS (precessional state) region and the coexisting parallel and precessional (P & PS) region indicates $\cos \theta$ for stable precession. We have colored part of the P region green to indicate that it is a physically interesting region - both barriers are lower than the critical value, so the system jumps between the precessional and the parallel states rapidly on the experimental time scale and causes random switching of resistance, which is called telegraph noise. The inset at top shows details of part of this telegraph noise region. Graphical insets show $e(u)$ energy landscapes at the indicated points.

To conclude, by investigating the energy landscape of perpendicular anisotropy metallic tunneling junction, we found a telegraph noise region that has not been discovered before. Telegraph noise in these systems has been reported, though attributed to domain wall motion rather than the precessional-antiparallel alternation we found[14].

CHAPTER 3 SWITCHING RATE CALCULATION OF ANISOTROPY GRADED MEDIA

a. The bottleneck of High Density Magnetic Recording – Superparamagnetic Limit

Magnetic recording plays an important role in information storage technology. In history, magnetic storage media existed in the forms of magnetic core memory, magnetic tape, magnetic bubble memory and hard disk drive (HDD). The basic mechanism of magnetic information storage is the magnetization orientation in magnetic media. Due to scalability and cost problems HDD and magnetic tape are almost the only magnetic storage mediums still in use today. The areal density of HDD (usually in bits/inch²) has been fast increasing. Before 2005, in most commercial hard disks the information bit is stored parallel to the surface of disk, which is known as longitudinal recording (Fig. 9). In the context of longitudinal recording reducing the size of a bit, or equivalently increase the areal density, leads to stronger demagnetization field at the transition between bits and increases thermal instability. In 2005 Hitachi Global Storage realized perpendicular recording and made the areal density as large as 230 Gbits/in². In perpendicular recording, the bit information is stored with magnetization perpendicular to the media surface. With the increasing areal density, longitudinal recording requires smaller grain size and thinner media to reduce the demagnetization field. Even with these techniques taken into account, the coercivity field of magnetic material still went from several hundred Oe to over 3kOe. In perpendicular recording (Fig. 10), demagnetization field reduces as the grain size gets smaller. Plus, it also gives higher write field amplitude, higher write field gradients and well aligned recording layers, zero demagnetization field at bit transitions and higher playback amplitude.

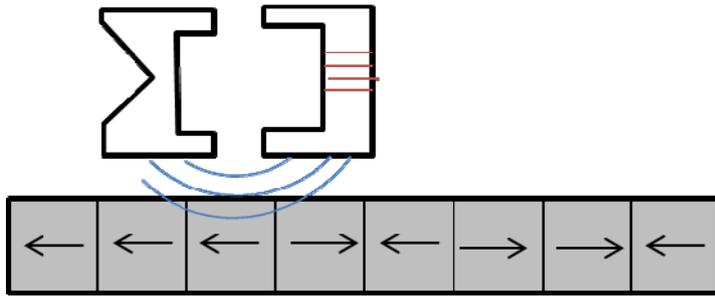


Figure 9 Longitudinal Recording. Magnetization orientations are in plane and can be switched by magnetic field from air gap of write head.

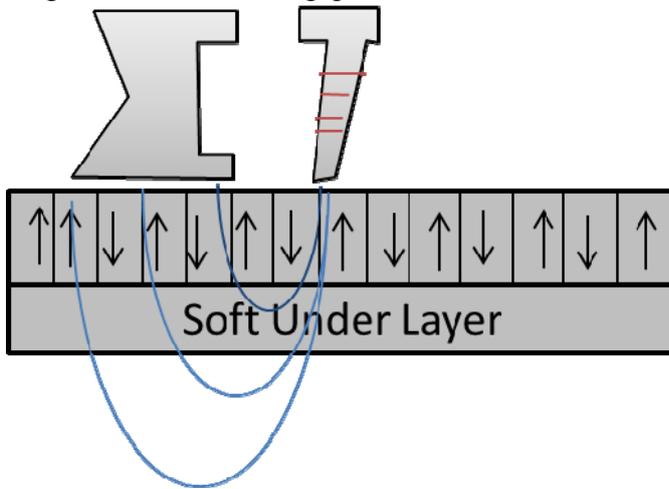


Figure 10 Perpendicular recording. Magnetization orientations are out of plane. The write field is enhanced by soft under layer due to mirror effect.

In magnetic recording, the “1” and “0” bits are represented by two magnetization states (spin up or spin down) of magnetic grains with an energy barrier laying between them. To switch from one state to the other, sufficient energy is required to conquer the energy barrier. In conventional magnetic recording, the state can be switched by applying a large enough external field. On the other hand, if the energy barrier is too low, thermal fluctuation may cause the grains to switch between states at room temperature and lead to the loss of recorded information.

The thermal stability can be expressed by the Néel-Arrhenius equation [11, 12]:

$$\frac{1}{\tau} = r = r_0 e^{-E_b/k_B T} . \quad (3.1)$$

where τ is the relaxation time, r_0 is called the attempt frequency which is of the order of the reciprocal of its gyromagnetic resonance frequency, i.e. $\sim 10^{10}$ Hz[1], E_b is the energy barrier between spin up and spin down state, k_B is the Boltzmann constant and T is the temperature. To make the material of practical use in industry, we need to make sure $\tau \geq 10$ years, corresponding to $E_b \approx 40k_B T$.

For a Stoner-Wohlfarth particle, the energy barrier is proportional to its magnetic anisotropy.

$$E_b = KV , \quad (3.2)$$

where K is the magnetic anisotropy[15] of the grain and V is the volume of the grain. Thus, for any given material (whose K is constant), the energy barrier decreases with the decrease of grain size. If the grain is too small it becomes thermally unstable. The critical size is called the superparamagnetic limit for the material.

Thus to increase the areal density and keep good thermal stability, the only way is to increase anisotropy. However, as the magnetic material gets harder, it is difficult to find a write head to write on it, because the theoretical maximum of a magnetic field that can be achieved depends on M_s of the head material, and is about 2.3 T[16]. On the other hand, for conventional thin film media, noise mainly comes from the transition region between two neighboring bits due to the interactions among grains[17, 18]. The Signal-to-Noise Ratio (SNR) of the reproduced signal is proportional to the number of grains in the recording region[19, 20]. Thus as the areal density increases, the size of the grains must decrease in order to maintain the SNR.

b. Possible Solutions to Superparamagnetic Limit

To overcome the superparamagnetic limit people have proposed several methods including heat assisted recording[21, 22], patterned media[23-26] and exchange coupled media[27, 28].

Heat assisted recording combines thermo-magnetic writing and magnetic reading. To increase the stability of the recorded information on extremely small grains, writing is done at an elevated temperature on the recording medium with a very high coercivity at room temperature, using a write head with an integrated light path incorporated into it. The media is locally heated by laser light through the light path during writing, and the coercivity of the medium decreases temporarily to a value such that recording with high saturation write heads is possible. With this technique, a smaller transition width and higher density or a better signal-to-noise ratio can be obtained than those achievable by conventional recording.

Patterned media stores data in an array of single-domain magnetic particles. It consists of a regular array of magnetic elements, each of which has uniaxial magnetic anisotropy. Different from the traditional thin film medium, the grains within each patterned element are strongly coupled so that the entire element behaves as a single magnetic domain, and equivalently increases the volumes of magnetic domains. This scheme can substantially overcome the superparamagnetic limit, and can enable higher recording densities. However, successful implementation of patterned media requires fabrication of sub-50-nm features over large media

areas and new design of recording systems which differ substantially from those used in conventional hard drives.

In this dissertation we study a different possible solution to the superparamagnetic limit, sometimes called composite or exchange-coupled media. The advantage of exchange coupled media can be described on a heuristic level. A composite perpendicular recording media has magnetically hard and soft regions within each grain. The hard region effectively sustains thermal stability of the whole grain when external field is not present. By applying an external field, one can cause the magnetization of the soft region to rotate and, thus, change the angle of the effective field applied to the hard region. This effectively applies an exchange field to the hard layer and makes it more likely to switch. Thus the external field necessary to switch the system is decreased.

Anisotropy graded media were suggested as a generalization of two-layer exchange-coupled media[28] to N layers by Suess[29]. In anisotropy graded media, magnetic multilayers are deposited in such a way that the anisotropy of each layer gradually increase from magnetic soft layer to hard layer. It has been proved as a general theoretical result (not dependent on any specific model) that, the figure of merit suggested by Victora (ratio of thermal switching energy barrier to coercivity, normalized to that of a Stoner-Wohlfarth particle) is rigorously less than 4[30] but better than that of exchange coupled media (about 2)[28]. In thin domain wall limit, the limit of 4 in figure of merit can be achieved if the anisotropy of n^{th} layer is proportional to n^2 [30]

As mentioned before, it is important to know the thermal stability of any media to decide if it is industrially applicable. There were reports around 2007 that the resulting thermal switching rates were much faster than expected for coherent switching over similar energy barriers, casting doubt on the stability of exchange coupled or anisotropy graded media.

To calculate the thermal switching rate, several methods have been proposed. One of them is to use the Hessian matrix of saddle point and energy minima to predict the attempt frequency[2].

Hessian matrix includes all the mixed second derivatives of the energy with respect to the magnetizations of the cell described by spherical coordinates (θ, ϕ) .

$$H = \begin{bmatrix} \frac{\partial^2 E}{\partial \theta_1^2} & \dots & \frac{\partial^2 E}{\partial \phi_n \partial \theta_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 E}{\partial \theta_1 \partial \phi_n} & \dots & \frac{\partial^2 E}{\partial \phi_n^2} \end{bmatrix}$$

Since the phase space volume can be represented by product of Hessian matrix eigenvalues, the attempt frequency prefactor is given by [31]

$$r_0 = \sqrt{\frac{\det H|_{\min}}{|\det H|_{sp}}}$$

The disadvantage of this method is that it is computationally expensive because it involves second derivatives in $2N$ dimensional space. It also involves approximations whose validity is hard to estimate because the derivation is quite formal.

In this paper we will propose a method that is both computationally efficient and physically intuitive, by projecting the multidimensional energy landscape onto one dimensional energy

landscape, and using mathematical tools such as Fokker-Planck equation to solve it semi-analytically.

c. One dimensional switching problem – Langevin Equation

The original Langevin equation[32] describes the apparently random motion of a small particle immersed in a fluid due to collisions with fluid molecules, which is called Brownian motion.

$$m \frac{d^2 \mathbf{x}}{dt^2} = -\lambda \frac{d\mathbf{x}}{dt} + \boldsymbol{\eta}(t). \quad (3.3)$$

where m denotes the particle's mass. The force acting on the particle is written as a sum of a viscous force proportional to the particle's velocity with damping coefficient λ , and a noise term named $\boldsymbol{\eta}$ representing the effect of the random collisions with the fluid molecules. The time variation $\boldsymbol{\eta}(t)$ has a Gaussian probability distribution with correlation function

$$\langle \eta_i(t) \eta_j(t') \rangle = 2\lambda k_B T \delta_{i,j} \delta(t - t'), \quad (3.4)$$

where k_B is Boltzmann's constant, T is the temperature and i, j are the indices of Cartesian coordinates. The δ -function of the time difference represents an ideal white noise. Though any actual random force has a finite correlation time corresponding to the collision time of the molecules, the Langevin equation is used to describe the motion of a "macroscopic" particle at a very long time scale, and in this context the δ -correlation and the Langevin equation become good approximation. A generic feature of the Langevin equation is the presence of the damping coefficient λ in the correlation function of $\boldsymbol{\eta}(t)$, which is the random force.

In 1963 Brown first calculated the switching rate of a Stoner-Wohlfarth particle at finite temperature[1]. He assumed that the thermal fluctuation spectrum can be taken as white noise up

to a frequency of $\frac{kT}{h} \sim 10^{13}$ Hz. The intrinsic frequency of a single-domain particle is of the same order of its gyromagnetic resonance frequency, i.e. $H_K \gamma \sim 10^{10}$ Hz. Thus the thermal fluctuation effect can be treated as Brownian motion. For the case of slow switching over a high barrier, Brown used a method introduced by Kramers[3] in 1940.

In this dissertation, we will generalize the Kramers-Brown method to a general (many-macrospin) micromagnetic system, governed by a LLG equation describing computational cells indexed by l :

$$\frac{d\mathbf{M}}{dt} = -\gamma \mathbf{M}^l \times \mathbf{H}^l - \frac{\alpha\gamma}{M_s} \mathbf{M}^l \times (\mathbf{M}^l \times \mathbf{H}^l) + \left(\frac{d\mathbf{M}^l}{dt} \right)_{random}, \quad (3.5)$$

where \mathbf{H}^l is the field felt by the l^{th} cell. To get an equation for the total magnetic moment m_x , we add this up over all cells:

$$\frac{d\mathbf{m}}{dt} = \frac{d(V \sum_l \mathbf{M}^l)}{dt} = -\gamma V \sum_l \mathbf{M}^l \times \mathbf{H}^l - \frac{\alpha\gamma V}{M_s} \sum_l \mathbf{M}^l \times (\mathbf{M}^l \times \mathbf{H}^l) + V \left(\frac{d \sum_l \mathbf{M}^l}{dt} \right)_{random}, \quad (3.6)$$

where V is the volume of one computational cell.

Eq. 3.6 is a complicated expression, but it depends linearly on the field. We will write the total field acting at the l^{th} cell as $\mathbf{H}^l = \mathbf{H} + \mathbf{H}^{l,int}$ where \mathbf{H} without a superscript is the external field, and the internal field includes the exchange field of neighboring cells and the anisotropy field. In principle we should also include the magnetostatic field, but we neglect it here. This is because our work is largely methodological, and we want to use the simplest model that has the correct

qualitative behavior. To make quantitative comparisons to experiment results, we would need to add the magnetostatic field.

Now Eq. 3.6 can be written as

$$\frac{dm_i}{dt} = \left(\frac{dm_i}{dt} \right)_{\text{int}} + \sum_k C_{ik} H_k + \left(\frac{dm_i}{dt} \right)_{\text{random}}, \quad (3.7)$$

where we have (according to LLG equation)

$$C_{ik} = -\gamma V \sum_{l,j} \varepsilon_{ijk} M_j^l - \sum_l \frac{\alpha \gamma V}{M_s} [M_i^l M_k^l - \delta_{ik} (M^l)^2]. \quad (3.8)$$

We are interested in the component m_x ($i = x$), and the external field is in the x direction ($k = x$), so we need

$$C_{xx} = -\frac{\alpha \gamma V}{M_s} \sum_l [(M_x^l)^2 - (M^l)^2] = \frac{\alpha \gamma V}{M_s} \sum_l [(M_y^l)^2 + (M_z^l)^2]. \quad (3.9)$$

In Eq. 3.7, this plays the role of a field mobility, and we will denote it by μ :

$$\mu = C_{xx} = \frac{\alpha \gamma V}{M_s} \sum_l [(M_y^l)^2 + (M_z^l)^2], \quad (3.10)$$

and we can rewrite the x-component of Eq. 3.7 as

$$\frac{dm_x}{dt} = \mu(m_x) [H_x - H_{\text{pin}}(m_x)] + \left(\frac{dm_x}{dt} \right)_{\text{random}}, \quad (3.11)$$

where the pinning field $H_{\text{pin}}(m_x)$ is defined as the ratio of time derivative of m_x caused by internal field to the mobility in the x direction:

$$H_{\text{pin}} \equiv \left(\frac{dm_x}{dt} \right)_{\text{int}} / \mu(m_x). \quad (3.12)$$

We refer to this field as \mathbf{H}_{pin} because an external field with equal magnitude but opposite direction can “pin” the domain motion by creating a local energy minimum, which can be

verified in micromagnetic simulations. Eq. 3.11 is now a one-dimensional Langevin equation for a single variable m_x .

