Multiscale approach to spintronics - from nanomagnet to topological excitations

---

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment
of the requirements for the degree

Doctor of Philosophy

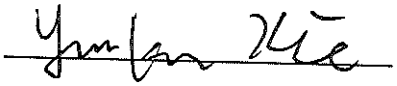
by

Yunkun Xie

May 2019

APPROVAL SHEET

This Dissertation
is submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Author Signature: 

This Dissertation has been read and approved by the examining committee:

Advisor: Avik Ghosh

Committee Member: Mircea Stan

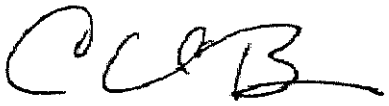
Committee Member: Andreas Beling

Committee Member: Bobby Weikle

Committee Member: Petra Reinke

Committee Member: _____

Accepted for the School of Engineering and Applied Science:



Craig H. Benson, School of Engineering and Applied Science

May 2019

Abstract

The rapid growth of information and data has been pushing the limit of our data storage and processing capabilities. At the hardware level, the physical scaling of the transistor technology has been increasingly difficult and expensive. Magnet, with its long history of being used in information storage, has re-branded itself through the so-called spintronics, which studies how to control and manipulate the magnetic moment (or electron spin) for modern computation. Unlike silicon, researchers have yet to identify an ideal material nor find a perfect way to take advantage of the extra degree of freedom from electron spins. This dissertation describes our study of the prospect and limitation of several spintronic systems as well as our methodology - a multiscale approach that goes from atomistic material properties to device-level performance. At the material level, we use the Non-Equilibrium Green's Function (NEGF) simulations to study the spin transport in nanostructures such as the magnetic tunnel junction, investigating the effect of different electrode materials and interfaces on its read/write performance. At the device level, we implemented a fast, general solver to simulate the thermal noise-induced switching error in nanomagnets, revealing its correlations with various material parameters and magnetic configurations. With an understanding of the limitations of magnetic tunnel junctions, we examined two emerging topological systems - topological insulator (TI) and magnetic skyrmion. In studying the topological insulator, we proposed an alternative way to verify the Klein-Tunneling physics in a PN junction setup on the TI surface. We also discussed its potential as a spin source. For skyrmions, we analyzed the conditions to achieve ultrasmall, fast, and stable skyrmions and offered few material candidates with their corresponding skyrmion phase diagrams.

Acknowledgments

For this long journey, I want to express my gratitude to my advisor, Professor Avik Ghosh, for instilling in me the qualities of being a good researcher and engineer. His infectious enthusiasm and unlimited passion have been major driving forces through my graduate career. He has been very patient and has given me advice beyond my research. It has been quite a journey.

I also want to thank Prof. Mircea Stan (ECE), Prof. Andreas Beling (ECE), Prof. Bobby Weikle (ECE), and Prof. Petra Reinke (MSE) for serving on my proposal and defense committee and providing me with constructive advice. I have taken courses from some of them and was impressed by their knowledge and devotion to teaching.

I would like to also thank my fellow graduate students, past and present, from the VINO group — for their companionship, feedback, and friendship. Some of them have guided me through the beginning of my PhD; some of them are with me from the start to the end; some of them have given me the chance to pass along the traditions of our group.

Finally but most importantly, I would like to thank my family and old friends. I thank my parents for their wise counsel and sympathetic ear all along my Ph.D. journey. I thank my dear wife for her never-ending love and caring, for the strength and hope she gave me. Last but not least, I want to thank my old friends, who are now scattered around the globe but always a phone call away, for their support and all the possibilities they've shown me.

Contents

List of Tables

List of Figures

List of Abbreviations

DFT	Density Functional Theory
FPE	Fokker-Planck Equation
GSHE	Giant Spin Hall Effect
KT	Klein Tunneling
NEGF	Non-Equilibrium Green's Function
STT	Spin Transfer Torque
TI	Topological Insulator
TIPNJ	Topological Insulator PN Junction
TMR	Tunnel Magneto-Resistance
WER	Write Error Rate

Chapter 1

Introduction

1.1 Background

The physical and electrical scaling of transistor prophesied by Moore [6] has fueled the semiconductor industry for decades. We have reached a point where each upgrade proves to be increasingly hard due to complications in the fabrication process and side effects such as increased power leakage and endurance problems from the super densely packed transistors. At the same time, the explosive growth of information and data from emerging technologies/markets, such as cloud computing (and services), 5G and Internet of Things, medical monitoring systems, autonomous cars, and so on, brings opportunities as well as challenges to the semiconductor industry as the demand for larger and faster storage, energy efficiency, persistent memory, and near data processing increases. Current storage hierarchy, consisting mostly of Static Random Access Memory (SRAM), Dynamic Random Access Memory (DRAM), and hard disk (or solid-state disk) are either volatile or slow. A paradigm shift is needed to address the needs of those emerging technologies.

Magnets are seen as a natural information recording medium and have been widely used since the advent of the first generation of modern computers via the magnetic-core memory. Magnetic hard disks are still the top choice for large data storage due to its high capacity, endurance, and cheap cost. However, the traditional way to write information into a magnet is through a magnetic field, which is not scalable due to the growing current requirement for switching scaled magnets. It is also hard to localize the magnetic field to prevent accidental

switch of adjacent cells in a densely packed array. In the past decade, the discovery of new physics and the development of material research have offered various ways to manipulate a magnet that overcomes the limitation of magnetic field-based switching. Spintronics, a portmanteau combining the electronics and its associated spin freedom, describes an approach to control and manipulate electron spins to store, transmit, and process information. One prominent milestone in the short history of spintronics is the discovery of the Spin Transfer Torque (STT) effect [7, 8], which states that a spin-polarized current can transfer its angular momentum to a magnet via a torque. With enough current, the magnetic moment can be flipped. This is a more scalable solution to the write process because the required current to flip a magnet scales with the volume of the magnet. Combined with the Tunnel Magneto-Resistance (TMR) effect for reading information from a magnet, the spin transfer torque magnetic random access memory (STT-MRAM) is believed to replace the conventional MRAM at advanced technology nodes. Recently, several vendors have already started shipping commercial STT-MRAM prototype for sampling. While material research and advancement in the fabrication process will continue to improve STT-MRAM performance, it is hardly the ideal universal memory people are searching for in its current status. To start with, although STT offers a much more energy efficient write process than conventional field-based switching, it still consumes a large current on the order of $10^6 - 10^7\,\mathrm{A/cm^2}$ to switch a magnet. The energy dissipation can be two orders of magnitude higher than a transistor switching. When it comes to integration with the CMOS technology, how to supply a sufficient current with a transistor of the same size can be an issue too. The state of art research focuses on how to inject spins more efficiently to guarantee speed and reliability while cutting down the charge current. There are efforts to explore electrical field induced switch in multiferroic materials [9, 10] or field-assisted switch in voltage-controlled magneto-anisotropy systems (VCMA) [11]. These systems try to minimize the electrical current while maintaining the reliability of the magnet. Other efforts aim at enhancing the conversion of charge current to spin current through the spin-orbit coupling, e.g. in Giant Spin Hall Effect (GSHE) [12], or Topological Insulator.

Recent development in understanding various topological excitations has spurred great interests in exploiting the topology of materials for spintronics applications that could

potentially lead to better devices. Two of those systems will be explored in this dissertation - Topological Insulator (TI) and Magnetic Skyrmion. For topological insulator, its unique bandstructure can provide a way to generate spin-polarized current and its efficiency can be tuned in a gated PN junction structure. Magnetic skyrmions are topological excitations that can be used as the information carrier. Magnetic skyrmions share similar properties as magnetic domains: they can be nucleated, annihilated, and driven by various mechanisms. Magnetic skyrmions can stabilize in smaller sizes and are essential zero-dimensional objects compared to their counterpart, which could mean higher packing density and other advantages.

1.2 Dissertation Organization

In the remainder of the dissertation, we will discuss the development of computational tools as well as a broad physical understanding of the magnetic tunnel junction, topological insulator, and magnetic skyrmion. This will culminate with a multiscale approach that connects material properties with device performance. In chapter two, we will start by exploring the key component in spintronics applications - the magnetic tunnel junctions. We first introduce a simple free electron transport model for a qualitative understanding of the MTJ. Then we move to a full atomistic Non-Equilibrium Green's Function (NEGF) based approach to calculate the I-V and STT in real material systems. Following that, we discuss the reliability, mainly focusing on the Write Error Rate (WER), of STT switching. We demonstrate a general, fast 2D Fokker-Planck solver for STT switching and apply it to various write scenarios. We conclude this chapter with a simple analytical equation to capture the WER in spin-torque switching, which hints us on why STT switching in MTJs is hardly energy efficient. In chapter three and four, we extend our approach to study two topological systems: For topological insulator (chapter three), we investigate the unique transport properties (Klein Tunneling) of electron/spins in TI PN junction and propose an experiment to detect it. We will also discuss its potential and limitations as a spin source in spintronics applications. For magnetic skyrmions (chapter four), we will look for ways to

achieve stable, ultra-small, and fast skyrmion through analyzing its energetics followed by a search in the Heusler family for ideal material candidates.