Since the behavior of magnetic moment can be taken as Brownian random walk, and the number of magnetic grains in the ensemble is a constant, the one dimensional Fokker-Planck equation[10] for this Langevin equation takes the form

$$\frac{d\rho(m_x)}{dt} = -\frac{dJ}{dm_x}, \quad (3.13)$$

where ρ is the distribution function of magnetic moment and J is the probability current and can be expressed as

$$J = \rho\mu(H_x - H_{pin}(m_x)) - D\frac{d\rho}{dm_x}. \quad (3.14)$$

The first term on the right hand side comes from the deterministic part of LLG equation and the second term originates from thermal fluctuation, where D is the diffusivity defined

$$\text{as } \langle \Delta m_x^2 \rangle = 2D\Delta t.$$

If $J = 0$, the system is at its equilibrium state, and

$$\rho(m_x) = Ce^{-[E(m_x) - \mu_0 m_x H]/k_B T} \quad (3.15)$$

for some constant C where $E(m_x)$ is the minimum energy with m_x constrained at zero field, and we have added a Zeeman term for nonzero field. Substituting Eq. 3.15 into Eq. 3.14 and noting that this current should be zero in equilibrium gives

$$0 = \left[\mu(m_x)H - \mu(m_x)H_{pin} + \frac{D}{kT} \frac{\partial E}{\partial m_x} - \frac{D}{kT} \mu_0 H \right] Ce^{-[E(m_x) - \mu_0 m_x H]/k_B T}. \quad (3.16)$$

The terms proportional to H give an Einstein relation

$$kT \mu(m_x) = \mu_0 D(m_x). \quad (3.17)$$

and the other terms give a relation between $E(m_x)$ and $H_{pin}(m_x)$:

$$\mu_0 H_{pin} = \frac{\partial E(m_x)}{\partial m_x}. \quad (3.18)$$

At this point, since we are investigating thermal stability problem the external field is set to zero.

To get the rate for a steady state with a small current, Kramers assumed J is a small constant[3].

Then he wrote the distribution function in the form of Eq. 3.15, but with C as a function of m_x .

Following his approach, we substitute Eq. 3.15 into Equation 3.14 and set $H = 0$ (in the case of thermal switching) and get

$$J = -D(m_x) \frac{dC(m_x)}{dm_x} e^{-E(m_x)/k_B T}. \quad (3.19)$$

This gives an equation for $dC(m_x)/dm_x$:

$$\frac{dC(m_x)}{dm_x} = -\frac{J}{D(m_x)} e^{+E(m_x)/k_B T}. \quad (3.20)$$

The boundary condition is that $C(+m_s)=0$ (in the empty energy well) so that

$$C(m_x) = \int_{m_x}^{m_s} \frac{J}{D(m_x)} e^{+E(m_x)/k_B T} dm_x. \quad (3.21)$$

The switching rate is then $r = J /$ (well population):

$$r = \frac{J}{\int_{-m_s}^{+m_s} C(-m_s) e^{-E(m_x)/k_B T} dm_x} = \frac{1}{\int_{-m_s}^{m_s} \frac{1}{D(m_x)} e^{+E(m_x)/k_B T} dm_x \int_{-m_s}^{+m_s} e^{-E(m_x)/k_B T} dm_x}, \quad (3.22)$$

where m_x^b is the x component of magnetic moment at the energy barrier. The integral of well population stops at m_x^b because initially well population is located to the left of the energy barrier. Here we have used the fact that C is nearly constant in the initial well in computing its population because the thermal equilibrium is well established in the initial well. We can substitute Eq. 3.22 into the Arrhenius form $r = r_0 \exp(-E_b/k_B T)$, and have

$$r_0 = \frac{1}{\int_{m_x}^{m_s} \frac{1}{D(m_x)} e^{+[E(m_x)-E_b]/k_B T} dm_x \int_{-m_s}^{+m_s} e^{-E(m_x)/k_B T} dm_x}. \quad (3.23)$$

This rate prefactor or ‘‘attempt frequency’’ r_0 can be calculated numerically once we know $E(m_x)$ and $D(m_x)$, which can be calculated within our micromagnetic calculation along the switching path. However, it is instructive to look at a simple analytic approximation. Note that the first integral in the denominator is dominated by energies near the barrier. Thus we can approximate the energy by a parabola near the barrier: $E(m_x) = E_b (1 - G^2 m_x^2 / m_s^2)$, and integrate to infinity: the integral turns to be

$$\int_{-\infty}^{\infty} \frac{1}{D} e^{-E_b G^2 m_x^2 / m_s^2 k_B T} dm_x = \frac{1}{D} \sqrt{\frac{\pi m_s^2 k_B T}{E_b G^2}} = \frac{m_s}{DG} \sqrt{\frac{\pi k_B T}{E_b}}. \quad (3.24)$$

This approximation gives good agreement with numerical integral as shown later in section i.

Now the rate prefactor changes to

$$r_0 = \frac{GD}{m_s \sqrt{\frac{\pi k_B T}{E_b}} \int_{-m_s}^{m_s} e^{-E(m_x)/k_B T} dm_x}. \quad (3.25)$$

The integral of $\int_{-m_s}^{m_s} e^{-E(m_x)/k_B T} dm_x$ is easy if we approximate $E(m_x)$ by a linear function

$\mu_0 H_{pin}(m_x + m_s)$ so that $E(-m_s) = 0$ and once again take advantage of the exponential function by changing the integration upper limit to infinity:

$$\int_{-m_s}^{m_s} e^{-E(m_x)/k_B T} dm_x \approx \int_0^\infty e^{-\mu_0 H_{pin} m_x / k_B T} dm_x = \frac{k_B T}{\mu_0 H_{pin}}. \quad (3.26)$$

H_{pin} is almost a constant at $m_x \approx -m_s$ due to the fact that uniaxial anisotropy $K(x) \sim x^2$ [30], and in thin domain wall limit the energy is proportional to square root of $K(x)$: $E \propto \sqrt{AK(x)} \propto x \propto m_x$,

so $H_{pin} = \frac{\partial E}{\partial m_x} = const.$, as verified in the modeling[30].

According to the Einstein relation (Eq. 3.17), the diffusivity D is proportional to the field mobility μ (Eq. 3.10), in which the expression being summed over cells involves the hard axis magnetization, which is zero outside the domain wall. Thus the diffusivity can be written as:

$$D = \frac{\alpha \gamma k_B T m_s}{\mu_0} W \text{ with } W = \sum_l V [(M_y^l)^2 + (M_z^l)^2] / M_s^2 V_{tot}, \quad (3.27)$$

in terms of a domain wall volume fraction W , or equivalently the ratio of domain wall width to grain length. Here V_{tot} is the total volume of the magnetic grain, and $m_s = M_s V_{tot}$ is the total magnetic moment. Thus

$$r_0 = \frac{GD}{m_s \sqrt{\frac{\pi k_B T}{E_b} \frac{k_B T}{\mu_0 H_{pin}}}} = \alpha \gamma H_{pin} \sqrt{\frac{E_b}{\pi k_B T}} WG. \quad (3.28)$$

This result differs from Brown's 1963 result [3] only in that there is an extra factor WG (and the coercivity H_K is replaced by H_{pin}). The domain wall width factor (which is the ratio of domain wall width to the entire length of the system) W is 1 in the case of coherent switching and is 0 in the case of infinitely thin domain wall. G^2 is the curvature of energy landscape at maximum, which is also 1 in the case of coherent switching. Thus our result for incoherent switching reduces to Brown's result for coherent switching in the appropriate limit.

d. Methods for calculating Minimum Energy Path

To calculate the switching rate we need to know the most probable path for switching. This is known as “minimum energy path” (MEP) problem in theoretical chemistry and condensed matter physics. In this context people often define a “reaction coordinate” for transitions from one state to another, such as chemical reactions, changes in molecules conformation, or solid state diffusion processes. The maximum energy along the MEP is the saddle point energy, which defines the activation energy barrier. To find reaction paths and saddle points, many different methods have been proposed. One of the schemes, known as the drag method or “sweeping algorithm”, uses some subset of the coordinates in the system as a progress variable, often through linear interpolation between the initial and final configurations. This one degree of freedom is then varied stepwise from the initial to the final condition, and at each step a minimization will be done with respect to the remaining degrees of freedom ($3N - 1$ if there are N atoms in 3-dimensional space, the same with our problem if we have N computational cells)[33, 34]. However, this method does not necessarily lead to minimum energy paths or saddle points. Fig. 11 shows a counterexample. Two black lines represent energy paths found by sweeping algorithm. The red line represents the true minimum energy path. The result shows that sweeping algorithm gives a discontinuous path, which depends on the direction of dragging. Therefore in some cases, due to hysteresis effect, the path found may be discontinuous and depending on the direction of the drag. In some other cases, the saddle point configuration will be missed, due to the effect that atomic coordinates may “slip” near the saddle point region. This phenomenon will be illustrated later in section g of this chapter.

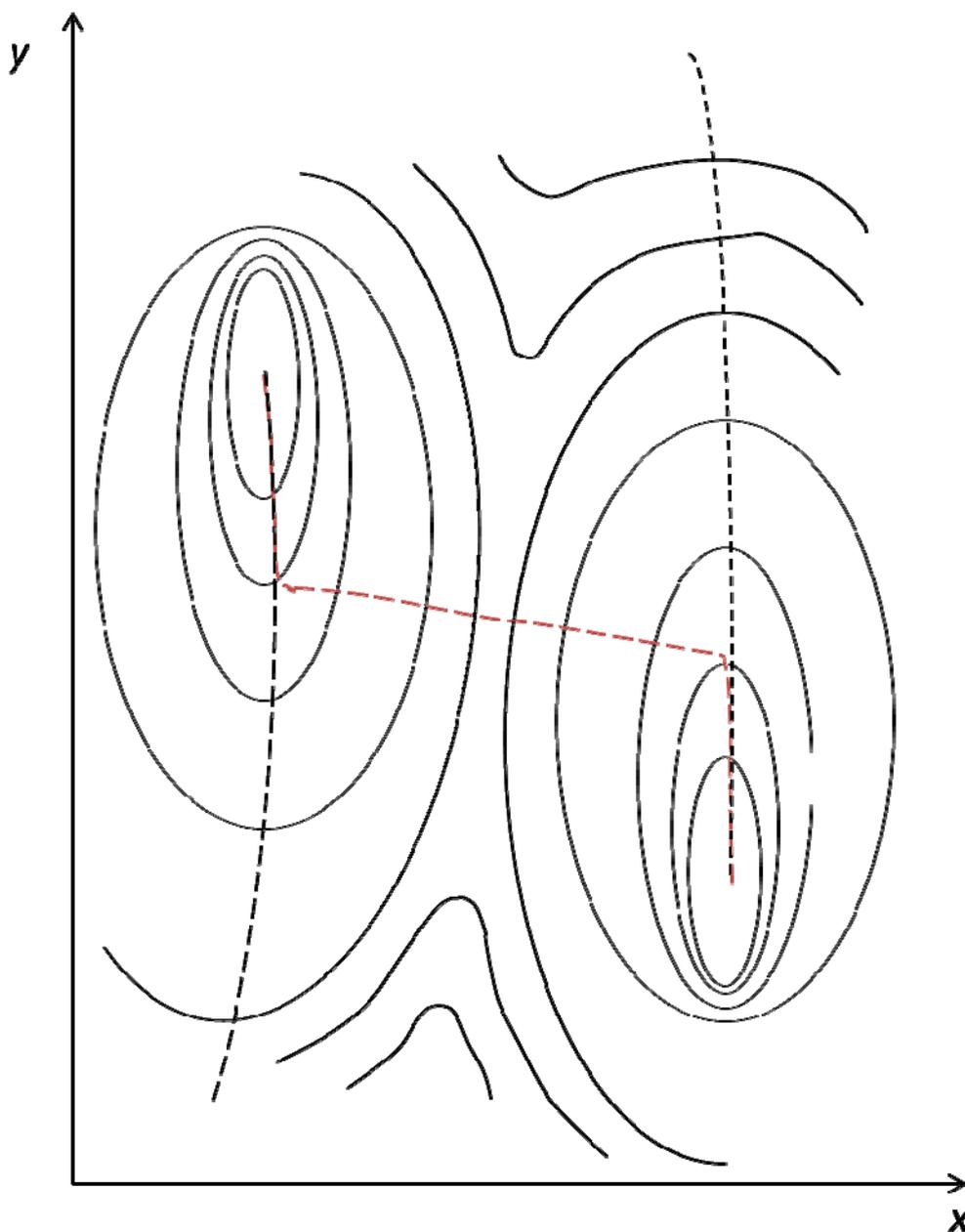


Figure 11 A contour plot of counterexample[35]. This is an energy potential model mimicking a reaction involving atom bond formation.

Another commonly used method is to first do a local harmonic approximation of the energy surface, then find its normal modes and follow each of the modes until a saddle point is found[36]. Evaluation and diagonalization of the second derivative matrix is involved in each

step of this method. So it is only subject to rather small systems and systems where second derivatives are available, and is computationally expensive for a micromagnetic system involving N by N demagnetization tensor matrix if N is large.

Some other methods use a two point boundary condition, i.e. both the initial and final configurations for the reaction are given. Each of these configurations corresponds to a local minimum on the multidimensional potential energy surface. A good example of these methods, the nudged elastic band (NEB) method[35], has been used to solve some micromagnetic problems[37].

The simplest of the methods discussed above is the sweeping algorithm. It was also used in our work as an alternative attempt to find MEP.

e. Convergence of Micromagnetic System with Constraints – Generalized Newton-Raphson Method

As mentioned in the previous section, to calculate the switching rate by Langevin equation, one needs a complete energy minimum path. Since the whole system has $2N$ degrees of freedom (each of the cells has 2 degrees of freedom (θ, ϕ) and M_s is generally assumed to be a constant). If N (the number of computational cells) is too many it would be over-complicated to calculate energy landscape as a function of all degrees of freedom.