Chapter 2

Building Block - Magnetic Tunnel Junction

2.1 Simple MTJ Model

2.1.1 Read/Write Information in Magnetic Tunnel Junction

Magnetic tunnel junction is an essential part of most spintronics applications. It has a sandwich structure consisting of two ferromagnetic layers separated by a thin non-magnetic insulator as shown in Fig. 2.1. One of the magnetic layers has a pinned (or fixed) magnetic moment that is hard to flip. The other magnetic layer is called the 'free' layer with a magnetic moment easier to flip. Both magnetic layers are engineered to have uniaxial anisotropies collinearly aligned with each other. Two energetically stable states emerge: the two magnetic layers either have parallel (P) magnetic moments or anti-parallel (AP) magnetic moments. Binary information can be stored in those states. For readout, a small current is passed through the tunnel junction. The MTJ has lower resistance in the parallel state and higher resistance in the anti-parallel state. By comparing the junction voltage with a reference voltage, one can detect the status of the magnetic moment in the free layer. For the write operation, there are two scenarios: To switch the MTJ from AP to P, a large current is passed from the pinned layer to the free layer. The pinned layer first polarizes the electron spins. After that, the polarized electrons tunnel through the insulator and interact

with the magnetic moment of the free layer via the exchange coupling. This interaction will try to align the incoming spins with the local spins (magnetization), therefore exerts a torque on the local magnetic moment. If the torque is large enough to overcome the energy barrier from the anisotropy of the free layer, the magnetic moment is reversed. To switch the magnet from P to AP, the direction of the current is reversed. The majority spins (those aligned with the free layer) pass through the MTJ much easier than the minority spins. When the accumulation of minority spins at the barrier/free layer interface exceeds a threshold, the magnetic moment of the free layer can be flipped. These two processes are shown schematically in Fig. 2.1

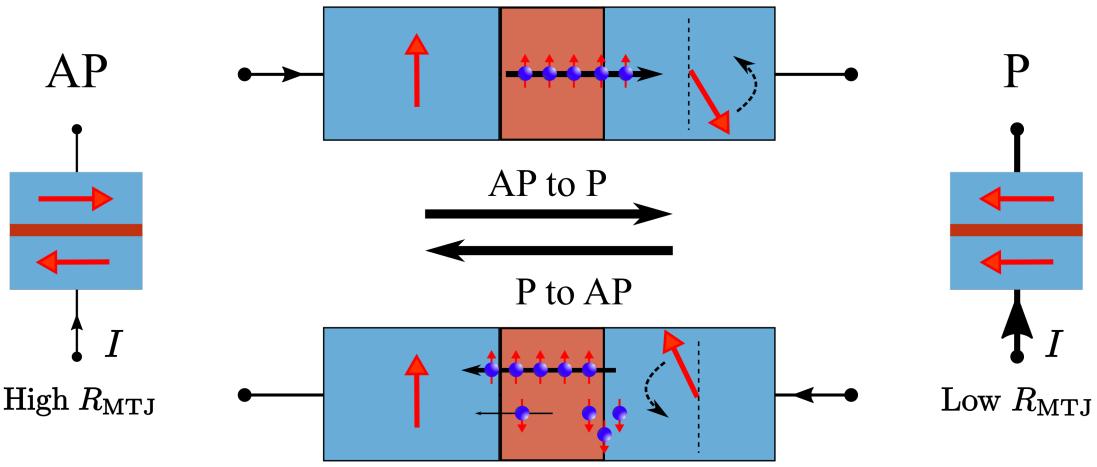


Figure 2.1: Schematics of the structure and basic read/write processes in a magnetic tunnel junction.

The performance of an MTJ is determined by two related phenomena - TMR for read and STT for write. Both phenomena depend on the material properties of the ferromagnetic as well as the non-magnetic layer. In the next section, we will use a modified Simmons' model to work out the I-V of MTJs within the free electron approximation for the magnetic layers.

2.1.2 Free Electron Model for MTJ

The I-V characteristics of MTJs can be qualitatively understood in terms of a simple free electron model with a spin-dependent barrier (Slonczewski, 1989 [7]). Consider a

magnetic tunnel junction. In the free-electron approximation, the longitudinal part of the spin-polarized electron Hamiltonian across the MTJ can be written as:

$$\begin{aligned} H &= \frac{\hbar^2 k_\perp^2}{2m} - \frac{1}{2}\vec{\Delta}_s \cdot \vec{\sigma}, \qquad x < 0 \ \text{ or } \ x > d \qquad \text{(FM layers)} \\ H &= \frac{\hbar^2 k_\perp^2}{2m^*} + U(x), \qquad 0 \le x \le d \qquad \text{(Nonmagnetic barrier)} \end{aligned} \tag{2.1}$$

where m and m^* are the electron effective masses in the ferromagnets and the barrier respectively, $\vec{\Delta}_s$ is the exchange field and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. If we choose the local z axis to be along the magnetization direction, the energy dispersions of the longitudinal part for two spin channels are two parabolic bands shifted with respect to each other, as shown in fig 2.2. In the following we will simply write $k_{\perp,\uparrow}, k_{\perp,\downarrow}$ as $k_\uparrow, k_\downarrow$.

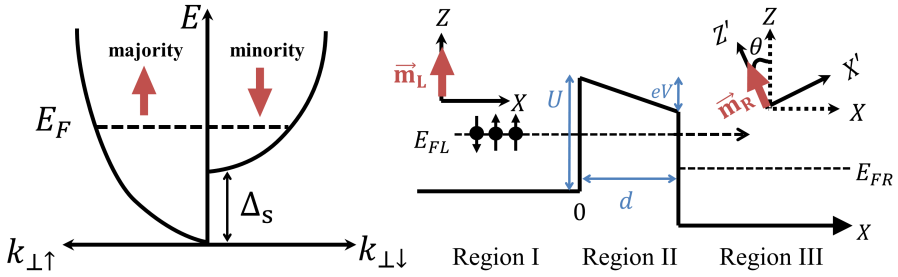


Figure 2.2: Simple barrier model for magnetic tunnel junction. Left: The band structure of the ferromagnetic contact. The bottom of ↑ and ↓ conduction bands in ferromagnetic contacts are separated by Δ_s. E_F is the Fermi energy. Right: When a bias is applied on the MTJ, the insulating barrier has a linear ramp potential. d is the width of the insulating barrier and U is the barrier offset between the contact and the spacer. The magnetization of the right contact is rotated to an angle θ from the magnetization of the left contact.

The magnetic tunnel junction can be broken up into three regions: (I) $x < 0$: the left ferromagnetic layer where the magnetization is pinned to the +z axis, (II) $0 \le x \le d$: the nonmagnetic tunnel barrier and (III) $x > d$: the right ferromagnetic layer whose magnetization is free to rotate and is defined by the angle θ measured with respect to the positive +z axis. The magnetization of the right ferromagnet is parallel to the z' axis of the coordinate system x', y', z', which is rotated at θ degrees to the original z axis. For simplicity, we omit the transverse momentum $k_{||}$ and solve for the longitudinal part. The

wave functions in the three regions can be written as:

$$\begin{aligned}
&\text{Region I: } \psi_\uparrow = \tfrac{1}{\sqrt{k_\uparrow^l}} e^{ik_\uparrow^l x} + R_\uparrow e^{-ik_\uparrow^l x} \\
&\qquad\qquad \psi_\downarrow = R_\downarrow e^{-ik_\downarrow^l x} \\
&\text{Region II: } \psi_\sigma = \tfrac{1}{\sqrt{\kappa(E,x)}} \left[A_\sigma e^{-E_b(x)} + B_\sigma e^{E_b(x)} \right] \\
&\text{Region III: } \psi'_\sigma = C_\sigma e^{ik_\sigma^r x} \quad \sigma = \uparrow, \downarrow
\end{aligned} \tag{2.2}$$

The longitudinal spin-polarized electron momentum in each of the three regions can be expressed as

$$\begin{aligned}
\text{Region I: } & k_\uparrow^l = \frac{1}{\hbar}\sqrt{2mE}, k_\downarrow^l = \frac{1}{\hbar}\sqrt{2m(E-\Delta_s)} \\
\text{Region II: } & \kappa(E,x) = \frac{1}{\hbar}\sqrt{2m^*[U - eVx/d - E]}, \\
& E_b(x) = \int_0^x \kappa(E,x')dx' \\
\text{Region III: } & k_\uparrow^r = \frac{1}{\hbar}\sqrt{2m(E+eV)}, k_\downarrow^r = \frac{1}{\hbar}\sqrt{2m(E-\Delta_s+eV)}
\end{aligned} \tag{2.3}$$

Notice that the wave function in region III, $\psi'_\uparrow$ and $\psi'_\downarrow$, is written with respect to the local axes x', y', z'. In order to conform to the original axes, a spinor transformation is required,

$$\begin{aligned}
\psi_\uparrow &= \cos(\tfrac{\theta}{2})\psi'_\uparrow + \sin(\tfrac{\theta}{2})\psi'_\downarrow \\
\psi_\downarrow &= -\sin(\tfrac{\theta}{2})\psi'_\uparrow + \cos(\tfrac{\theta}{2})\psi'_\downarrow
\end{aligned} \tag{2.4}$$

By matching $\psi_{\uparrow,\downarrow}$ and $d\psi_{\uparrow,\downarrow}/dx$ at $x = 0$, d, the unknowns $R_{\uparrow,\downarrow}$, $A_{\uparrow,\downarrow}, B_{\uparrow,\downarrow}$ and $C_{\uparrow,\downarrow}$ can be solved for. The charge current is obtained from,

$$J_e = \frac{e\hbar}{2m^* i}\left[\begin{pmatrix} \psi_\uparrow^* & \psi_\downarrow^* \end{pmatrix} \begin{pmatrix} d\psi_\uparrow/dx \\ d\psi_\downarrow/dx \end{pmatrix} - \begin{pmatrix} \psi_\uparrow & \psi_\downarrow \end{pmatrix} \begin{pmatrix} d\psi_\uparrow^*/dx \\ d\psi_\downarrow^*/dx \end{pmatrix} \right] \tag{2.5}$$

Since the charge current is conserved throughout the junction, the equation can be evaluated at any point. Solving for the charge current to leading order in $e^{-E_b(d)}$, we get

$$J_e(E) = J_0(1 + P^2 \cos\theta) \tag{2.6}$$

where

$$J_0(E) = \frac{8e\hbar\kappa_L\kappa_R}{m^*}\frac{(\kappa_L^2+k_\uparrow^l k_\downarrow^l)(k_\uparrow^l+k_\downarrow^l)}{(\kappa_L^2+{k_\uparrow^l}^2)(\kappa_L^2+{k_\downarrow^l}^2)}\frac{(\kappa_R^2+k_\uparrow^r k_\downarrow^r)(k_\uparrow^r+k_\downarrow^r)}{(\kappa_R^2+{k_\uparrow^r}^2)(\kappa_R^2+{k_\downarrow^r}^2)}e^{-2E_b(d)} \tag{2.7}$$

$$\begin{aligned} P(E)^2 =& \frac{(\kappa_L^2-k_\uparrow^l k_\downarrow^l)(k_\uparrow^l-k_\downarrow^l)}{(\kappa_L^2+k_\uparrow^l k_\downarrow^l)(k_\uparrow^l+k_\downarrow^l)}\cdot\frac{(\kappa_R^2-k_\uparrow^r k_\downarrow^r)(k_\uparrow^r-k_\downarrow^r)}{(\kappa_R^2+k_\uparrow^r k_\downarrow^r)(k_\uparrow^r+k_\downarrow^r)} \\ =& P^l(\kappa_L,k_\uparrow^l,k_\downarrow^l)P^r(\kappa_R,k_\uparrow^r,k_\downarrow^r) \end{aligned} \tag{2.8}$$

$P^i = (\kappa_i^2-k_\uparrow^i k_\downarrow^i)(k_\uparrow^i-k_\downarrow^i)(\kappa_i^2+k_\uparrow^i k_\downarrow^i)^{-1}(k_\uparrow^i+k_\downarrow^i)^{-1}, \quad (i=l,r)$ are 'effective polarization' because it is the product of the FM contact polarization $(k_\uparrow^i-k_\downarrow^i)(k_\uparrow^i+k_\downarrow^i)^{-1}$ and the coupling between the spacer and the FM contact, $(\kappa_i^2-k_\uparrow^i k_\downarrow^i)(\kappa_i^2+k_\uparrow^i k_\downarrow^i)^{-1}$. The TMR is defined in the following and it can be related to the effective polarizations of the FM layers:

$$\text{TMR} = \frac{R_{ap}-R_p}{R_p} = \frac{I_P-I_{AP}}{I_{AP}} \propto \frac{2P_LP_R}{1-P_LP_R} \tag{2.9}$$

Eq. 2.9 states that higher spin-polarized magnetic layers lead to higher TMR, which improves the preformance of the readout in MTJs. In the extreme case of half-metallic magnets, which have 100% spin-polarized conducting electrons, the TMR diverges due to $R_{ap}\to\infty$.

2.2 *ab-initio* Spin Transport in MTJs

2.2.1 Motivation for atomistic MTJ model

While the simple free electron model captures the principle of the MTJ, it ignores some physics that can only be captured by taking into account the full bandstructure. For instance, early adoption of MTJs mainly used Al_2O_3 as the insulating layer, which has been replaced by MgO later for improved TMR. The reason for the improved TMR is from the 'symmetry filtering' effect predicted from atomistic calculations [13]. Fig. 2.3 shows part of the bandstructure of bcc Fe near the Fermi energy. The spin polarization of bulk Fe is not impressive due to the existence of several bands in both spin channels crossing the Fermi energy. However, in a magnetic tunnel junction with MgO as the tunneling insulator, the tunneling effect from one Fe layer to the other is constrained by the symmetry of the bands.

Those Fe bands with different space symmetries from the MgO bands have different decay lengths inside the MgO. It turns out that the Δ_1 band has the greatest transmission and dominates the tunneling current. Fig. 2.3 indicates the dominating Δ_1 band is indeed highly spin-polarized near the Fermi level.

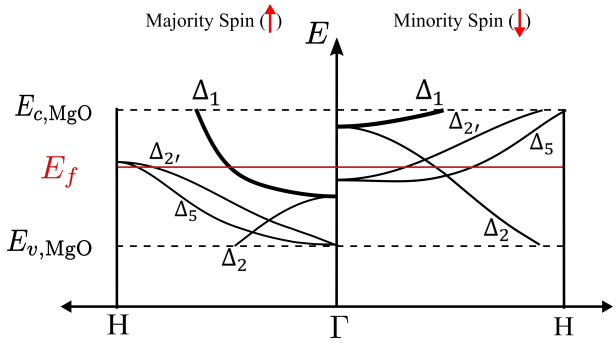


Figure 2.3: Band origin of the symmetry filtering. The band structure of bulk bcc Fe is plotted in the energy window of the MgO bandgap. The Δ_1 bands in both spin channels are highlighted.

Therefore it is necessary to have a full atomistic model to study the transport phenomena such as TMR and STT when dealing with real material candidates. In the next section, we write down the central equations for the NEGF approach to calculate the charge/spin transport in two terminal structures. One advantage of the NEGF approach is that it can couple with either full atomistic model or simplified tight-binding model. We will apply the NEGF approach to study MTJs in this chapter and topological insulator in chapter 3. It is a key element in connecting equilibrium material properties to non-equilibrium transport in our multiscale framework. Part of this work was published in ref. [14] with coauthors Ivan Rungger, Kamaram Munira, Maria Stamenova, Stefano Sanvito, and Avik W. Ghosh. My contributions are summarized at the end of this chapter.

2.2.2 Non-Equilibrium Green's Function Approach

We formulate the NEGF formalism starting from the Hamiltonian. The NEGF Hamiltonian is obtained either from Density Functional Theory (DFT) program or from a simpler tight-binding (TB) model, projected over a localized orbitals basis set. The obtained single particle Hamiltonian H, and the calculated density matrix (shown shortly) ρ, can be written as a

sum of a spin-dependent part and a spin-independent part

$$\begin{aligned} H &= H_0 \cdot \mathbb{I} + \vec{H}_{\mathrm{S}} \cdot \vec{\sigma} \\ \rho &= \rho_0 \cdot \mathbb{I} + \vec{\rho}_{\mathrm{S}} \cdot \vec{\sigma} \end{aligned} \tag{2.10}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, and $\mathbb{I}$ is the 2×2 unity matrix. H_0 is the spin-independent part of H, while $\vec{H}_{\mathrm{S}} = (H_x, H_y, H_z)$ represents the spin-components of H corresponding to the exchange field. In the same way ρ is split into its spin-independent part ρ_0 and its spin-vector $\vec{\rho}_S = (\rho_x, \rho_y, \rho_z)$. Note that H_α and ρ_α, with $\alpha \in \{0, x, y, z\}$, are $N_{\mathrm{o}} \times N_{\mathrm{o}}$ matrices, with N_{o} being the number of orbitals in the simulation cell.

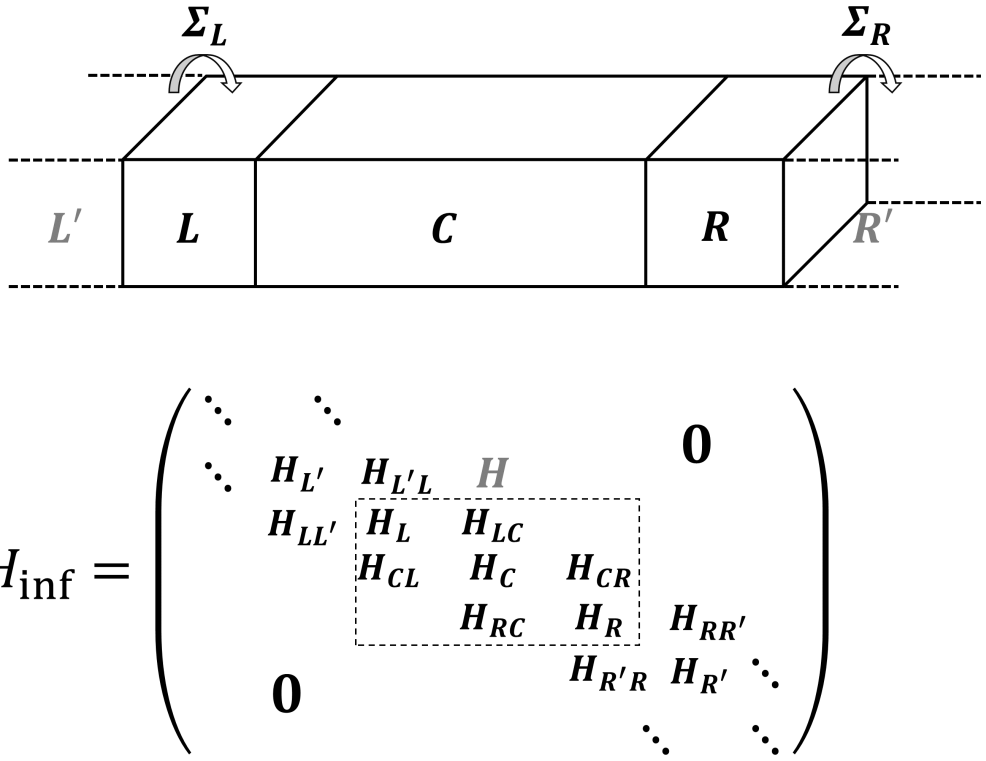


Figure 2.4: NEGF Hamiltonian. The complete Hamiltonian is an infinite matrix due to the two semi-infinite left/right contacts. The device Hamiltonian H (enclosed in dashed box) is chosen to include the whole channel and part of the contacts (L, R) while the rest of the semi-infinite contacts are modeled by self-energies Σ_L, Σ_R. $\Sigma_{L,R}$ are calculated from the contact Hamiltonian block $H_L, H_{L'L}$ and $H_R, H_{R'R}$.

calculation of density matrix ρ

The density matrix is calculated from the Hamiltonian plus the boundary condition. As shown in Fig. 2.4, in a typical two-terminal device setup, the two boundaries (contacts) are assumed to be infinitely extended electron reservoirs that maintain their own electron distribution according to the Fermi-Dirac distribution function f_L, f_R respectively. The retarded Green's function $G(E)$ as a function of energy is given by:

$$\begin{aligned} G(E) =& \Big(ES - H - \Sigma_L(E) - \Sigma_R(E)\Big)^{-1}, \\ \Gamma_{L,R}(E) =& \, i\Big(\Sigma_{L,R}(E) - \Sigma_{L,R}^{\dagger}(E)\Big) \end{aligned}$$

where S is the overlap matrix for non-orthogonal basis and it reduces to an identity matrix for orthogonal basis. $\Sigma_{L,R}$ are *self-energy* matrices [15]. From those quantities, we can define the lesser Green's function, $G^{<}(E)$, as

$$G^{<}(E) = iG\Gamma_L G^{\dagger} f_L + iG\Gamma_R G^{\dagger} f_R. \tag{2.11}$$

The density matrix is then related to $G^{<}(E)$ by

$$\rho = \frac{1}{2\pi i} \int_{-\infty}^{\infty} G^{<}(E) dE. \tag{2.12}$$

Analogous to H we can split $G^{<}(E)$ into its spin components

$$G^{<}(E) = G_0^{<}(E) \cdot \mathbb{I} + \vec{G}_{\mathrm{S}}^{<}(E) \cdot \vec{\sigma}, \tag{2.13}$$

with $\vec{G}_{\mathrm{S}}^{<}(E) = \big(G_x^{<}(E), G_y^{<}(E), G_z^{<}(E)\big)$.