Though we will only use the case of one or two such moments, we will establish our notation assuming an arbitrary number of moments, n . Consider a micromagnetic system, whose state is defined by the magnetization configuration $\mathcal{M}(\mathbf{r})$. In a finite-difference calculation, it is

represented by giving the magnetization vectors of many computational cells, which we will collectively refer to below as \mathcal{R} . The field applied to the system is $\mathbf{H}(\mathbf{r})$, which we will express in terms of n normal modes indexed by $i=0,1,\dots,n-1$.

$$\mathbf{H}(\mathbf{r}) = \sum_i h_i \mathbf{F}_i(\mathbf{r}), \quad (3.29)$$

where the mode shape functions \mathbf{F}_i may be sinusoidal (to induce spin waves or stripe domains, for example) or linear in \mathbf{r} (so there is a field gradient). We will refer to the mode amplitudes h_i collectively as an n -component vector \mathbf{h} . We will restrict ourselves to simple field distributions which can be expressed in Eq. 3.29. Each of the modes has a corresponding moment m_i (canonically conjugate to the field h_i , as shown in Eq. 3.31 later)

$$m_i = \int \mathbf{M}(\mathbf{r}) \cdot \mathbf{F}_i(\mathbf{r}) d^3\mathbf{r}, \quad (3.30)$$

which we again refer to collectively as \mathbf{m} . Then the energy of the system is

$$E(\mathbf{M}, \mathbf{h}) = E_0(\mathbf{M}) - \mu_0 \int \mathbf{M}(\mathbf{r}) \cdot \mathbf{H}(\mathbf{r}) d^3\mathbf{r} = E_0(\mathbf{M}) - \mu_0 \sum_i h_i m_i(\mathbf{M}). \quad (3.31)$$

Our objective is to compute an energy landscape $E(\mathbf{m})$ as a function of the moments m_i , i.e., of the n -component vector \mathbf{m} . This is a constrained minimum energy – $E(\mathbf{m})$ is the minimum of $E_0(\mathcal{R}) = E(\mathcal{R}, \mathbf{h} = 0)$ when \mathcal{R} is allowed to vary over all configurations having the desired values of the m_i . Minimizing at $\mathbf{h} = 0$ is the same as minimizing at an arbitrary \mathbf{h} , since we have constrained the m_i so the Zeemann term $-\mu_0 \sum_i h_i m_i$ is just a constant. That is, minimum value of the energy $E(\mathcal{R}, \mathbf{h})$ at constrained m_i is $E(\mathbf{m}) - \mu_0 \sum_i h_i m_i$ and we do not need to calculate the minimum separately from each \mathbf{h} . So the configuration \mathcal{R} minimizing the energy for fixed \mathbf{m} is independent of \mathbf{h} .

The field \mathbf{h} plays the role of a Lagrange multiplier here. As in the Lagrange multiplier method, if we now minimize the energy without any constraint (at nonzero \mathbf{h}) we can adjust \mathbf{h} so that the moments \mathbf{m} have any desired value at this minimum. Because this is a minimum over all possible configurations \mathcal{X} , it must also be a minimum when \mathcal{X} is constrained, so we have found our desired constrained minimum. The unconstrained minimization is done with the LLG time evolution algorithm with large damping, which is well known to be a very efficient way to find energy minima. If we remove the field \mathbf{h} , the energy is no longer a minimum with respect to changes in the moment \mathbf{m} ; rather, its derivative is nonzero and

$$\mu_0 h_i = -\frac{\partial E(\mathbf{m})}{\partial m_i}. \quad (3.32)$$

Thus we need an algorithm that evolves the system according to LLG, while at the same time adjusting \mathbf{h} so that \mathbf{m} approaches its desired value.

Suppose we are given a desired set of moments \mathbf{m}^{want} , and we want to find a configuration \mathcal{X} that has these moments and the minimum energy $E(\mathbf{m}^{want})$ consistent with these moments. The LLG equation will take care of relaxing the other degrees of freedom – the configuration

$$\mathcal{X}(t + \Delta t) = \mathcal{X}(t) + \frac{d\mathcal{X}}{dt}(\mathbf{M}, \mathbf{h})\Delta t$$

will be closer to the minimum than is $\mathcal{X}(t)$. However, we

must keep track of the changes in moments \mathbf{m} and move them towards \mathbf{m}^{want} . We do this by changing the field amplitudes h_i . To determine how to change them, we use an analog of the classical Newton-Raphson method. It is important to note that this is different from applying the standard Newton-Raphson method for solving a 2N dimensional equation. We may be dealing with values of N for which that would be impossible (and it is cumbersome even for fairly small

N). We are using the Newton-Raphson method only for n ($= 1$ or 2) of the variables – the rest of them will relax according to the (much more efficient) LLG equation. We start with a particular configuration \mathcal{K}_0 , with moments \mathbf{m}_0 , and a particular set of values for the field amplitudes \mathbf{h}_0 . We calculate the expected change in moments during Δt , $\mathbf{m}(\mathcal{K}_0, \mathbf{h}_0)$, and how this change is altered by an additional field \mathbf{h} – that is, the matrix of derivatives

$$\mathbf{J} = \frac{\partial(\Delta\mathbf{m})}{\partial\mathbf{h}} \quad \text{or} \quad J_{ij} = \frac{\partial(\Delta m_i)}{\partial h_j}$$

so that

$$\Delta m(\mathcal{K}_0, \mathbf{h}_0 + \delta\mathbf{h}) = \Delta m(\mathcal{K}_0, \mathbf{h}_0) + \mathbf{J} \delta\mathbf{h} . \quad (3.33)$$

We then choose the additional field \mathbf{h} so that the time evolution gives the desired moments \mathbf{m}^{want} , as illustrated schematically in Fig. 12

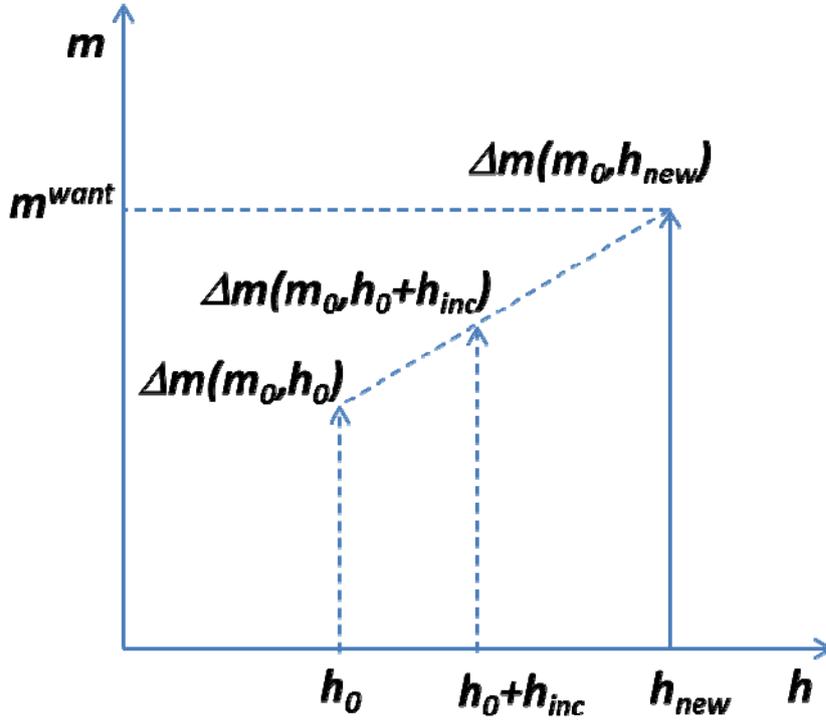


Figure 12 Illustration of the Newton-Raphson Method. The time derivative of magnet moment Δm at applied field h_0 and h_0+h_{inc} can be calculated to find the slope of Δm versus applied field. Then accordingly we can find the right field to apply to reach desired magnetic moment m^{want} .

$$m^{want} - m_0 = \Delta m(\mathcal{K}_0, h_0) + J \delta h, \quad (3.34)$$

which requires

$$\delta h = J^{-1}(m^{want} - m_0 - \Delta m(\mathcal{K}_0, h_0)). \quad (3.35)$$

If we start out far from m^{want} , it may be over-optimistic to expect to get to m^{want} in one time interval Δt . This will be reflected in the fact that the field h may be very large. If so, we scale it down to some maximum absolute value h_{max} , and it may take several time intervals to get near to m^{want} .

When we apply this algorithm, the convergence takes place in two stages. At first the moments are far from the desired ones m^{want} , and the requested fields exceed the maximum, so we evolve with the maximum allowed field until the moments are close. Then the Newton-Raphson algorithm takes over, and we converge to the correct moments in a few time steps. It may take a while longer for the rest of the configuration to relax (for short-wavelength spin waves to die out, for example) – this convergence is similar to that in conventional applications of the LLG equation to find equilibrium configurations.

f. One dimensional application of Newton-Raphson Algorithm: The 1D Energy Landscape

To model anisotropy graded media and investigate the switching rate problem, we construct a micromagnetic system with 20 computational cells whose magnetizations are initially aligned in the x direction.

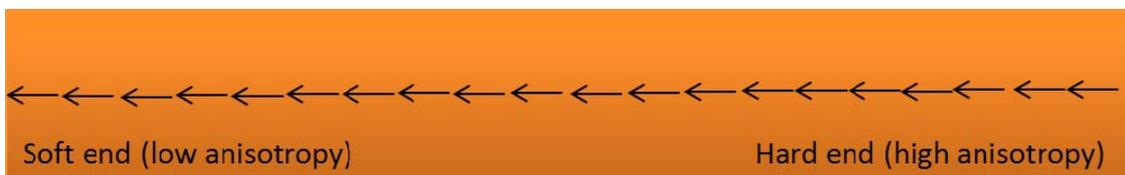


Figure 13 Anisotropy graded media. Each vector corresponds to the magnetization of a computational cell. The initial state is taken as negative saturation, i.e. $m_x \sim -m_s$. The left end is the soft end, i.e. $H_K = 0$, and the right end is the hard end, with $H_K = 9550$ kA/m.

The anisotropy is chosen as a quadratic function of x coordinate to maximize the figure of merit as mentioned in section b of this chapter.

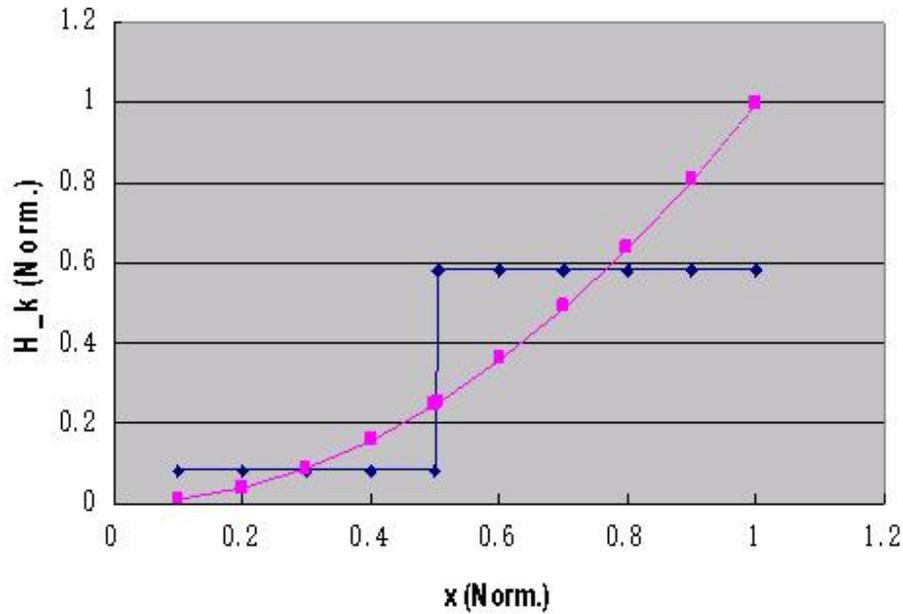


Figure 14 Anisotropy as a function of x . The blue curve represents the anisotropy of exchange coupled media as a function of x coordinate, and the pink curve represents that of the anisotropy graded media. The anisotropy of exchange coupled media is chosen in such a way that the average of each layer is the same as that of anisotropy graded media.

The purple line represents the anisotropy as a function of x coordinate in anisotropy graded media, while the blue line represents the anisotropy of exchange coupled media for comparison purpose. In our settings the following values are chosen: $M_s = 1520$ kA/m (L1₀ FePt), exchange constant[15] is estimated as $A = 1 \times 10^{-11}$ J/m, $H_{K_max} = 9550$ kA/m. The size of system is set to be $10\text{nm} \times 5\text{nm} \times 5\text{nm}$, and the smallest computational cell is $0.5\text{nm} \times 5\text{nm} \times 5\text{nm}$, and the system is divided into 20 cells.

In the case of switching, suppose the initial magnetization of the system points to the left (this is a symmetric problem with respect to the x coordinate, so it doesn't matter if the magnetization actually points to the right). The domain wall first appears in the soft end of the media (on the left of Fig. 13) then moves to the hard end (on the right) and finally disappears, and the switching is completed.

An obvious magnetic moment to constrain is the x component of total magnetization of the system (denoted by m_x from now on), because $m_x = -m_s$ and $m_x = m_s$ are the two energy minima of the system. So the basic scheme is, we start from a domain nucleation from the left (where m_x is close to $-m_s$), and constrain the system with respect to m_x until the energy minimum is reached. Then we slowly increase m_x and make sure the system converge to the energy minimum after every time m_x is incremented.

The constraint of m_x can be easily realized by applying an external field in the x direction. According to the LLG equation, Eq. 3.11 and Eq. 3.12, we can consider C_{xx} as “mobility” of the system in the x direction if a horizontal external field is applied.

Suppose the initial condition of the system has magnetization in the x direction, whose value is m_{x0} . We would like to constrain it to m_{x_seek} . According to Eq. 3.11, we need to apply a field in the x direction:

$$H_x = \frac{m_{x_seek} - m_{x_0}}{\frac{\alpha\gamma V}{M_s} \sum_l [(M_y^l)^2 + (M_z^l)^2] \Delta t}. \quad (3.36)$$

One could argue that the mobility C_{xx} is not a constant over the increment. However, in realizing the algorithm, the increment of m_x is can be set small and the change in C_{xx} can be neglected.

On the other hand, we have assumed that we can get to m_{x_seek} in a single time step. If we are starting from an m_x too far away, Eq. 3.36 will give a huge H_x , which will cause instability in the discretized LLG equation. More specifically, the angle of rotation of a magnetization is roughly $\gamma H_x \Delta t$; to keep this angle below some θ_{max} , we limit H_x to $H_{max} \sim \theta_{max} / \gamma \Delta t$, where we have found

$\theta_{max} = 0.1$ is small to avoid artificial instability in the numerical simulation. The effect of this algorithm is to impose a very large field that changes m_x rapidly until we get near m_{x_seek} , and then switch to the discussed algorithm.

An alternative way to constraint the system to a given m_{x_seek} is the Newton-Raphson method discussed in previous section. First we apply a field in the x direction H_{x0} , and get the change in x component of magnetization Δm_{x0} . Second we restore the system to its starting point and apply another field $H_{x0} + H_{inc}$, record the change in m_x as Δm_{x1} . Finally apply a field using the formula below:

$$H_{new} = \frac{m_{x_seek} - m_{x0}}{\Delta m_{x1} - \Delta m_{x0}} H_{inc} + H_{x0}. \quad (3.37)$$

Once we have the system adjusted to the correct m_x , the remaining work is to make it relax to the energy minimum. As mentioned in the previous sections, this can be done by constraining m_x while continuing the LLG iteration of the system.

Instead of carrying out the algorithm described above to a series of m_x sequentially, another implementation of this algorithm is doing a slow “sweep” over m_x . “Slow” means m_x is incremented at constant speed with respect to time, but the system will have sufficient time to relax to its energy minimum. This is an approximation of the energy minimum, because the system has to relax while m_x is still slowly increasing. However, as shown later, for most part of the minimum energy path, this is a good approximation.

The sweeping algorithm is also related to the Newton-Raphson algorithm. As discussed in section c of this chapter, LLG equation is a linear equation with respect to the applied field H at zero temperature. So Newton-Raphson method is well applicable here and proved to be working. Suppose at zero applied field, the change in m_x over one time step is

$$\Delta m_x = m_x(0, t + \Delta t) - m_x(0, t), \quad (3.38)$$

where the first parameter in the parentheses is applied field and the second one is time. Then starting from the same initial condition, we apply a trial field H_{trial} , and let the change in m_x in one time step be

$$\Delta m_x' = m_x(H_{trial}, t + dt) - m_x(H_{trial}, t). \quad (3.39)$$

Then as shown in Fig.15, the field used to have $\Delta m_{x_seek} = \Delta m_x$ is

$$H_x = \frac{\Delta m_{x_seek} - \Delta m_x}{\Delta m_x' - \Delta m_x} H_{trial}. \quad (3.40)$$

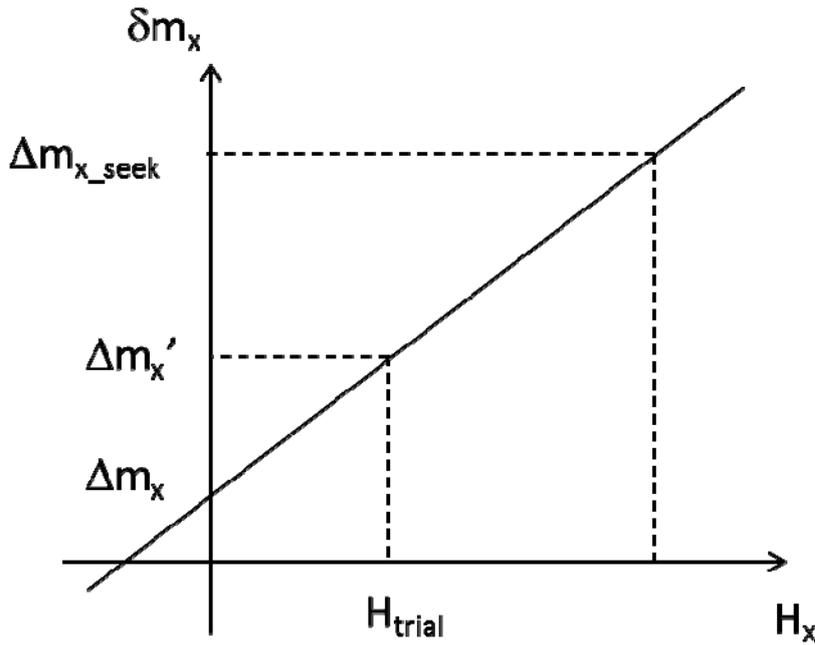


Figure 15 Linear extrapolation of m_{x_seek} . It turns out to work efficiently, as the time derivative of m_x reaches the desired value in a few time steps.

The minimum energy path generated by this algorithm is shown in Fig. 16. As stated previously, first we start from a saturated magnetization pointing to the left (marked as position 1 in the figure), and sweep m_x incrementally to the right. At position 2 we get energy maximum, and right after that, the system loses its stability i.e. cannot maintain its original path for sweeping algorithm and quickly “slips” into another minimum energy path, and finally reaches its saturated magnetization on the right, labeled as position 3. The other energy minimum path is also drawn here. Its similarity to the first path is caused by the left-right symmetry, i.e.

$$E(m_x) = E(-m_x). \quad (3.41)$$

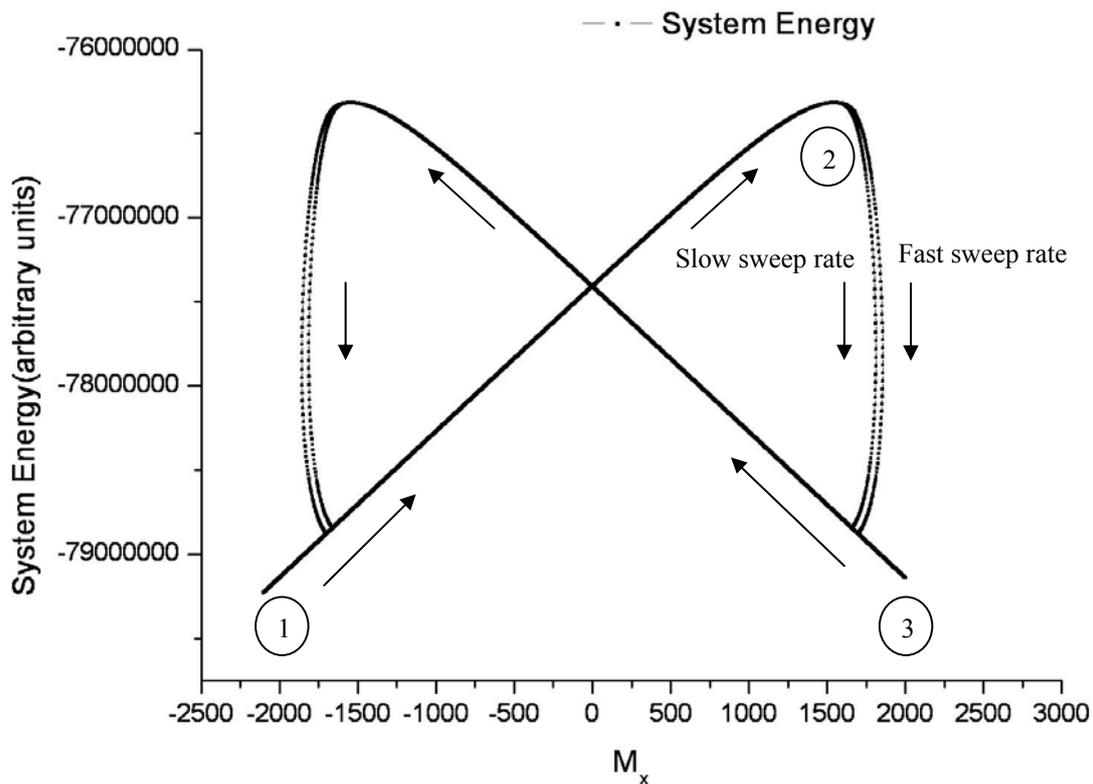


Figure 16 Energy landscape generated by sweeping method. Dashed curve represents fast switching of the system from one branch to the other, and there is no way to do m_x constraint during this fast switching, unlike the situation in the case of solid lines.

The two dotted paths on the downward part of Fig.17 represent different sweep rates – it can be seen that the ascending part of the curve is independent of sweep rate and represents a true energy minimum. It is apparent that something happens on the downward part and the sweep algorithm is unable to maintain the energy minimum. We will see later in section g that this can be understood in terms of a two-dimensional energy landscape – close to the saddle point, there simply is no minimum in the energy at constant m_x .

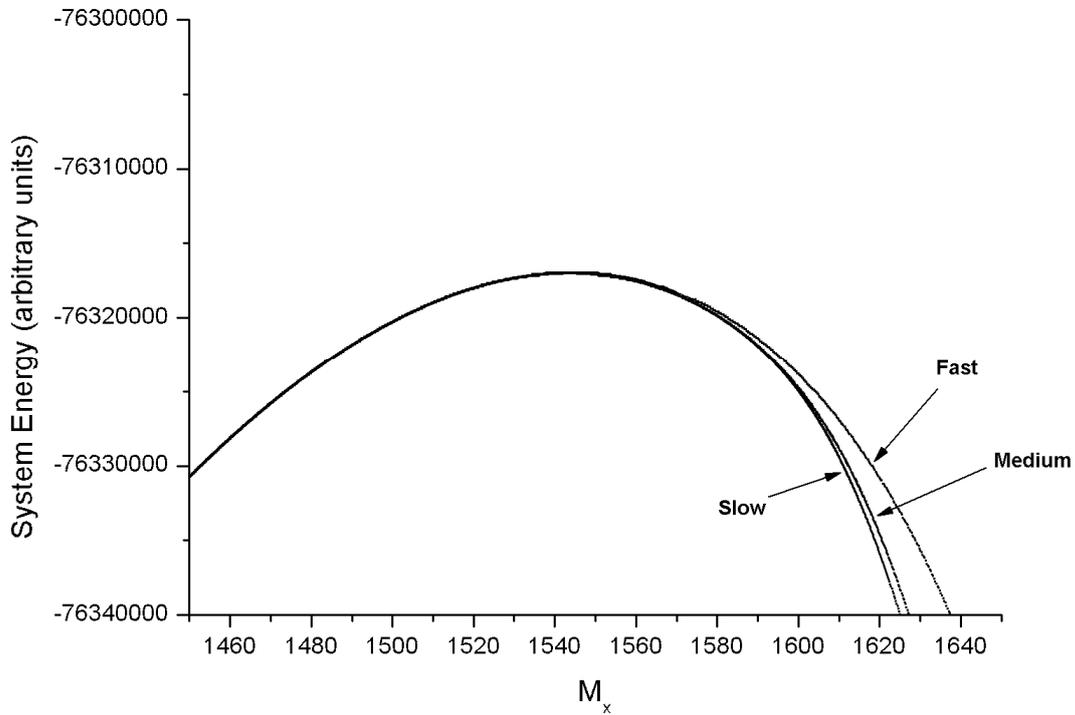


Figure 17 Energy paths at different sweeping speed before and after energy maximum (enlargement of part of Fig. 16).

g. Two-dimensional Energy Landscape – a Semi-analytical Way to Describe Energy Barrier in Switching

After observing the behavior of the system near the saddle point we noticed that the width of domain significantly increases when switching. So we conclude that we need an extra variable describing the domain wall width.

To formulate the width of the domain wall we chose the following expression:

$$Q_{xx} = \int_{-L/2}^{L/2} M_x x dx dy dz , \quad (3.42)$$

where L is the length of the whole grain (the soft end has $x = -L/2$ and the hard end has $x = L/2$), or the thickness of anisotropy graded media. Here Q_{xx} is the quadrupole moment of the system magnetization, which is analogous to quadrupole moment tensor Q_{ij} in electrostatics[38]. We take the origin to be at the center of a grain to make $Q_{xx} = 0$ for homogeneous magnetization distribution. Imagine that we have a fixed m_x , and the domain wall is infinitely wide, i.e. the magnetization at each cell points to the same direction, as shown in Fig. 18. Then do the integration as Eq. 3.42, we have $Q_{xx} = 0$.

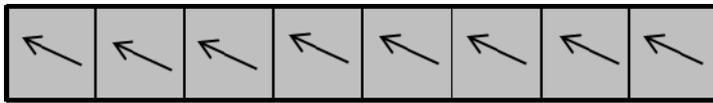


Figure 18 A demonstration of $Q_{xx} = 0$.

Then imagine a system with the same geometry but the domain wall is infinitely thin, which means the left half and the right half of the system is mirrored by the central line. Then the system looks like Fig. 19 and Q_{xx} is the maximum value for given m_x .

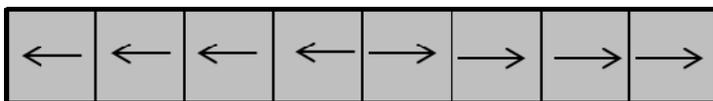


Figure 19 A demonstration of $Q_{xx} = \text{maximum}$.

One could see that the system configuration is not unique for given (m_x, Q_{xx}) : a system with alternating up-and-down magnetization will also have $m_x = 0$, $Q_{xx} = 0$. The ambiguity can be avoided if we constrain the problem to only one domain wall case, as in the simulation we start from a domain wall nucleating from the soft end and do not allow multiple domain walls.

We can work out a formula for the maximum Q_{xx} as a function of m_x . The result is a function like the shape of an American football bounded by two parabolas.

$$\begin{aligned}
|Q_{xx}|_{\max} &= \int_{-L/2}^{L_{dw}} -M_s \cdot \text{area} \cdot x dx + \int_{L_{dw}}^{L/2} M_s \cdot \text{area} \cdot x dx \\
&= -\frac{M_s \cdot \text{area} \cdot x^2}{2} \Big|_{-L/2}^{L_{dw}} + \frac{M_s \cdot \text{area} \cdot x^2}{2} \Big|_{L_{dw}}^{L/2} \\
&= M_s \cdot \text{area} \left(\frac{L^2}{4} - L_{dw}^2 \right) \\
&= \frac{M_s \cdot \text{area}}{4} \left(L^2 - \frac{m_x^2}{M_s^2 \cdot \text{area}^2} \right)
\end{aligned} \tag{3.43}$$

where L_{dw} is the x coordinate of the domain wall, area is the cross section area of the system, and L is the length of the whole system (in our setting, $\text{area} = 5\text{nm} \times 5\text{nm} = 25\text{nm}^2$, and $L = 10\text{nm}$).

The derivation above used the fact that $m_x = 2L_{dw} \cdot M_s \cdot \text{area}$.

Now we have two free variables to describe the behavior of the system and thus its energy landscape $E(m_x, Q_{xx})$, which is a constrained minimum energy – the minimum as we vary all magnetizations keeping m_x and Q_{xx} fixed. The remaining question is how to converge to both given m_x and Q_{xx} .

The basic approach is still the Newton-Raphson method, described above for a general vector shape function $\mathbf{F}_i(r)$, using the specific choices $\mathbf{F}_1(r) = (1, 0, 0)$ and $\mathbf{F}_2(r) = (x, 0, 0)$, as $\mathbf{F}_1(r)$ represents the mode for m_x and $\mathbf{F}_2(r)$ for Q_{xx} . Note that $Q_{xx} = \int M_x dx$ and to change it, we need to apply a field proportional to the shape function $\mathbf{F}_2(r)$, which in this case is a field gradient

$$H_x(x) = H_{x0} + x H_{xx}, \tag{3.44}$$

where H_{x0} is a constant and H_{xx} is the gradient.

We can represent the dependence of $d\mathbf{M}/dt$ on \mathbf{H} by

$$\begin{pmatrix} \frac{dm_x}{dt} \\ \frac{dQ_{xx}}{dt} \end{pmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \begin{pmatrix} H_x \\ H_{xx} \end{pmatrix}. \quad (3.45)$$

Below is an example of how to get J_{11} and J_{21} ; as mentioned before, the algorithm is still based on Newton-Raphson method.

First we apply a trial field H_{x_trial1} , record the response of the system $(\Delta m_{x1}, \Delta Q_{xx1})$ and then from the original state, apply another trial field H_{x_trial2} , and record the response again $(\Delta m_{x2}, \Delta Q_{xx2})$.