charge and spin current

We can evaluate the charge and spin current resolved between any two sites (orbitals) i and j in the system. This is denoted as the bond current J_{ij} [16], and can be separated in its spin components $J_{ij}(E) = J_{e,ij}(E) \cdot \mathbb{I} + \vec{J}_{\mathrm{S},ij}(E) \cdot \vec{\sigma}$, with $\vec{J}_{\mathrm{S},ij}(E) = (J_{x,ij}(E), J_{y,ij}(E), J_{z,ij}(E))$, and the energy dependent electron current $J_{e,ij}(E)$. Within the NEGF formalism, and for

systems with general overlap matrix S (of dimension $N_\mathrm{o} \times N_\mathrm{o}$), the spin-dependent bond current is given by

$$\begin{aligned} J_{e,ij}(E) = \frac{4e}{h}\mathbf{Re}\Big[&(H_{0,ij} - ES_{ij})G^{<}_{0,ji}(E) + H_{x,ij}G^{<}_{x,ji}(E) \\ &+H_{y,ij}G^{<}_{y,ji}(E) + H_{z,ij}G^{<}_{z,ji}(E)\Big] \\ J_{\alpha,ij}(E) = \frac{4e}{h}\mathbf{Re}\Big[&H_{\alpha,ij}G^{<}_{0,ji}(E) + (H_{0,ij} - ES_{ij})\,G^{<}_{\alpha,ji}(E)\Big], \end{aligned} \tag{2.14}$$

with $\alpha \in \{x, y, z\}$, and $\mathbf{Re}[]$ denotes the real part. The total charge/spin current is obtained by integrating over all energies, and results to

$$\begin{aligned} I_{e,ij} &= -\frac{4e}{\hbar}\mathbf{Im}\left[H_{0,ij}\rho_{0,ji} - S_{ij}F_{0,ji} + H_{x,ij}\rho_{x,ji} + H_{y,ij}\rho_{y,ji} + H_{z,ij}\rho_{z,ji}\right] \\ I_{\alpha,ij} &= -\frac{4e}{\hbar}\mathbf{Im}\left[H_{\alpha,ij}\rho_{0,ji} + (H_{0,ij}\rho_{\alpha,ji} - S_{ij}F_{\alpha,ji})\right], \end{aligned} \tag{2.15}$$

with $\alpha \in \{x, y, z\}$, and $\mathbf{Im}[]$ denotes the imaginary part. Here we have introduced the energy density matrix, F, which is given by [17]

$$F = \frac{1}{2\pi i}\int_{-\infty}^{\infty} EG^{<}dE \tag{2.16}$$

In order to obtain the current from a subsystem, denoted as SS_1, to another subsystem, denoted as SS_2, one needs to sum over all the possible site-to-site currents

$$I_{\alpha,SS_1,SS_2} = \sum_{i\in SS_1}\sum_{j\in SS_2} I_{\alpha,ij}, \tag{2.17}$$

with $\alpha \in \{e, x, y, z\}$. SS_1 includes all sites/orbitals within subsystem SS_1, while SS_2 includes all sites/orbitals within subsystem SS_2.

2.2.3 From Bandstructure to I-V: First-Principle TMR and STT

Evaluating TMR is rather straightforward once we can calculate the total charge current flowing through the MTJ stack. Since charge current is conserved throughout the junction, the total charge current is equal to the charge current across any transverse plane within the MTJ (e.g. choosing SS_1 to include all sites to the left of the chosen plane while SS_2

to include all sites to the right of the plane). Alternatively, one can directly calculate the terminal current from NEGF [15]. TMR can then be evaluated by setting the MTJ Hamiltonian to parallel and anti-parallel spin configurations and obtain the two currents $I_{p,\,ap}$ respectively. For Spin transfer torque, there are multiple ways to calculate the torque exerted on one or an ensemble of magnetic moments [18]. Here we adopt the approach from [19] but implement it on a non-orthogonal atomic basis. Based on [19] and the original definition from Slonczewski's [20], we can evaluate the total STT on any sub-system by evaluating the difference of the incoming and outgoing spin current to/from this sub-system, which is based on the conservation of the total spin angular momentum. Fig. 2.5 is a schematics to show how the STT can be calculated from the spin current across the MTJ. We have worked with the *Smeagol* [21] group to implement this approach in their code. The code can be obtained for free by contacting the *Smeagol* developers.

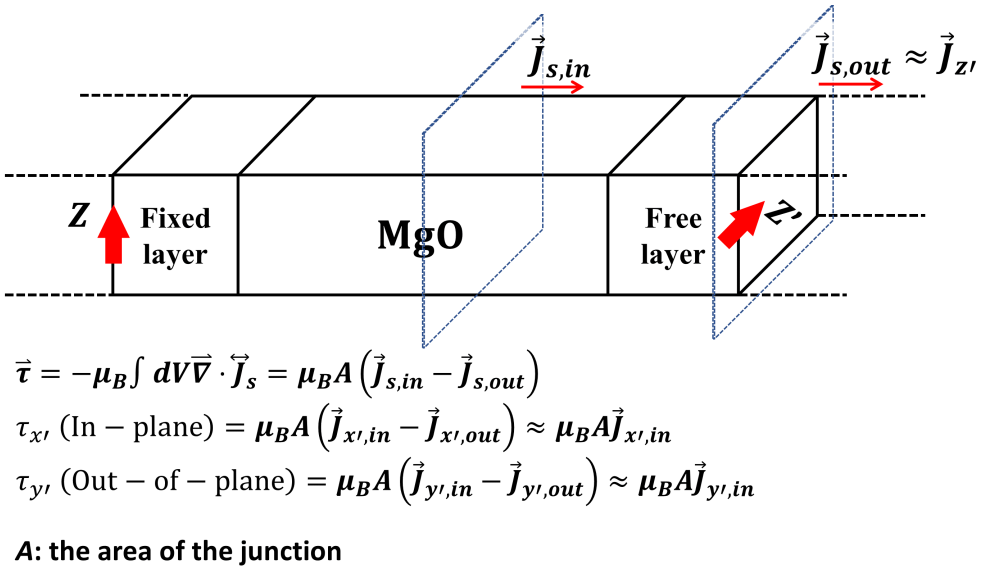


Figure 2.5: Evaluation of the STT at the right-hand-side electrode through spin-current fluxes. The magnetic moment of the fixed layer is pinned along z axis. The magnetic moment of the free layer is assumed to rotate within the x-z plane (transverse) plane.

First-Principle STT in Fe/MgO/Fe

We evaluate the STT for a defect-free Fe/MgO/Fe tunnel junction with 6 MgO monolayers as illustrated in Fig.2.6(a). In Fig. 2.6(b), the layer resolved spin-current is shown for different layers in our simulation cell. The bias voltage corresponds to 0.5 V, and the magnetic

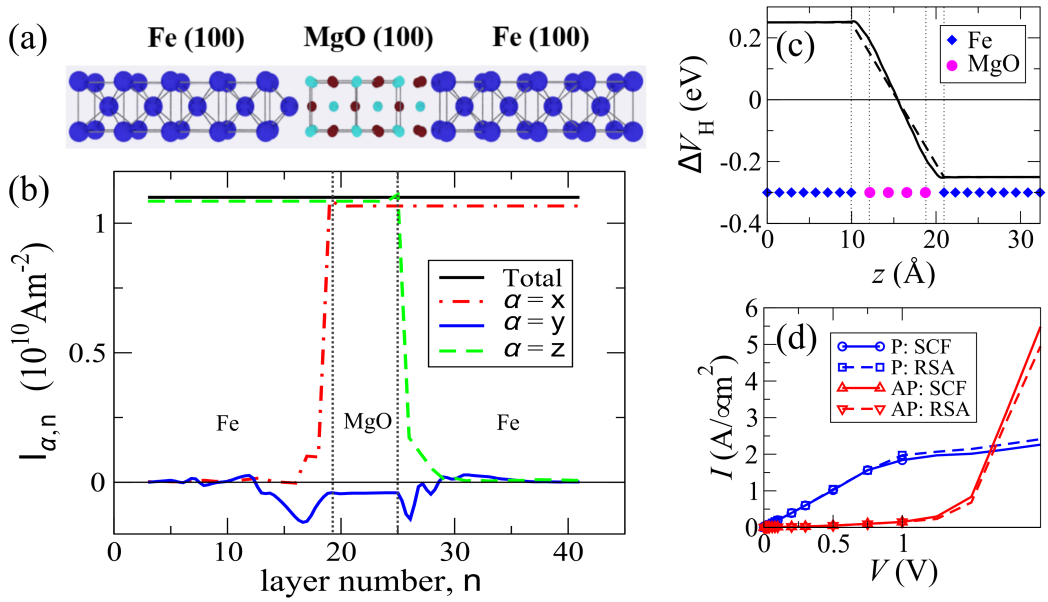


Figure 2.6: (a) Fe/MgO/Fe magnetic tunnel junction with 6 atomic layers of MgO. (b) Layer-resolved spin current at 0.5 V for a Fe/MgO/Fe junction, with the layer index denoted by n ($n \in [1, 19]$ correspond to the Fe layers in the left electrode, $n \in [20, 25]$ are the MgO monolayers, and $n \in [26, 41]$ correspond to the Fe layers in the right electrode). (c) Planar-averaged potential V_H profile at 0.5 V in the self-consistent calculation (solid line) and the rigid shift approximation (dashed line). Diamonds and dots indicate the location of the Fe and MgO layers, respectively.(d) Comparison of the I-V curves for the P and AP configurations obtained with the self-consistent solution (solid lines) and with the rigid shift approximation (dashed lines).

moment of the left electrode is along the z direction, while the one for the right electrode is along the x direction. We see that the current is fully polarized along z (x) deep in the left (right) electrode, showing that the local spin-current is parallel to the magnetization in the electrode. A torque is exerted at the layer index n where $I_{\alpha,n}$ changes significantly. We see that the torque is localized mainly at the Fe/MgO interfaces. For example, $I_{z,n}$ drops from its left-electrode bulk value to approximately zero within the first 4 Fe layers of the right electrode, which implies that the total torque along z (the in-plane torque for the right electrode) acts within these first 4 Fe layers. The out-of plane torque is determined by the y components of the spin-current, and acts also mostly close to the MgO/Fe interfaces, although it protrudes deeper into the electrodes. The total electron current is constant over the whole system, as expected. we consider the total torque acting on the right electrode, which is simply given by the spin-currents along z (y) in the middle of the MgO spacer for