The changes in response Δm allow us to calculate the components of J .

Then by applying Newton-Raphson formulas we get:

$$J_{11} = \frac{\delta\left(\frac{dm_x}{dt}\right)}{\delta H_x} = \frac{\Delta m_{x2} - \Delta m_{x1}}{H_{x_trial2} - H_{x_trial1}},$$

$$J_{21} = \frac{\delta\left(\frac{dQ_{xx}}{dt}\right)}{\delta H_1} = \frac{\Delta Q_{xx2} - \Delta Q_{xx1}}{H_{x_trial2} - H_{x_trial1}}. \quad (3.46)$$

We need to take into account the cross effect here i.e. change in m_x caused by H_{xx} and change in Q_{xx} caused by H_x .

Now we have an initial point (m_{x0}, Q_{xx0}) and would like to converge to energy minimum at given $(m_{x_seek}, Q_{xx_seek})$. According to Eq. 3.45, this can be done with the J matrix known:

$$\begin{pmatrix} H_x \\ H_{xx} \end{pmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}^{-1} \begin{pmatrix} \frac{m_{1_seek} - m_{1_0}}{dt} \\ \frac{m_{2_seek} - m_{2_0}}{dt} \end{pmatrix}. \quad (3.47)$$

That involves calculation of matrix inverse:

$$\begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}^{-1} = \frac{1}{J_{11}J_{22} - J_{12}J_{21}} \begin{bmatrix} J_{22} & -J_{12} \\ -J_{21} & J_{11} \end{bmatrix}. \quad (3.48)$$

which is simple to implement in the code.

Now we can easily get the energy minimum at any given (m_x, Q_{xx}) by using LLG equation to relax over other degrees of freedom. So we made a grid of (m_x, Q_{xx}) and calculated the corresponding energy. In this way we did a contour plot of energy as a function of (m_x, Q_{xx}) and mark the minimum energy path obtained by sweeping method on the plot, as shown in Fig. 20. In the figure we can see that at first magnetization is tilted at the left while domain wall starts to nucleate. Then the domain wall slowly moves to the right. At the saddle point, the domain wall is close to the right end of the system. After the trajectory passes through the saddle point the domain wall vanishes from the right end. That is the whole scheme of domain wall switching for anisotropy graded media.

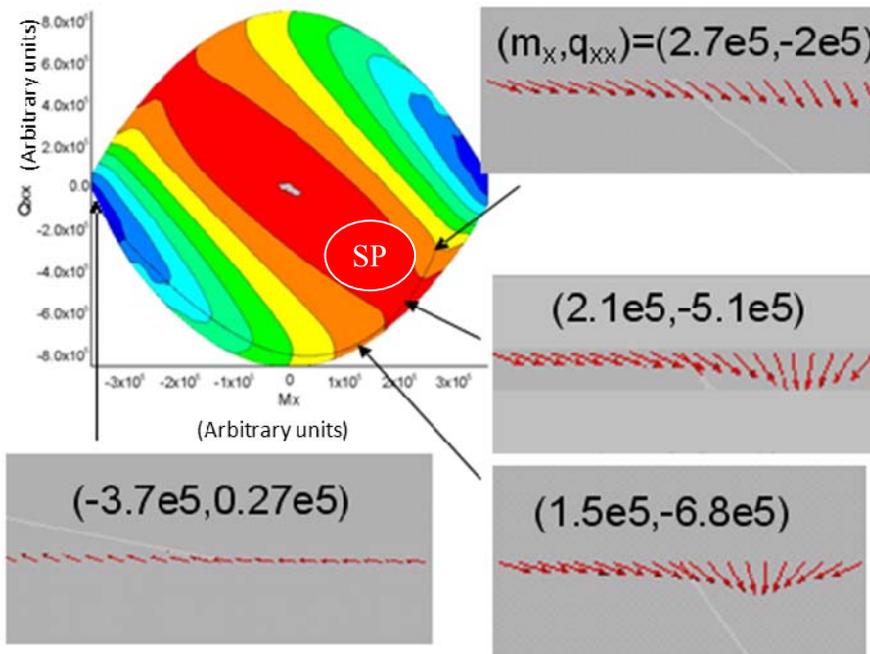


Figure 20 System magnetizations at different stages of switching. Saddle point (SP) is marked on the contour. Blue regions represent low energy and red represents high energy. The switching trajectory is also shown here. One can see the domain wall starts to nucleate from negative saturation, and moves to the right (hard end) and eventually disappears there.

It is helpful to study the energy path generated by the sweeping algorithm (recall that this generates a minimum energy at constrained m_x) in the context of 2D energy landscape. Starting from the left saturation, this method works well for most part of the sweeping but loses its way after passes the saddle point. This can be understood by drawing energy as a function of Q_{xx} at fixed m_x , as shown in Fig. 21.

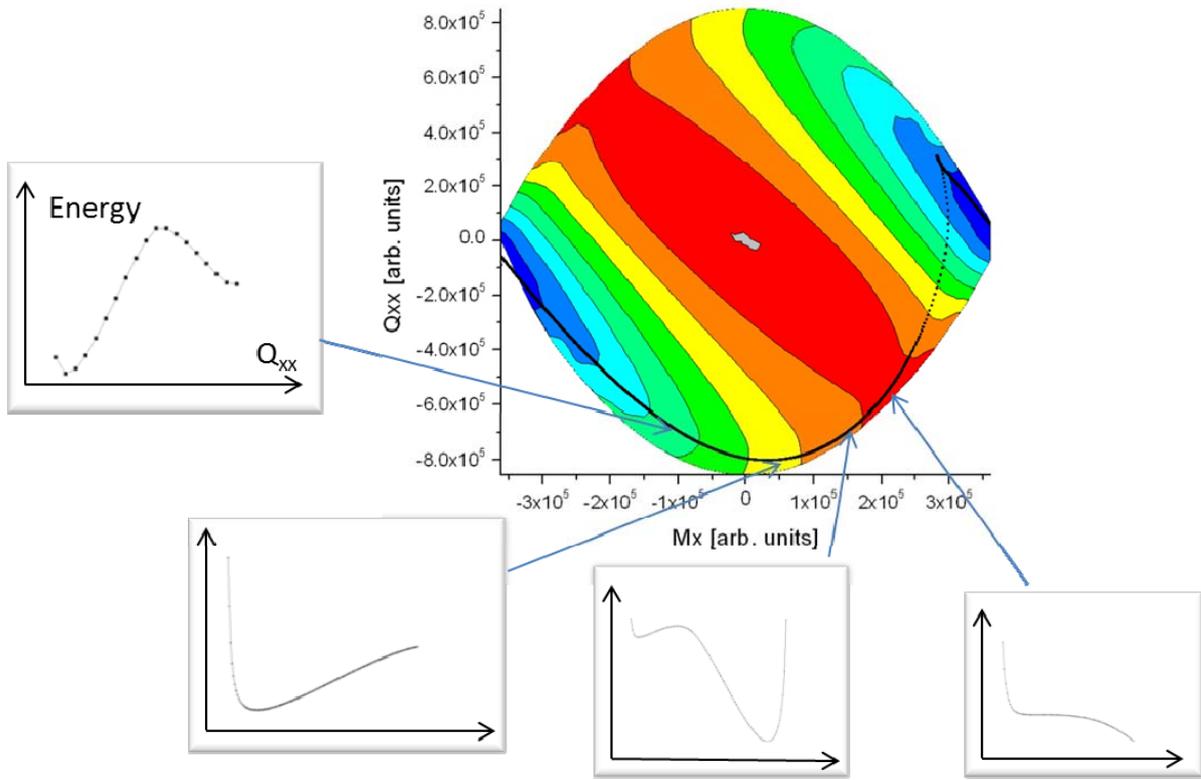


Figure 21 System magnetizations at different stages of switching. The y axes of insets are energy and the x axes are Q_{xx} , as shown in the top left inset.

From the energy functions we can investigate the reason why the sweeping algorithm doesn't work for anisotropy graded media. It is clear that before coming to the saddle point (in the leftmost insets in Fig. 21) the function has a local minimum, describing a domain wall state. However, when the system reaches the saddle point, this local minimum vanishes. In one dimensional sweeping algorithm, the system will "slip" to the other local energy minimum, i.e. another domain nucleates from the right end. This is the reason for failure of sweeping algorithm, which is the lack of free variables to fully control the domain wall.

Now that we have the full grid of the 2D energy landscape, it is easy to find the energy minimum path according to the landscape. The most intuitive way is to do a saddle point vicinity fit. We assume the energy potential surface takes quadratic form

$$E(m_x, Q_{xx}) = z_0 + am_x + bQ_{xx} + cm_x^2 + dQ_{xx}^2 + fm_xQ_{xx}. \quad (3.49)$$

By fitting vicinity of saddle point we get

z_0 [10^{-27} J]	a [10^{-3} J/A/m ³]	b [10^6 J/A/m ³]	c [10^{21} J/A ² /m ⁴]	d [10^{39} J/A ² m ⁶]	f [10^{30} J/A ² m ⁵]
3.36721e8	-1523.02604	941.07915	-9.74702e-4	1.53291e-4	-0.00376

Table 1. Parameters of 2D energy landscape fitting.

At the saddle point we have

$$\begin{cases} \frac{\partial E(m_x, Q_{xx})}{\partial M_x} = 0 \\ \frac{\partial E(m_x, Q_{xx})}{\partial Q_{xx}} = 0 \end{cases} \text{ or } \begin{cases} a + 2cm_x + fQ_{xx} = 0 \\ b + 2dQ_{xx} + fm_x = 0 \end{cases} \quad (3.50)$$

Solving the equation above gives saddle point: $(m_x, Q_{xx}) = (2.08e5, -5.13e5)$. The units for m_x are kA/m, and for Q_{xx} are kA.

h. Barbell Algorithm – a 3N Dimensional Algorithm to Find the Saddle Point in Switching
It is worth noting that saddle point calculated previously is based on the approximation of energy surface grids. The switching rate given by Langevin equation is very sensitive at the vicinity of saddle point because it integrates over exponential function of energy. Therefore, while 2D energy landscape can give us a good understanding of the switching process, we still need a systematic way of finding the saddle point at high precision.

As mentioned previously, different methods have been applied to find the minimum energy path, including the nudged elastic band method which has been successfully applied in micromagnetic system[37]. However, these “image chain” methods require large number of images for high accuracy calculation and are computationally expensive. Here we propose a new method to find the saddle point, which we call the “barbell” algorithm.

The basic idea of barbell algorithm is as shown in Fig. 22. Our objective is to find the saddle point, and also the local minimum-energy path through the saddle point. We can specify these two things by giving two starting points (actually magnetizations in $3N$ dimensional space) \mathbf{M}_1 and \mathbf{M}_2 , whose midpoint is the saddle point, and such that the difference vector $\mathbf{r} = \mathbf{M}_2 - \mathbf{M}_1$ gives the minimum-energy path. Suppose we start from some estimates of \mathbf{M}_1 and \mathbf{M}_2 , as shown in Fig. 22, then the projection of $\frac{d\mathbf{M}_1}{dt}$ and $\frac{d\mathbf{M}_2}{dt}$ on \mathbf{r} at zero applied field must be of opposite direction.

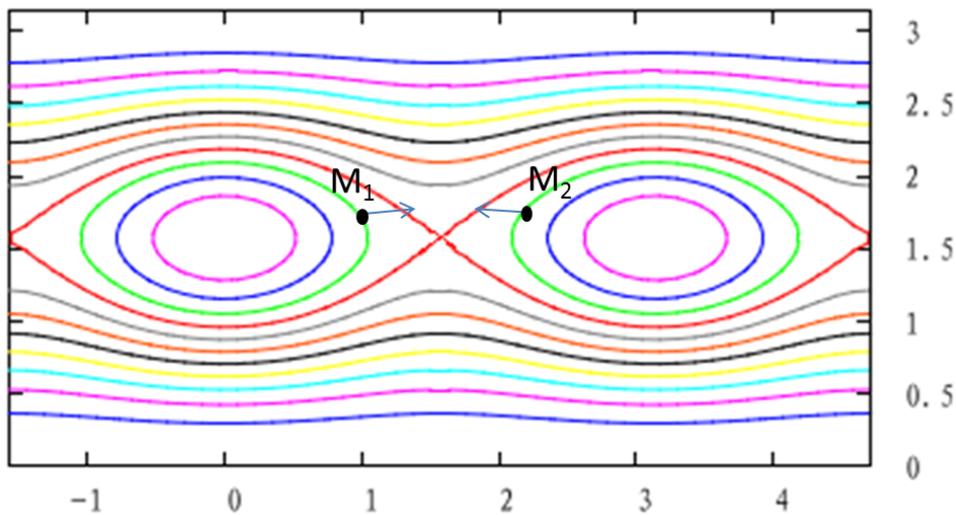


Figure 22 Illustration of barbell algorithm (black dots are M_1 and M_2 blue arrows are dM/dt there).

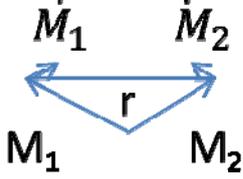


Figure 23 Illustration of M_1 , M_2 and r .

The idea of the barbell algorithm is that the energy rises as we move up or down near the saddle point in Fig. 22, so that the LLG algorithm (which always moves us toward lower energy) will take care of the vertical motion. We need another mechanism to adjust the horizontal position, i.e., the position along the switching coordinate, since LLG will take us outward from the saddle point. The horizontal position is the component along the vector \mathbf{r} . We also want M_1 and M_2 to be separated by a given distance, denoted by ϵ . If we start with estimates for M_1 and M_2 , we want to find better estimates $M_1 + \delta M_1$, and $M_2 + \delta M_2$, satisfying

$$(M_1 + \delta M_1 - M_2 - \delta M_2) \cdot (M_1 + \delta M_1 - M_2 - \delta M_2) = \epsilon^2,$$

and the condition that $dM/dt = 0$ at the saddle point requires

$$\dot{M}(M_1 + \delta M_1) \cdot \mathbf{r} + \dot{M}(M_2 + \delta M_2) \cdot \mathbf{r} = 0.$$

We do this by assuming that $(dM/dt) \cdot \mathbf{r}$ depends linearly on $M \cdot \mathbf{r}$ (a heuristic assumption, justified by the fact that the resulting algorithm indeed converges to the saddle point), so that the slope

$$\frac{\dot{M}(M_2) \cdot \mathbf{r} - \dot{M}(M_1) \cdot \mathbf{r}}{M_2 \cdot \mathbf{r} - M_1 \cdot \mathbf{r}} = \frac{\dot{M}(M_1 + \delta M_1) \cdot \mathbf{r} - \dot{M}(M_1) \cdot \mathbf{r}}{\delta M_1 \cdot \mathbf{r}} = \frac{\dot{M}(M_2 + \delta M_2) \cdot \mathbf{r} - \dot{M}(M_2) \cdot \mathbf{r}}{\delta M_2 \cdot \mathbf{r}}. \quad (3.51)$$

Solving the equations gives

$$\begin{aligned}\delta\mathbf{M}_1 \cdot \mathbf{r} &= \frac{1}{2}(-\varepsilon|\mathbf{r}| - r^2 \frac{2\dot{\mathbf{M}}_1 \cdot \mathbf{r}}{\dot{\mathbf{M}}_2 \cdot \mathbf{r} - \dot{\mathbf{M}}_1 \cdot \mathbf{r}}), \\ \delta\mathbf{M}_2 \cdot \mathbf{r} &= \frac{1}{2}(\varepsilon|\mathbf{r}| - r^2 \frac{2\dot{\mathbf{M}}_2 \cdot \mathbf{r}}{\dot{\mathbf{M}}_2 \cdot \mathbf{r} - \dot{\mathbf{M}}_1 \cdot \mathbf{r}}).\end{aligned}\quad (3.52)$$

which can be achieved by applying a field along the direction of \mathbf{r} . Consider the component of the LLG equation in the direction of \mathbf{r} and using Eq. 3.9, we can get the mobility

$$J = -\sum_i \frac{\alpha\gamma}{M_s} [(r_i M_i)^2 - r_i^2 M_i^2], \quad (3.53)$$

where i is the index of computational cells.