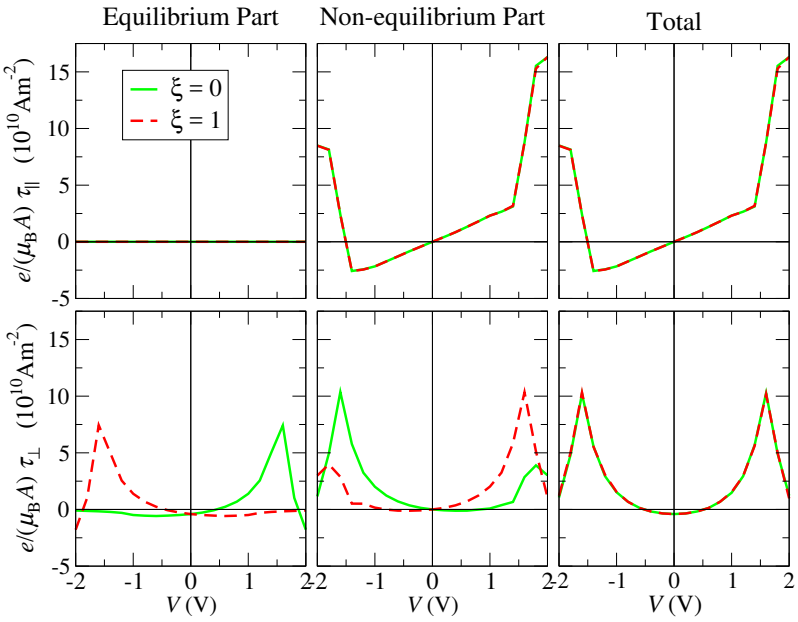


Figure 2.7: Bias dependence of the 'equilibrium' and 'non-equilibrium' components (defined in the text) of the in- and out-of-plane components of STT for a Fe/MgO/Fe junction with 6 MgO monolayers, with ξ defined in Eqs. (2.18) and (2.19).

the in-plane (out-of-plane) torque, since the total spin-current deep in the right electrode is fully polarized along the x direction (see Fig.2.5).

It is worth mentioning that we applied the bias voltage non-selfconsistently within the rigid shift approximation (RSA) for our calculations. We assume the potential drops linearly across the insulating barrier as shown in Fig. 2.6(c). We have also calculated it self-consistently to check the validity of the approximation. It gives a good approximation for the current when compared to the fully self-consistent solution (see Fig. 2.6(c)(d)). The RSA approximation dramatically reduces the computation time.

Fig.2.7 shows the bias dependent STT. As low bias, current driven STT (in-plane) is linear in V while the field like STT (out-of-plane) is quadratic in V, with a small non-zero component even at equilibrium. This small $V = 0$ out-of-plane torque is due to the exchange interaction between the left and right Fe electrodes across the MgO. At high bias above $\sim$ 1.5 V the torque shows a highly non-monotonic behavior and can even change sign. This is due to the fact that at these high voltages the current in the anti-parallel configuration increases rapidly and eventually becomes larger than the one for parallel alignment.

Out-of-Plane Torque: Equilibrium vs. Non-equilibrium

In the literature, the total electron density is often split into a 'condensate' or 'equilibrium part' (EP), and a 'non-equilibrium' part (NEP) [21]. This division comes from the NEGF formalism and usually doesn't have any effect when calculating any non-equilibrium properties such as I-V. However, as we have seen that the out-of-plane STT has a non-zero 'equilibrium' component, this division of EP versus NEP has caused some confusion in the literature, because some formalism only calculate the 'non-equilibrium' component of the STT. Therefore, the discrepancy in the definition can produce inconsitent results. Here we try to clarify this division with a definition that is consistent and physical.

Within NEGF the EP is usually written as

$$\rho_{\mathrm{EP}} = \frac{1}{2\pi i}\int_{-\infty}^{\infty}(-1)\left(G(E)-G^{\dagger}(E)\right)\left[\xi f_{\mathrm{R}}+(1-\xi)f_{\mathrm{L}}\right]dE, \tag{2.18}$$

and the NEP as

$$\rho_{\mathrm{NEP}} = \frac{1}{2\pi i}\int_{-\infty}^{\infty} iG(E)\left[\xi\Gamma_{\mathrm{L}}-(1-\xi)\Gamma_{\mathrm{R}}\right]G^{\dagger}(E)\left(f_{\mathrm{L}}-f_{\mathrm{R}}\right)dE, \tag{2.19}$$

with $\xi \in [0,1]$, so that $\rho = \rho_{\mathrm{EP}} + \rho_{\mathrm{NEP}}$. Such a partitioning is not physically motivated and the choice of the terms 'equilibrium' and 'nonequilibrium' is somewhat of a misnomer. The true 'equilibrium part' would correspond to $f_{\mathrm{L}} = f_{\mathrm{R}}$ and the partitioning thereafter would not be arbitrary, provided the voltage division among the contacts is consistent with the intermediate Laplace potential matching all boundary conditions. The drive to make the EP and NEP division is primarily computational, since the so-called EP part in Eq. 2.18 involves just the imaginary part of G, whose poles are localized entirely in one half of the complex energy plane. The simple pole structure makes the corresponding EP integral easy to evaluate on a contour, leaving just an energy window over which the residual NEP integral needs to be computed brute force. Since ξ can be chosen arbitarily in the range from 0 to 1, the splitting in EP and NEP is not unique. In the same way the energy density is split into EP and NEP. The EP of the torque is then obtained by using ρ_{EP} and F_{EP} in Eq. (2.15), while the NEP torque is obtained by using ρ_{NEP} and F_{NEP}, so that the total

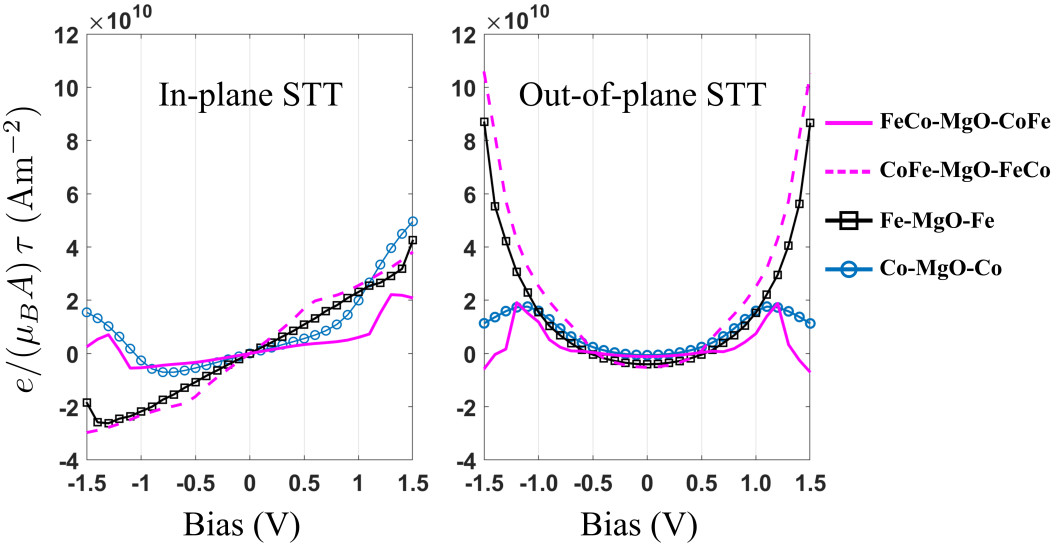


Figure 2.8: Comparison of in-plane and out-of-plane STT for different electrode compositions and interfaces.

torque is the sum of EP and NEP torques. The results are shown in Fig. 2.7 for $\xi = 0$ and $\xi = 1$. While the in-plane torque is identical for any choice of ξ, it can be seen that the individual out-of-plane components change completely depending on the choice of ξ. Importantly, the total out-of-plane torque is independent of the choice of ξ, indicating that *the only meaningful quantity is the total torque*, and it is not really meaningful to split it into the arbitrary EP and NEP parts.

STT in Different Electrode/Interface

While Fe/MgO/Fe is the benchmark system for MTJ calculations. We look further into the dependence of the torque on the electrode composition. In most recent experiments, typically a mixture of Co and Fe is used as electrodes. We consider four different systems here: 1) Fe-MgO-Fe is the system considered so far, 2) Co-MgO-Co is the system where we replace the Fe atoms with Co atoms, 3) CoFe-MgO-FeCo is a system where we put alternating layers of Fe and Co, and where Fe atoms are placed at the interface to MgO on both sides, 4) FeCo-MgO-CoFe is the same as system 3), but where we put Co atoms at both interfaces to MgO. The resulting STT data are shown in Fig. 2.8. The general trends are similar for all junctions, but there are quantitative differences in the magnitude of the

STT. For the Co-MgO-Co junction, we find the onset of the non-linear behavior for the in-plane torque already at about 1 V. The early onset of nonlinearity arises from the fact that Co has one more electron than Fe so that the Fermi energy is effectively shifted to higher energies. For the mixed systems, we see that the metal layer adjacent to the MgO is of key importance: CoFe-MgO-FeCo shows the highest low-bias torque, while FeCo-MgO-CoFe shows the smallest one. For randomly mixed FeCo systems, we expect that the overall torque corresponds to some average of the shown results, although clearly, the local torque will still be strongly dependent on the vicinal atomic structure. A comparison between our calculation and an experimental measurement [22] shows that the measured STT lies somewhere between system Fe-MgO-Fe and CoFe-MgO-FeCo. Although this is a difficult experiment and there is still debate on the precision of the measurements, DFT-based calculations are the only practical means to evaluate the material and bias-dependent variations in the STT for realistic interfaces. The microscopic insights and physical understanding we drawn from these simulations are of clear significance to the development of STT-MRAM technology.

Half-Metal Magnetic Tunnel Junctions

High-throughput material genome projects have provided us with a family of potential magnetic materials for spintronics applications. For exploratory analysis, we have selected six Half Heuslers (HH) that are half-metallic or near half-metallic. They are selected with minimum lattice mismatch ($< \%1$) with MgO[100] to reduce interfacial strain. There are more than one way for half-heusler alloys (with chemical formula XYZ) to form MTJ with the MgO (e.g. interfacial configuration -YY-MgO-YY-, or -XY-MgO-XY-, and etc.). In the absence of experimental evidence, my colleague Jianhua has studied the energetics of different interface type and determined a particular configuration -YY-MgO-YY- has the lowest total energy compared to other configurations [23]. Fig. 2.9(i)(ii) show the I-V and TMR of the six HH-MgO-HH junctions with the -YY-MgO-YY- configuration. Given that those calculations assume defect-free interfaces, four of them with the top TMRs can be considered to have good readout capabilities. For comparison, calculations of Fe/MgO/Fe MTJs have shown TMRs well above 6000% while experiments in expitaxially-grown Fe-MgO and CoFeB-MgO MTJs have demonstrated about 600% TMR at room temperature and

1140% at 5K [24–27]. For the write part, most of these alloys have relatively small current except for NiMnSb, which also has the highest TMR. Out of those six alloys, it seems that NiMnSb with -YY-MgO-YY configuration has the most potential to build a high performance MTJ. Fig. 2.10 summarizes the transport properties of NiMnSb/MgO/NiMnSb junction including the TMR (read) and STT (write).