Since we have the mobility of the system, to achieve the change in \mathbf{M}_1 and \mathbf{M}_2 mentioned previously, we can apply a field of which the magnitude is

$$h = \frac{1}{J\Delta t} (\mathbf{r} \cdot \delta\mathbf{M} - \mathbf{r} \cdot \dot{\mathbf{M}}\Delta t), \quad (3.54)$$

on each computational cell of the system.

We implemented the barbell algorithm in the micromagnetic code we have been using. The result proved that this algorithm finds the saddle point in a few thousand iterations, which takes less than 10 CPU minutes to complete (a 2.8 GHz Intel x86 processor), as shown in Fig. 24.

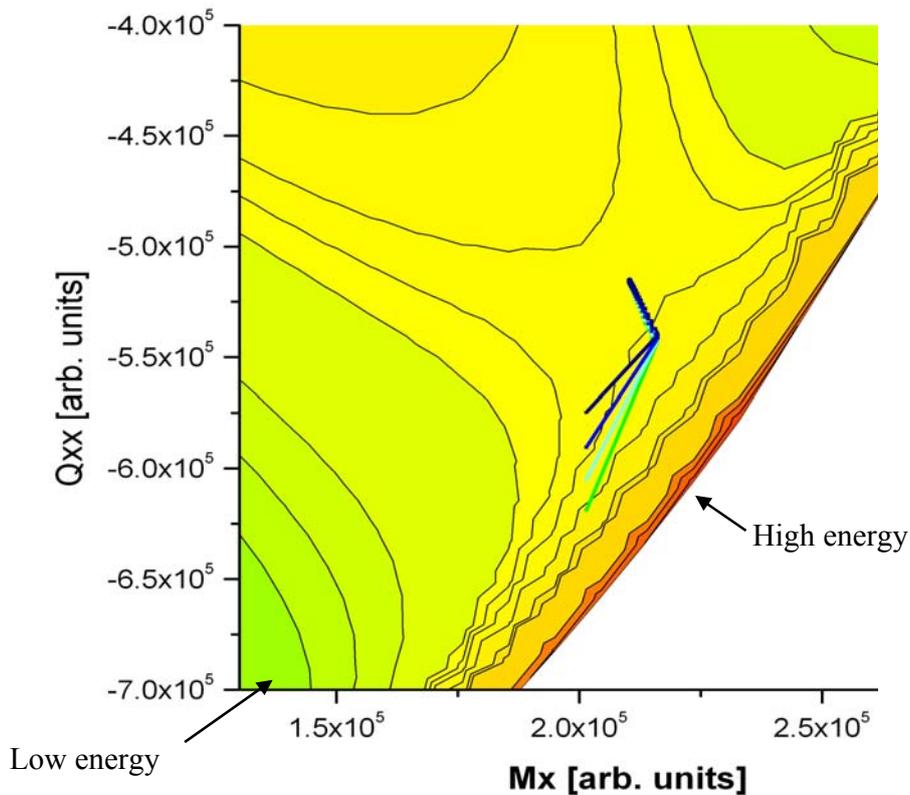


Figure 24 “Barbell algorithm” to find the energy saddle point. Convergence trajectory in different colors represent different initial $(M_1 + M_2)/2$. Energy of the total grain diverges to infinity at the boundary (or maximum Q_{xx} at fixed M_x), so this contour plot based on finite grid gives jaggedness there.

In the figure shown above each line gives the iteration of $(M_1 + M_2)/2$ till convergence. As shown in the figure, the first step is large. The rest of convergence is mostly contributed by Landau-Lifshitz damping with a small artificial correction field given by barbell algorithm. The final convergence point is very close to the saddle point achieved by 2D energy landscape fitting, and provides evidence of its correctness.

While applying this algorithm we must take caution because it is sensitive to initial conditions.

The reason can be understood in the figure shown below.

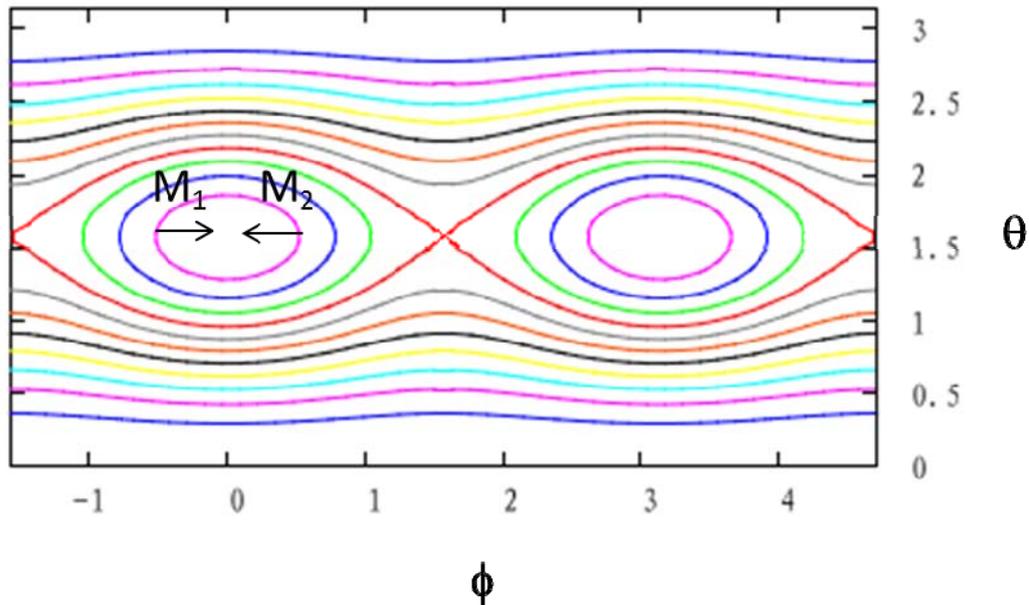


Figure 25 Initial condition for which barbell algorithm converges to energy minimum. This is an energy contour diagram for a thin film magnetic element with in plane anisotropy.

If we start from \mathbf{M}_1 and \mathbf{M}_2 close to a local energy minimum (well), the condition of barbell

algorithm $\frac{d\mathbf{M}_1}{dt} \cdot \mathbf{r} = -\frac{d\mathbf{M}_2}{dt} \cdot \mathbf{r}$ can also be satisfied. However, in this case $\frac{d\mathbf{M}_1}{dt}$ and $\frac{d\mathbf{M}_2}{dt}$ will be

pointing towards each other. This will eventually cause \mathbf{M}_1 and \mathbf{M}_2 to fall down to the energy wall instead of the saddle point.

To avoid such situation, the initial values of \mathbf{M}_1 and \mathbf{M}_2 must be carefully chosen. One approach is to do a sweeping algorithm and make snapshots of the system right before and after it passes the energy maximum, then let the two configurations of snapshots be \mathbf{M}_1 and \mathbf{M}_2 .

i. Switching Rate Calculation Results

From the saddle point obtained by barbell algorithm, we did a steepest descent path tracing by making the damping constant α infinitely large. In practice we remove the precessional term in LLG equation and just preserve the damping term. In this case $\gamma\alpha$ coefficient in LLG damping term is still finite but the ratio $\gamma\alpha/\gamma$ is infinite. The energy landscape of 20 computational cells (layers) $E(m_x)$ we computed is shown in Fig. 26.

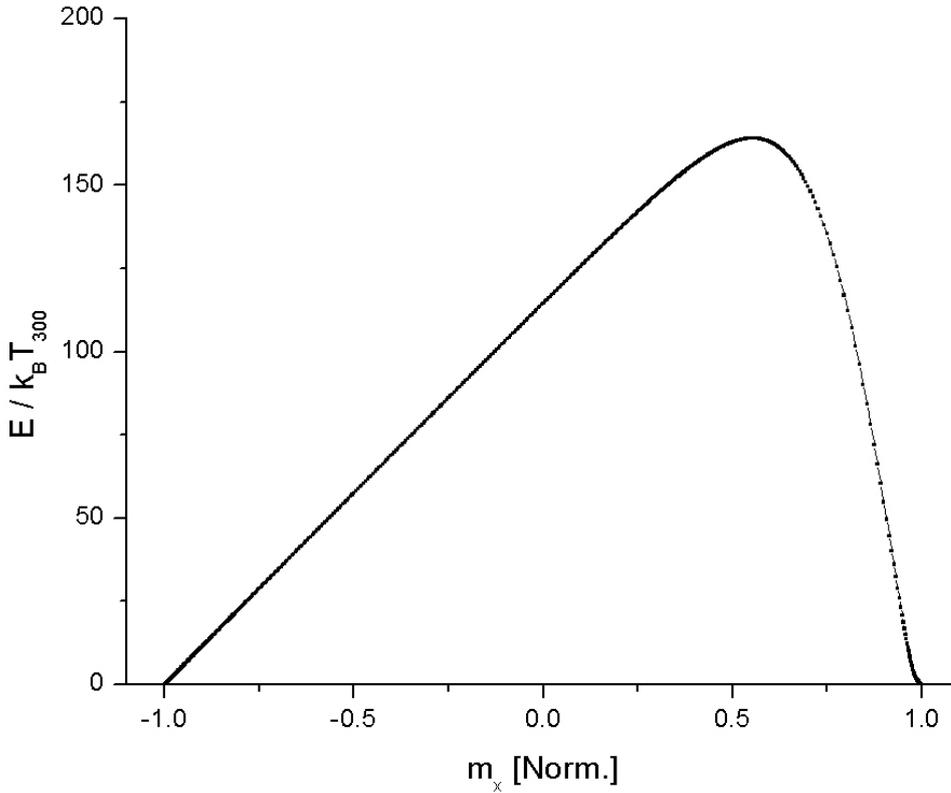


Figure 26 Energy landscape obtained by barbell algorithm.

In the equation (Eq. 3.28) for the prefactor, the only factor unknown now is the second derivative of energy as a function of m_x at the maximum (G^2 , defined before Eq. 3.24). This can be easily done by numerical fitting.

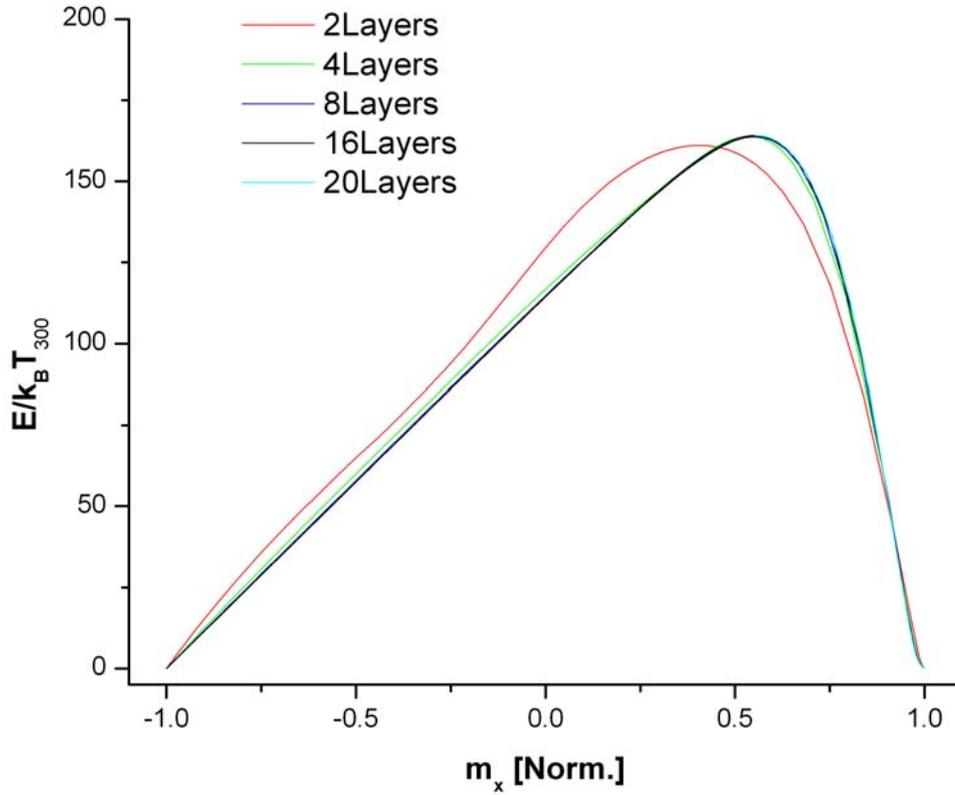


Figure 27 Energy landscape for different layered systems.

For comparison purpose, besides the 20-cell system we have modeled, we also did simulations of systems with fewer layers, while keeping the anisotropy of the hard end unchanged (9550 kA/m). Fig 27 shows energy landscape varying as the number of anisotropy graded layers changes, and Fig. 28 shows thermal switching rate prefactor as a function of number of anisotropy graded layers. The extreme high energy barriers shown in Fig. 27 are caused by the material parameters we have chosen. The anisotropy field we used in the simulation $H_K = 9550$ kA/m for the hard end is similar to that of bulk FePt, and is probably over optimistic for a thin film. But the results could be scaled to a lower H_K and the results would not be qualitatively different. One may also

notice that energy barriers for exchange coupled media and anisotropy graded media are only slightly different.

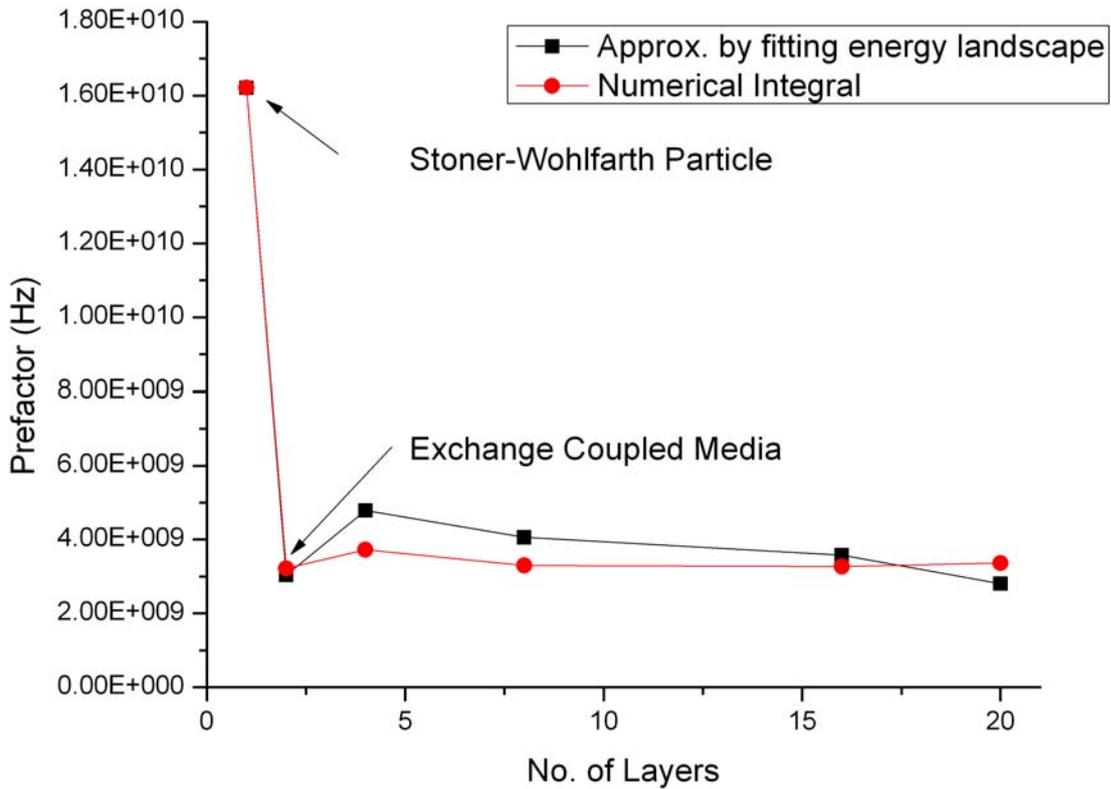


Figure 28 Thermal switching rate prefactor as function of number of layers. The black curve is obtained by parameters from parabolic fitting of energy landscape (see Eq. 3.28). The red curve is obtained by direct numerical integral of energy landscape (see Eq. 3.23). The switching rate prefactor of Stoner-Wohlfarth particle is obtained by analytical calculation of the energy landscape assuming that the energy barrier is the same as 20-cell anisotropy graded media system.

As we can see in Fig. 28, the thermal stability of anisotropy graded media is significantly better than that of single layered media, for the same energy barrier

CHAPTER 4 MICROMAGNETIC MODELING OF PERMANENT MAGNET NANOSTRUCTURES

a. History and Modern application of Permanent Magnets

Permanent magnet materials have been known for their applications in traditional electronic devices including power generators and actuators. In the coming era of new alternative energy, it regains attention because it will play an important role in devices such as electric automobile motors and wind turbines.

Historically, natural lodestones and high carbon steels were used for permanent magnets.

However, due to the low coercivity of these materials, magnetic domains can be easily formed by their own demagnetization field. The spontaneous magnetization of steel magnet is around 800kA/m, while its coercivity is around 5- 10kA/m. To solve the problem that magnetic domains are easily formed due to low coercivity reducing the total magnetic moment, people developed horseshoe, bar and needle shapes to reduce the demagnetization field.

To make a permanent magnet arbitrary shaped without breaking its uniformity, a coercivity close to its magnetization is required (the demagnetization field can be as large as the magnetization for a perpendicularly magnetized film). The figure of merit of a permanent magnet is its energy product $|BH|_{\max}$, which is twice the maximum energy stored in unit volume by a magnet [15].

Thus high performance magnetic materials must be developed to increase $|BH|_{\max}$.

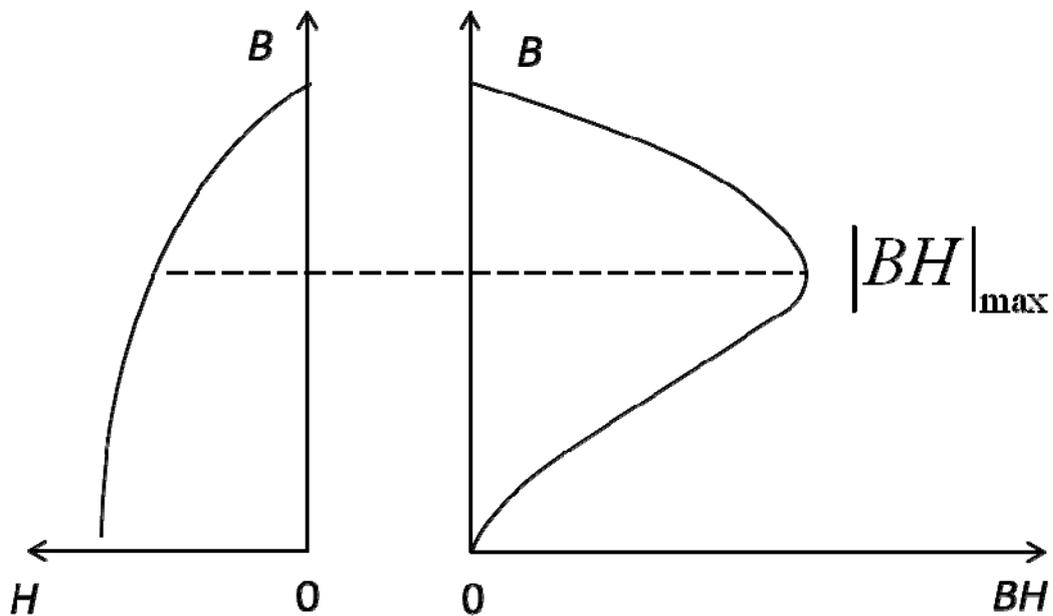


Figure 29 The demagnetization curve in B-H loop and maximum energy product $|BH|_{\max}$. The energy product $|BH|$ represents the free energy stored in free space around the magnet[38, 39].

This problem can be solved by rare earth magnets such as NdFeB and SmCo. For comparison purpose, table 2 lists a few commonly used permanent magnets.

Material	H_c (kA/m)	$ BH _{\max}$ (kJ/m ³)
FeCo, 5% Co	11	3.3
FeCo, 40% Co	21	8.2
BaFe ₁₂ O ₁₉ (Isotropic)	255 ~ 310	7.96 ~ 10.3
SrFe ₁₂ O ₁₉ (Anisotropic)	199 ~ 239	26.3 ~ 30.2
Sm ₂ Co ₁₇	550	250
Nd ₂ Fe ₁₄ B	960	360

Table 2 List of Parameters of Permanent Magnets[40] (H_c represents the coercivity of the magnet).

However, the use of rare earth material has been causing environment pollution, and in recent years, the control of rare earth material export in China has raised its price significantly. Thus the possibilities of alternative permanent magnets have been studied recently.

b. Modeling of MnAl-FeCo Two Phase Nanostructured Permanent Magnet

In this chapter we did the micromagnetic modeling of MnAl-FeCo permanent magnet nanostructure. This two-phase nanostructure has been proposed since the early 1990s[41]. The system is composed of an aligned hard phase and a soft phase. The hard phase (MnAl) has high uniaxial anisotropy ($H_k \sim 3000$ kA/m) and moderate saturation magnetization ($M_s = 800$ kA/m). The soft phase (FeCo) has low anisotropy (H_k is on the order of 10 kA/m, and taken as zero in simulation) and high saturation magnetization ($M_s = 1930$ kA/m). These two layers are exchange coupled at the interface. In field switching, the hard phase provides a strong exchange field, and prevents the soft phase from switching at low field. On the other hand, the strong saturation magnetization from the soft phase gives large remanent magnetization. So the key characteristic $|BH|_{\max}$ of this nanostructure should far exceed that of either single phase.

The simulation is carried out on the system described above and the geometry is shown as in Fig. 30.

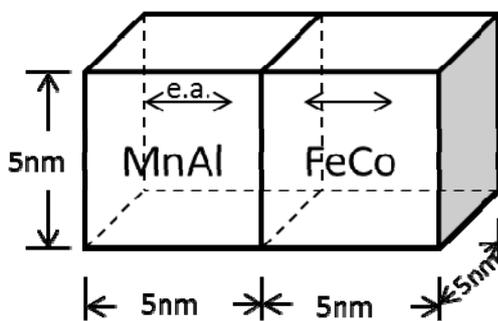


Figure 30 MnAl-FeCo two-phase nanostructured permanent magnet.

The simulation result gives M-H easy axis loop as shown in Fig. 31. In the figure, m_x represents average magnetization of the whole system. The deviation from a rectangle-shaped M-H loop at $H_x \sim 500$ kA/m is caused by the nucleation of a wall in the soft layer, and its pinning and depinning at the interface, as the anisotropy is different in two phases.

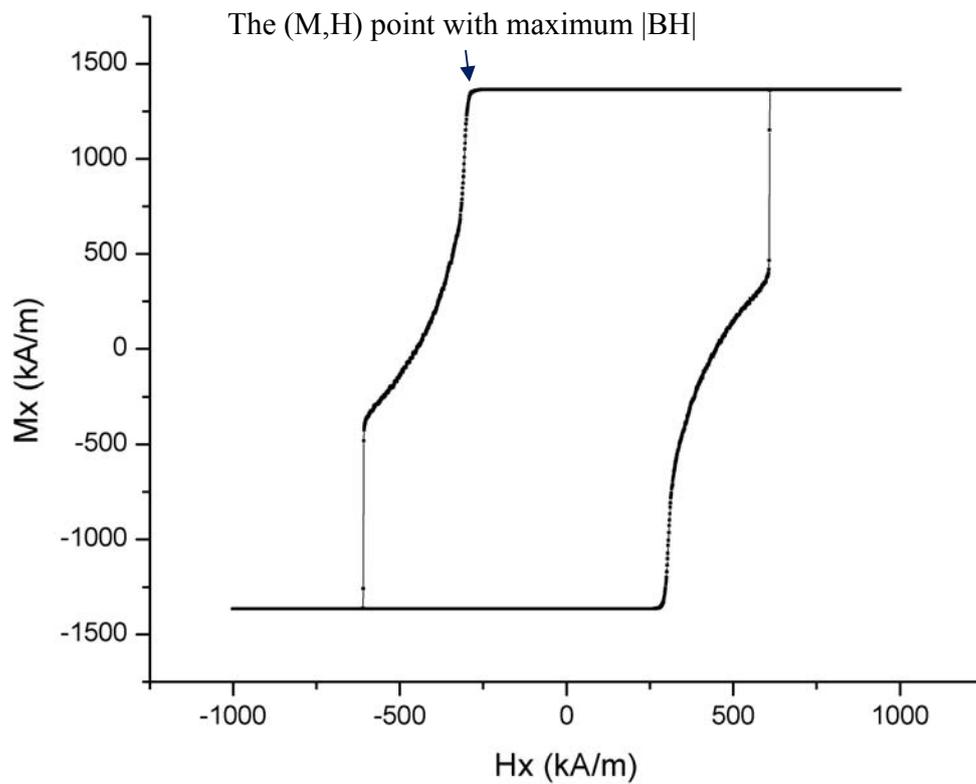


Figure 31 M-H loop of MnAl-FeCo permanent

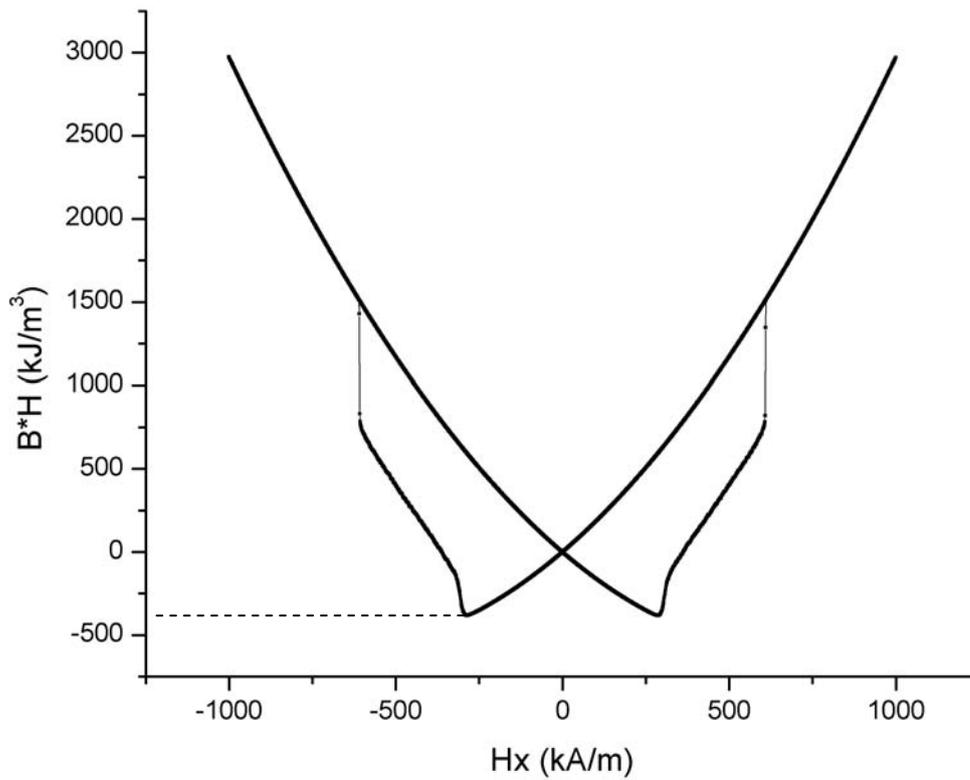


Figure 32 BH curve of MnAl-FeCo permanent magnet.

The $|BH|_{\max}$ can be easily calculated once we have the M-H loop. As shown in Fig. 32, the $|BH|_{\max} \sim 380 \text{ kJ/m}^3$ for this material, which corresponds to a point (M,H) with the maximum $|BH|_{\max}$ in second quadrant of Fig. 31. If we were trying to optimize this material, we might try decreasing the FeCo thickness so the hard layer can hold it at higher field. In this case the material will lose some M but gain H_c and the energy product might increase.

c. Energy Landscape Interpretation of Permanent Magnets' Switching Mechanism

Most previous work on micromagnetics of switching mechanism has fallen into two categories. The first one is for much idealized models, i.e. continuum models with simple geometry such as spherical inclusions, plane layers[42] for which nucleation fields can be computed. The other category is large-scale high-resolution calculations on more realistic models[43], such as core-shell particles, which are useful but time consuming. Besides, it does not provide much insight into determination of the switching mechanism. We feel that a different approach is needed to gain insight into switching of realistic systems.