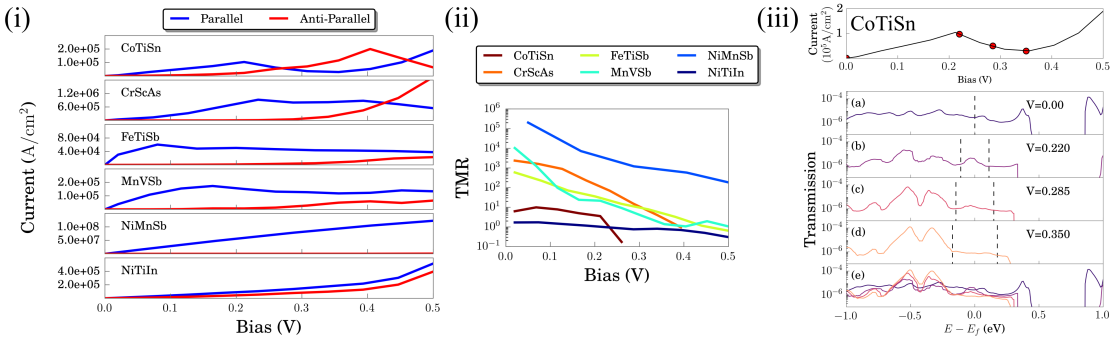


Figure 2.9: (i) I-V of Half-heusler/MgO/Half-heusler junction for P and AP configuration. (ii) Bias dependent TMRs for the six MTJs. (iii) Top. I-V of CoTiSn/MgO/CoTiSn in parallel configuration. (a)-(e). Transmission as a function of energy at different voltage. Bias points indicated by red circle in the top IV plot. (e) is (a)-(d) put together for comparison.

Some I-Vs show negative differential resistance (NDR) effects in our calculations. After careful examination, we argue that it is mostly due to a peak in the transmission that is shrinking with bias (because the left and right interface states getting out of resonance). Considering all six systems under study are half-metallic or near half-metallic, there is a gap in at least one spin channel very close to the Fermi energy. Therefore, in some bias range the widened bias window doesn't add many more conducting channels while the out-of-sync resonance state reduces the conductance of existing channels as shown for CoTiSn in Fig. 2.9(iii). In real experiments, it is hard for the resonance states to exist due to defects at the interface and therefore the NDR effect is usually washed out.

2.3 Write Error Rate in MTJ

Once we can calculate the spin transfer torque, we are able to study the process of magnetization switching. Here we move away from the atomistic details of the MTJ and treat the free magnet, which consists of a large number of spins, as one giant spin (macrospin

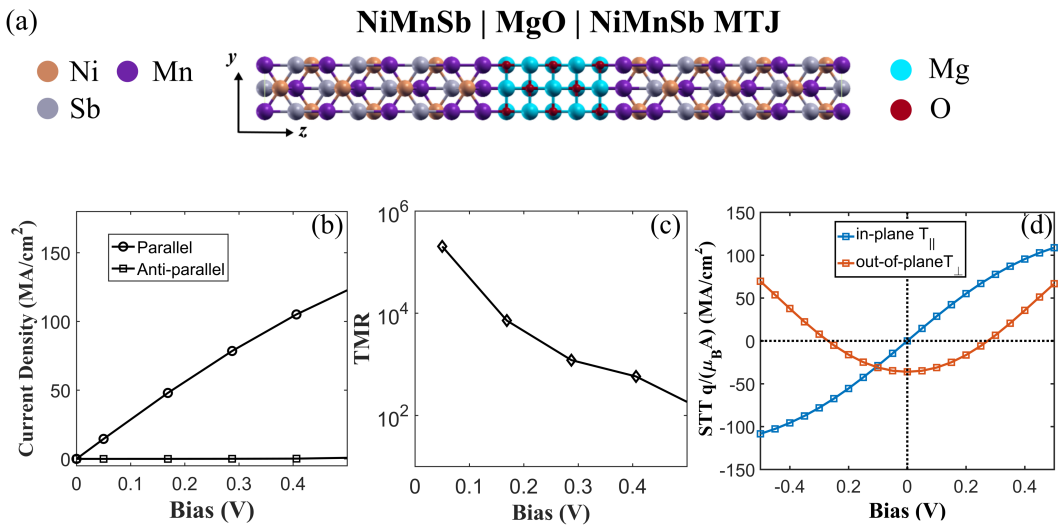


Figure 2.10: Transport calculation for a NiMnSb-MgO-NiMnSb (100) junction. (a) Unit cell used for NiMnSb-MgO-NiMnSb heterojunction with 5 monolayers of MgO. (b) The low bias I-V characteristics. (c) Low bias TMR ratio calculated from the I-V: TMR $= (I_p - I_{ap})/I_{ap}$. (d) The bias-dependent spin transfer torque of the junction.

approximation). This approximation is generally valid for sub-100nm magnets, where all spins act in a coordinated fashion and thus can be seen as one. With this approximation, useful entities have been derived to characterize the performance of a nanomagnet in the write process [28]. Eq. 2.20 summarizes two important quantities: zero temperature critical switching current I_c and thermal stability factor Δ. The former dictates the minimal current required to flip the magnet while the latter describes the stability of the magnet against thermal noise. k_B is the Boltzmann constant and T is the temperature. M_s is the saturation magnetization. μ_0 is the permeability constant and Ω is the volume of the magnet. q is the single electron charge. α is the magnetic damping. η is the spin polarization of the injecting current. H_k is the anisotropy field. K_u is the uniaxial magnetic anisotropy.

$$\Delta = \frac{\mu_0 H_k M_s \Omega}{2k_B T}, \quad I_c = \frac{2\alpha q}{\eta \hbar} \mu_0 H_k M_s \Omega, \quad H_k = \frac{2K_u}{\mu_0 M_s} \tag{2.20}$$

The reliability of a nanomagnet consists of two parts: Firstly, the magnet should not switch itself when switching is not intended (e.g. in the idle state or in a read operation). Secondly, the magnet should switch reliably in a write operation. One ubiquitous source

of error comes from thermal fluctuation in the magnetic moment. In the first scenario, if the energy barrier between the parallel (P) and antiparallel (AP) states is small, thermal fluctuation can cause accidental flip of magnet known as thermal switching. It implies that the thermal stability factor Δ needs to be high enough (e.g. $\Delta > 40$ for approximately 10 years of lifetime). The write process is more complicated because thermal fluctuation actually helps initiate the switching process but the effect is random due to its stochasticity, therefore creating a distribution of switching delays in repeated switching events. A reliable system requires all magnetic cells to be able to switch within a given time budget. Extra current $I > I_c$ is usually needed to write reliably. It indicates that we need more than one number (I_c) to describe the reliability of magnetization switching under a certain time constraint.

A more useful figure is the write error rate (WER) as a function of voltage/current for a given current pulse width. Such a plot contains multiple messages: First, the plot shows the onset of switching which can indicate the average switching current. The plot also sets the approximate boundary for read current since a switching event in read operation causes read disturbance. Second, the lowest error rate achieved at certain current limit will determine the size of the memory array. Third, when the error rate is plotted in logarithmic scale, the slope of the 'error tail' shows how fast the error rate goes down with increasing current, which characterizes the error margin for the write operation. The following work on WER analysis were published in ref. [1] coauthored by Behtash Behin-Aein, Avik W Ghosh and ref. [29] coauthored by Jianhua Ma, Samiran Ganguly, Avik W Ghosh with only my contributions included in the dissertation.

2.3.1 Switching Regions and Analytical Models

In the literature, discussions on thermal effects in spin torque switching usually fall into two switching regions set by the ratio between injected current I and critical current I_c. The supercritical $I \gg I_c$ regime is called current dominated region while the subcritical regime $I \ll I_c$ is referred to as thermal region. This division allows approximate analytical models to be built in the two limits. However, practical MTJ applications mostly work in the intermediate regime at the moment because of three reasons - large MTJ resistance,

relatively large critical switching current, and need to avoid time-dependent dielectric breakdown (TDDB). Most analytical models do not have a 'smooth' transition between supercritical/subcritical regimes and often encounter mathematical singularities in the transition. Recently there have been efforts at formulating a brute force mathematical transition between the analytical equations [30]. Such a scheme offers a simple fix to the discontinuity of the analytical equations but lacks physical insights into the switching behavior at transition. The alternative is to use numerical solution to avoid mathematical approximations and take into account physical parameters at the same time. Numerical methods are quite universal and not limited to a specific region but are in general much less computationally efficient than analytical approaches. Among those numerical methods, we will show the advantages of the Fokker-Planck Equation (FPE) approach. FPE can be applied to all switching regions as demonstrated in Fig. 2.11 where the average switching current (measured by P_{sw} = WER = 0.5) is plotted against the switching time. The experimental data are fitted with different analytical models as well as the FPE simulation.

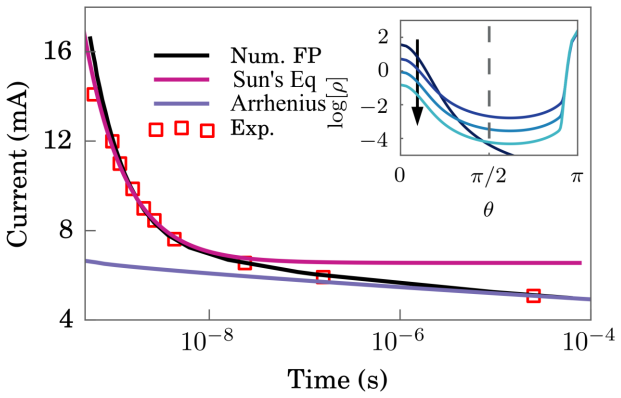


Figure 2.11: Average switching current ($P_{sw} = 0.5$) as a function of pulse width in a 100 nm spin-valve nano-pillar. The experimental data are extracted from Fig. 7(a) of ref [2]. Inset: Time-evolution of probability distribution at $t = 0.1, 5.0, 10, 15$ ns with current $I = 6.55$ mA. The arrow indicates increasing time. For analytical models, the supercritical regime is fitted with the Sun's equation 2.21 while the subcritical regime is fitted with a modified Arrhenius model [3]. The FPE parameters are $\Delta = 80$, $I_c = 8.3$, mA, $\tau_D = 0.25$ ns. See the definition of τ_D in Eq. 2.22. Figure reprinted from [1] with permission @2017 IEEE.

It is worth reviewing one analytical model - Sun's equation 2.21, which assumes a STT-dominated magnetic switching event. Although it can be less accurate when $I <= I_c$,

it is a very useful equation from which one can gain quantitative insights from, which we will revisit the end of this chapter.

$$P_{sw} = \exp\left\{-4\Delta \exp\left[-2\tau(i-1)\right]\right\} \tag{2.21}$$

where i, τ, Δ are scaled quantities defined as:

$$i = \frac{I}{I_c}, \qquad \tau = \frac{t}{\tau_D}, \qquad \tau_D = \frac{1+\alpha^2}{\alpha\gamma\mu_0 H_k} \tag{2.22}$$

where γ is the gyromagnetic ratio. Note that all models discussed here are based on the macrospin approximation. In reality STT switching can involve complications like sub-volume effects [28] or edge effects [31]. While it is crucial to understand those effects, accounting for both non-macrospin effects and thermal effects can be computationally challenging in simulations. Thus macrospin model is still valuable because one can typically approximate those complicated effects with effective parameters that are good enough to draw physical insights and at the same time interface with practical device or circuit simulations.

2.3.2 Fokker-Planck Approach

The phenomenological Landau-Lifshitz-Gilbert (LLG) equation describes the dynamics of the normalized magnetization $\mathbf{m} = \mathbf{M}/M_s$ determined by the torque from effective magnetic field $\mathbf{H}_{\text{eff}}$, magnetic damping, and Slonczewski spin torque $\mathbf{L}_{\text{STT}}$ (the in-plane component) in the case of STT switching:

$$\begin{aligned}
\frac{\partial \mathbf{m}}{\partial t} &= \mathbf{L}_{\text{prec}} + \mathbf{L}_{\text{damp}} + \mathbf{L}_{\text{STT}} \\
\mathbf{L}_{\text{prec}} &= -\mu_0\gamma\mathbf{m} \times \mathbf{H}_{\text{eff}}/(1+\alpha^2) \\
\mathbf{L}_{\text{damp}} &= -\alpha\mu_0\gamma\mathbf{m} \times (\mathbf{m} \times \mathbf{H}_{\text{eff}})/(1+\alpha^2) \\
\mathbf{L}_{\text{STT}} &= -\frac{\mu_B I\eta}{q\Omega M_s}\mathbf{m} \times (\mathbf{m} \times \mathbf{I}_s)/(1+\alpha^2)
\end{aligned} \tag{2.23}$$

μ_B is the Bohr magneton. I is the charge current and $\mathbf{I}_s$ is the unit vector along the injected spin orientation. To include the thermal effect, $\mathbf{H}_{\text{eff}}$ is replaced by $\mathbf{H}_{\text{eff}} + \mathbf{F}_{\text{th}}$ where $\mathbf{F}_{\text{th}}$ is a

random torque from the thermal noise.

An alternative way to quantify the statistical nature of STT switching under thermal fluctuations is to solve the corresponding Fokker-Planck equation. The Fokker-Planck method has been applied to describe thermally agitated magnet by Brown [32]. The method can be generalized to include spin transfer torque. We start from the LLG equation above and assume a gaussian distribution of random thermal noise: $\mathbf{F}_{\text{th}}$

$$\Pi\left(\mathbf{F}_{\text{th}}\right) = \frac{1}{\sqrt{8\pi^3 D^3}} \exp\left(\frac{-\left|\mathbf{F}_{\text{th}}\right|^2}{2D}\right) \tag{2.24}$$

Instead of keeping track of the random trajectory of $\mathbf{m}$, the Fokker-Plank equation solves for the probability distribution of magnetization:

$$\rho\left(\mathbf{m};t\right) = \int \Pi\left(\mathbf{F}_{\text{th}}\right) \delta\left(\mathbf{m} - \mathbf{m}_{\mathbf{F}}\right) d\mathbf{F}_{\text{th}} \tag{2.25}$$

Taking the derivative of Eq.2.25 with respect to time gives the general Fokker-Planck equation for nano-magnet in the form of a convection-diffusion equation on a 2D spherical surface:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\mathbf{L}\rho) + D\nabla^2 \rho \tag{2.26}$$

where $\rho\left(\theta, \phi; t\right)$ is the probability density of the magnetization in spherical coordinate. $\mathbf{L}\left(\mathbf{m}\right) = \mathbf{L}_{\text{prec}} + \mathbf{L}_{\text{damp}} + \mathbf{L}_{\text{STT}}$ includes all the deterministic torque terms in Eq. 2.23. The effective 'diffusion' constant describing the thermal effect is defined as:

$$D = \frac{\alpha\gamma k_B T}{(1+\alpha^2)\mu_0 M_s \Omega} \tag{2.27}$$

We solve Eq. 2.26 through finite-element method with triangular meshes generated on a unit spherical surface [33] [34]. The differential operator is discretized with Galerkin's method and the time evolution is calculated through Crank-Nicolson's method. The solution is a 2-dimensional probability density $\rho(\theta, \phi; t)$ evolving on the surface of a unit sphere. The WER (or non-switching probability P_{ns}) is then evaluated as the total integrated probability lingering on the upper hemisphere (assuming the initial magnetization points at

+z direction):

$$\mathrm{WER} = \int_0^{\pi/2} \int_0^{2\pi} \rho\left(\theta, \phi; t\right) d\phi d\theta \tag{2.28}$$

A Matlab version of the code can be found in the appendixB for benchmark and further developement.

Computational Efficiency

To accurately describe switching distribution, Eq. 2.23 needs to be solved for a large number of trials which is computationally expensive. Fokker-Planck equation only needs to be solved once. Even though a single run of FPE takes longer than a single run of LLG, FPE simulation shows much higher computational efficiency and accuracy than LLG approach at low WER. Fig. 2.12 a comparison between the stochastic LLG approach and the 2D Fokker-Planck approach in terms of accuracy and running time. Capturing low probability events with high accuracy in LLG method requires enormous number of runs. In comparison, FPE approach can achieve high accuracy with reasonable number of grid meshes. In a magnetic system with rotational symmetry, Eq. 2.26 can be easily reduced to a 1-D differential equation [35] which can be solved even more efficiently.

2.3.3 WER in Perpendicular STT-MRAM

In general, WER has a complex dependence on current, pulse width and material properties. Current and pulse width are often determined by the application. Therefore we will focus our discussion on a few other related parameters: temperature, anisotropy and damping, saturation magnetization, current polarization. Some parameters are prone to change and hard to control (e.g. temperature rise due to joule heating) while others are often tuned by experimentalists to achieve better performance. We hope to shed some light on how those parameters affect the WER. For the convenience of discussion, we have chosen a set of experimental data from [36] and constructed a 'reference' device - a perpendicular CoFeB magnetic tunnel junction. The physical parameters are shown in table 2.1. All the following discussions assume a fixed pulse width of $\tau_{pw} = 10\,\mathrm{ns}$. Before going into the discussion, it

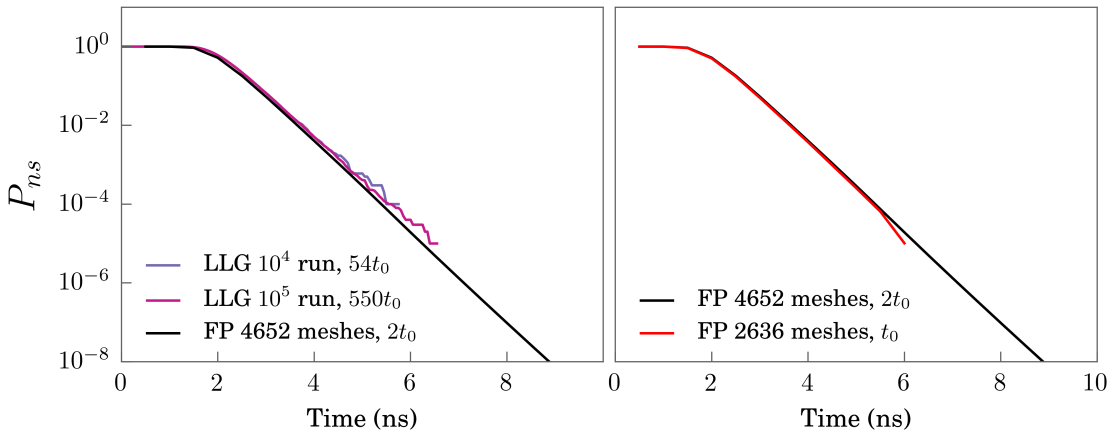


Figure 2.12: Non-switching probability P_{ns} as a function of switching time for $I = 6I_c$ where I_c is the critical current. Left. Compare Fokker-Planck result and LLG simulations with different number of trials. Right. FP simulations with different number of grid meshes. All simulation times are normalized by the shortest one t_0 for FP simulation with 2636 meshes. The initial magnetization distribution of our LLG simulations is sampled from a Boltzmann distribution. A very fine sampling grid is needed to capture the initial magnetization close to the easy axis. Figure reprinted from [1] with permission @2017 IEEE.

is worth mentioning that in general, specific WER targets need to be met by the memory array or various applications to build a viable product. The obvious way to achieve low WER is to increase the voltage but that comes with a power consumption penalty and more importantly the probability of dielectric breakdown over time. Switching efficiency is another aspect that can help but so far the highest efficiency is achieved in MTJs with smaller diameter [37] which is accompanied by very high resistance. In this section, we focus on the effect of material parameters on the write error and write slope. The actual WER engineering requires diligent handing of all such factors. Nevertheless the approach is quite general and can be easily applied to other discussions.