So we apply the energy landscape method, which we have developed for graded anisotropy media to determine switching mechanisms and coercivity. These are the same things one wants to know in the context of permanent magnets – the difference is that in recording media one wants to minimize the coercivity, whereas in permanent magnets one wants to maximize it. In systems whose dimensions are large compared to the domain wall width, it is convenient to describe switching in terms of domain wall nucleation and pinning – however, in the nanostructured media, these concepts are only useful in a qualitative way, because the domain wall width is often comparable to the structure size. Roughly, however, in media one wants a low nucleation field for reversed-domain formation, whereas in permanent magnets one wants a high nucleation field, or no nucleation at all (coherent switching). This crossover between domain-wall (DW) switching and coherent Stoner-Wohlfarth (SW) switching is controlled by the strength A of the exchange interaction, or more precisely by A/KL^2 , where L is a characteristic length (e.g., layer thickness)[15]. Consider for simplicity a layered system (or a 1D media grain) consisting of two layers, each of thickness $L/2$. The configuration space is infinite-dimensional – even if we restrict our consideration to magnetizations that are a function only of the coordinate x

normal to the layers, and discretize the system into N layers for a micromagnetic calculation, the space is $2N$ -dimensional.

The system is shown in Fig. 32. The soft layer has a uniaxial anisotropy field $H_k = 796$ kA/m, while the hard layer is seven times harder ($H_k = 5571$ kA/m).

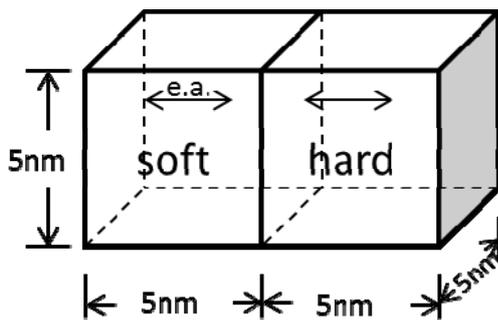


Figure 33 Illustration of two layer soft-hard composite permanent magnet nanostructure.

For strong exchange (exchange constant $A = 4 \times 10^{-11} J/m$), the minimum-energy switching path from one easy-axis direction to the other contains nearly-coherent configurations and passes over a saddle point at which the magnetization points nearly in the hard direction. This is shown in Fig. 33. The saddle point is located at where $Q_{xx} = 0$ which corresponds to infinitely wide domain wall or coherent switching.

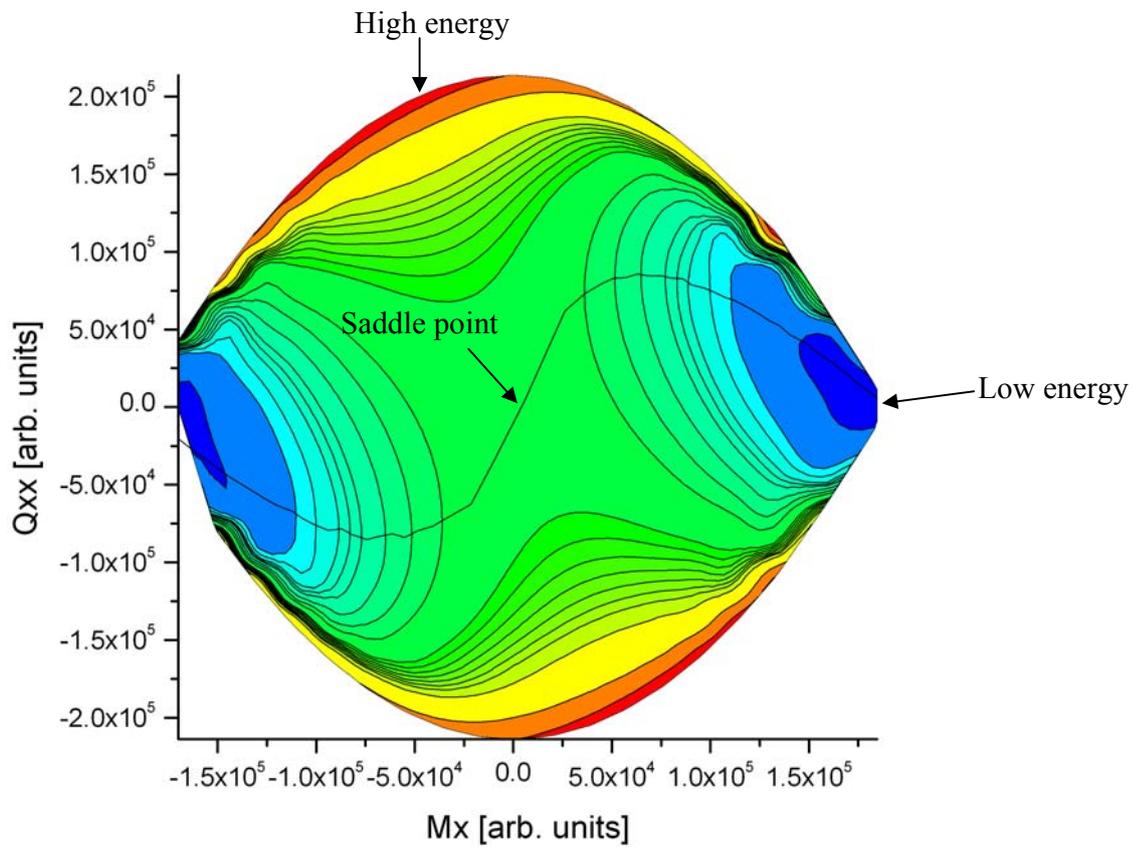


Figure 34 Coherent switching: Energy contours for a bilayer with $H_{K,hard} = 7 H_{K,soft}$ and strong exchange.

However, a smaller exchange constant will lead to domain wall switching, as shown in Fig. 34.

In this simulation the exchange constant is set to be $A = 1 \times 10^{-11} J / m$.

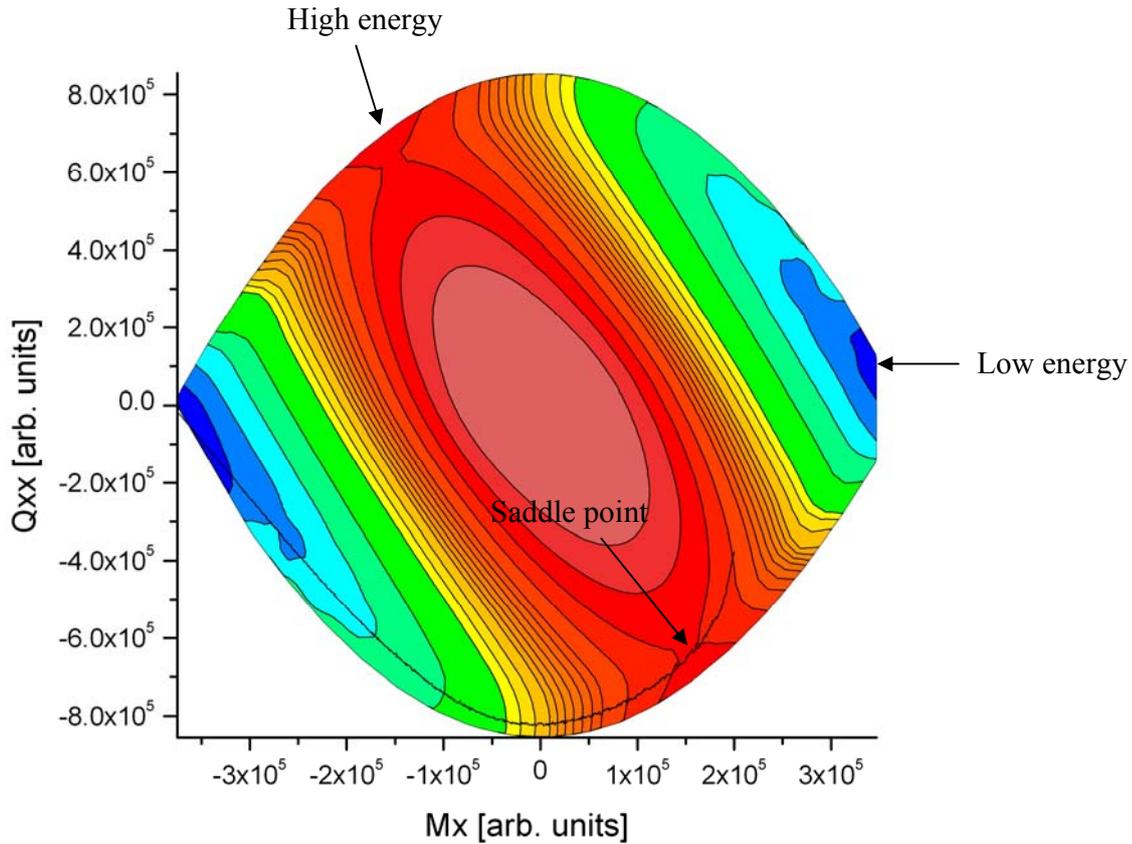


Figure 35 Domain wall switching in two layered permanent magnet nanostructure with a moderate exchange constant.

This approach is not just applicable to layered systems. It can be adapted to different systems by adjusting the order parameters we use to describe the system (dipole and quadrupole moments in the case described above.) For example, if we have a spherical core-shell particle, we would still use the dipole moment m_x as one of our variables for the energy landscape, but instead of the quadrupole moment, we would use the radial moment $m_r = \int r^2 M_x(x) d^3 r$. In a two-macrospin core-shell model, these would be linear combinations of the core and shell moments. However, the general landscape formulation will allow us to visualize the landscape and switching mechanism of a realistic micromagnetic model of a core-shell particle.

CHAPTER 5 SUMMARY AND POSSIBLE FUTURE WORK

In this dissertation we discussed several theoretical and numerical approaches to the energy landscapes of magnetic nanostructures and their application in solving problems such as thermal switching rate. We have developed a method for calculating a switching path in the 2N-dimensional configuration space, calculating the mobility and diffusivity (which define a 1D Langevin motion) along this path, and from these evaluating Kramers' formula for the switching rate. We proved that the thermal stability of anisotropy graded media is comparable to traditional single layered media and the previous suspicion of its instability is not correct. Although it is very difficult to make direct comparisons with direct simulation (because direct simulation is very difficult for these very slow rates), it is likely that this method is exact in the limit of low temperatures (large $E_b/k_B T$), because the deviations from the minimum-energy path vanish in this limit. To the best of our knowledge the multi-dimensional Newton-Raphson method has not been proposed before for this application. Further application of this method can be used on more complicated micromagnetics, such as the modeling of vortex mode of a domain wall.

In the future we could adjust the material parameter in the modeling to more realistic values from experiments and also evaluate the change introduced by including magnetostatic field in the calculation. Thus the simulation results of exchange coupled media and anisotropy graded media can be directly compared with those from experiments.

References

- [1] W. F. Brown, *Physical Review* **130**, 1677 (1963).
- [2] S. Fischer and M. Karplus, *Chemical Physics Letters* **194**, 252 (1992).
- [3] H. A. Kramers, *Physica* **7**, 284 (1940).
- [4] G. Henkelman, B. P. Uberuaga, and H. Jonsson, *The Journal of Chemical Physics* **113**, 9901 (2000).
- [5] M. Julliere, *Physics Letters A* **54**, 225 (1975).
- [6] H. Hidaka, *Embedded Memories for Nano-Scale VLSIs*, 241 (2009).
- [7] J. C. Slonczewski, *Journal of Magnetism and Magnetic Materials* **159**, L1 (1996).
- [8] S. Mangin *et al.*, *Nature Materials* **5**, 210 (2006).
- [9] J. Z. Sun *et al.*, *Journal of Applied Physics* **93**, 6859 (2003).
- [10] D. M. Apalkov and P. B. Visscher, *Physical Review B* **72**, 180405 (2005).
- [11] J. Z. Sun, *IBM journal of research and development* **50**, 81 (2006).
- [12] Y. B. Bazaliy, B. A. Jones, and S. C. Zhang, *Physical Review B* **69**, 94421 (2004).
- [13] P. B. Visscher, in *Proceedings of SPIE (Proc. SPIE, San Diego, CA, USA, 2008)*, p. 70360B.
- [14] J. Cucchiara *et al.*, *Applied Physics Letters* **94**, 102503 (2009).
- [15] R. C. O'Handley, *Modern magnetic materials: principles and applications* (Wiley New York, 2000).
- [16] S. Batra *et al.*, *IEEE Tran. Mag.* **40**, 319 (2004).
- [17] H. N. Bertram and J. G. Zhu, *IEEE Tran. Mag.* **27**, 5043 (1991).
- [18] J.-G. Zhu and H. Wang, *IEEE Tran. Mag.* **31**, 1065 (1995).

- [19] R. Wood, IEEE Tran. Mag. **36**, 36 (2000).
- [20] R. M. Metzger *et al.*, IEEE Tran. Mag. **36**, 30 (2000).
- [21] J. J. M. Ruigrok *et al.*, Journal of Applied Physics **87**, 5398 (2000).
- [22] M. Alex *et al.*, IEEE Tran. Mag. **37**, 1244 (2001).
- [23] C. A. Ross, Annual Review of Materials Research **31**, 203 (2001).
- [24] R. M. H. New, R. F. W. Pease, and R. L. White, Journal of magnetism and magnetic materials **155**, 140 (1996).
- [25] M. Todorovic *et al.*, Applied Physics Letters **74**, 2516 (1999).
- [26] G. F. Hughes, IEEE Tran. Mag. **36**, 521 (2000).
- [27] V. Skumryev *et al.*, nature **423**, 850 (2003).
- [28] R. H. Victora and X. Shen, IEEE Tran. Mag. **41**, 537 (2005).
- [29] D. Suess, Applied Physics Letters **89**, 113105 (2006).
- [30] Z. Lu, P. B. Visscher, and W. H. Butler, IEEE Tran. Mag. **43**, 2941 (2007).
- [31] P. G. Bolhuis, C. Dellago, and D. Chandler, Faraday Discussions **110**, 421 (1998).
- [32] W. Coffey, Y. P. Kalmykov, and J. T. Waldron, *The Langevin equation: with applications to stochastic problems in physics, chemistry, and electrical engineering* (World Scientific Pub Co Inc, 2004).
- [33] T. A. Halgren and W. N. Lipscomb, Chemical Physics Letters **49**, 225 (1977).
- [34] M. J. Rothman and L. L. Lohr Jr, Chemical Physics Letters **70**, 405 (1980).
- [35] H. Jonsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne (World Scientific, 1998).
- [36] J. Baker, Journal of Computational Chemistry **7**, 385 (1986).
- [37] W. Scholz *et al.*, Computational Materials Science **28**, 366 (2003).
- [38] J. D. Jackson, *Classical Electrodynamics Third Edition* (John Wiley, 1999).
- [39] J. M. D. Coey, Solid State Communications **102**, 101 (1997).

- [40] M. Yan, *Fundamentals of Magnetism and Magnetic Material* (Zhejiang University Press, 2006) (In Chinese).
- [41] R. Skomski and J. M. D. Coey, *Physical Review B* **48**, 15812 (1993).
- [42] R. Skomski and J. M. D. Coey, *Permanent magnetism* (Institute of Physics Publishing Bristol and Philadelphia, 1999).
- [43] N. N. Phuoc and T. Suzuki, *IEEE Tran. Mag.* **41**, 2715 (2005).