Let us now discuss the parameter dependences on the MTJ WERs. In the following WER-V plots, the voltages have been scaled by the average switching voltage $V_{sw}^{ref}(\text{WER} = 0.5)$ of the reference device to de-emphasize their exact values but to focus on the general trend.

Table 2.1: MTJ Parameters

parameter	value	remark
d	40 nm	free layer diameter
t	1 nm	free layer thickness
M_s	1.23×10^6A/m	saturation magnetization
α	0.027	magnetic damping
Δ	43	thermal stability
η	0.6	spin polarization
RA	$18\ \Omega \cdot \mu\text{m}^2$	resistance area product

Table 2.2: Table reused from [1] with permission @2017 IEEE.

Temperature

Practical applications of STT are inevitably tied to stochastic switching at room temperature. Indeed, STTs are prone to environmental change and joule heating during the writing process. The impact of thermal fluctuations on WERs is critical to the operation of STT based devices, the main effects being a change in the initial magnetization distribution and a change in the energy barrier $\Delta = E_b/k_BT$. Fig. 2.13 shows that the impact on WER of a drastic change in temperature is quite minimal. More complicated current driven temperature change could arise due to self-heating effects that are beyond the scope of this dissertation [38]. The weak dependence of WERs on temperature variation can be understood by recognizing that these simulations are performed in the fast switching regime (10 ns), during which the chance of thermal switching is very low. Therefore temperature variation only affects the initial magnetization distribution. This effect can be estimated from the Sun's Equation 2.21. One can show that the temperature difference from 300K to 425K only changes log[WER] by about log[425K/300K] from the Taylor series expansion of Eq. 2.21 at WER $\rightarrow$ 0 (or $\text{P}_{\text{sw}} \rightarrow 1$) limit.

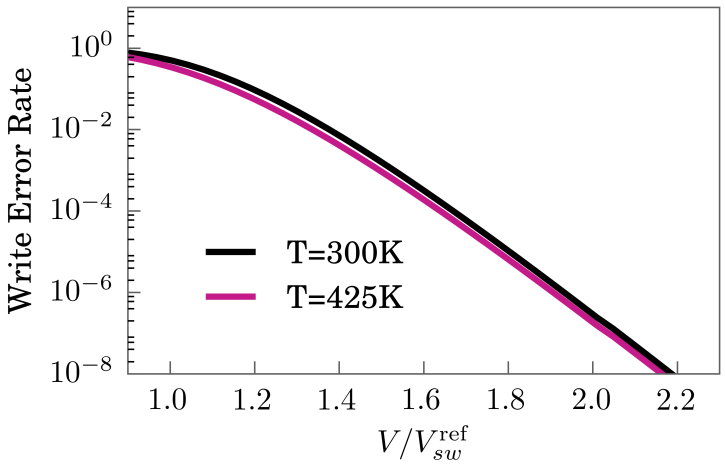


Figure 2.13: WER as a function of junction bias for different temperatures. Figure reprinted from [1] with permission @2017 IEEE.

Anisotropy, Damping and saturation magnetization

Besides the physical dimensions of the magnet, anisotropy, magnetic damping, and saturation magnetization are three important material parameters experimentalists often work with. Improvements in fabrication techniques as well as material modeling can give better control of these parameters over time. Fig. 2.14 shows that magnetic damping α and anisotropy field H_k have nearly identical effects on the WER, assuming M_s is held constant. Reducing either parameter would reduce the switching current. This is expected since these are the physical forces that oppose magnetic switching [39]. Fig. 2.14 also shows that reducing α or H_k does not change the slope of the WER-V curve, meaning the noise margin remains the same even when the average switching current is changed.

Fig. 2.15 shows how saturation magnetization M_s affects the WER. Since M_s and H_k are related through Eq. 2.20, two different scenarios emerge. In the left plot of Fig. 2.15, H_k is kept fixed. This means the overall magnetic anisotropy constant K_u changes with M_s (see Eq. 2.20), a reduction in M_s reduces the switching current and increases the error slope. In the right plot, K_u is kept unchanged while H_k changes with M_s accordingly. The average switching current is altered moderately while the slope of WER-V curve changes in a similar way as the left plot. It is important to notice that in the left scenario, the change

of saturation magnetization M_s would change the thermal stability Δ as well but in the right scenario (fixed K_u) Δ remains unchanged for different M_s according to Eq. 2.20.

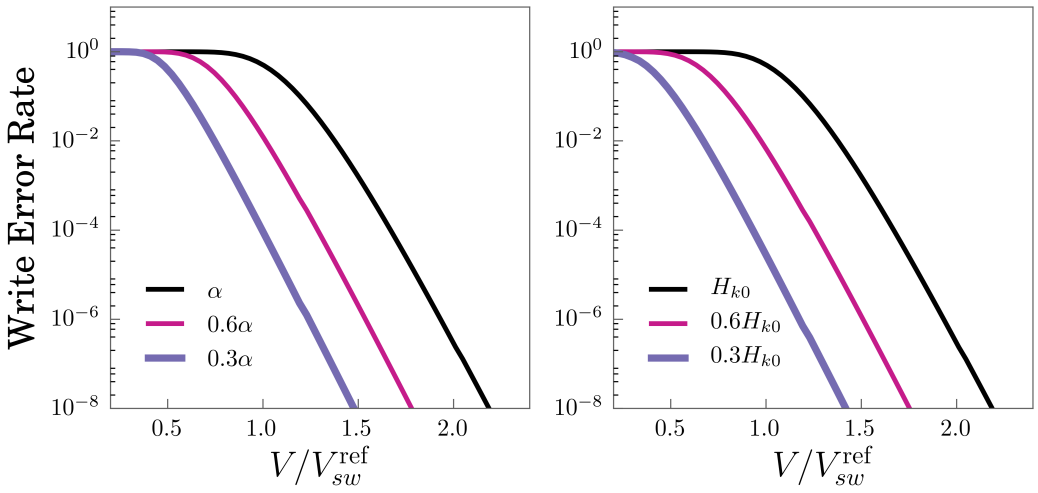


Figure 2.14: WER-V plot for various magnetic damping and anisotropy. Left: WER-V for different magnetic damping α. Right: WER-V for different magnetic anisotropy field H_k. Figure reprinted from [1] with permission @2017 IEEE.

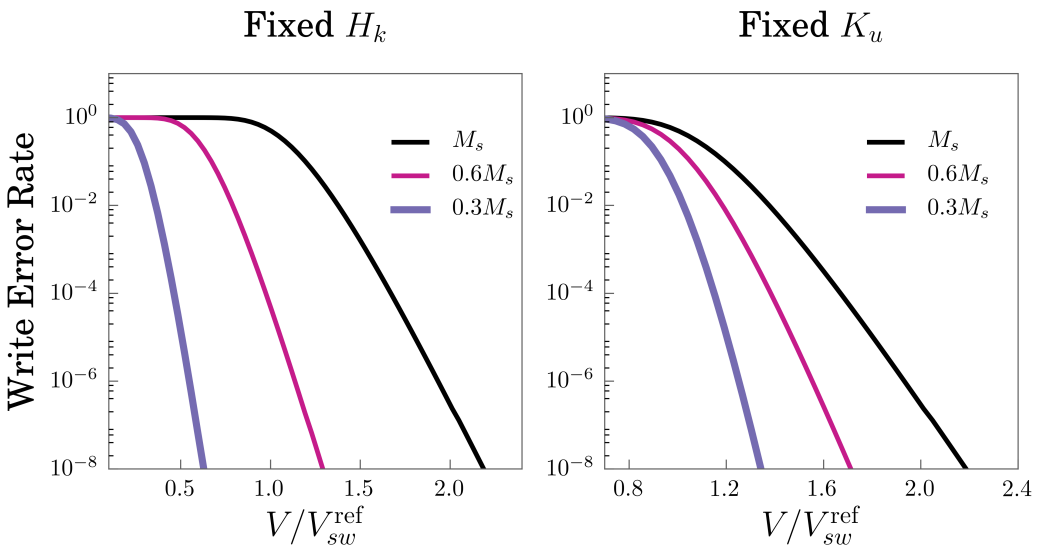


Figure 2.15: WER-V plot for various saturation magnetization. Left: WER-V with fixed H_k. Right: WER-V with fixed K_u. Figure reprinted from [1] with permission @2017 IEEE.

Spin Polarization

The degree of spin polarization is also a critical determinant of most spintronic applications, although affected by many factors and hard to control. In the case of MTJ, it is spin filtering by the fixed magnet that imposes a torque on a noncollinear free magnet. Typically the spin polarization is determined by the materials used but can be largely affected by the interfacial configurations such as symmetry filtering oxides as discussed in sec 2.1, or defects and strain [40] [41]. Fig. 2.16 plots the WER-V for various spin polarizations. Understandably, increased polarization lowers the critical current (from Eq. 2.22) as well as the average switching current (WER = 0.5). At the same time, the slope of WER-V increases with spin polarization, resulting in a narrower write margin. This improvement of performance saturates with $\eta \to 1$.

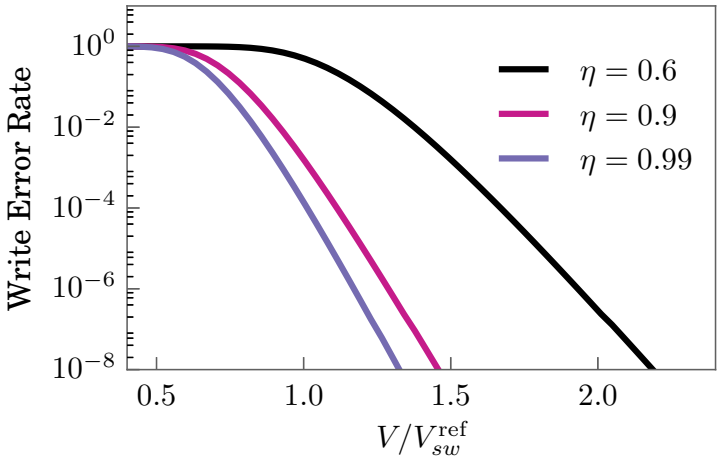


Figure 2.16: WER as a function of junction bias for various polarization. Higher polarization allows lower critical current and a sharper slope. Figure reprinted from [1] with permission @2017 IEEE.

In the real world, it is hard to change one parameter at a time. Indeed, many of the parameters discussed above are in fact, correlated. However, it is not difficult to comprehend their combined effect on the writing process. Notice that the average switching current I_{sw} is mostly determined by the intrinsic critical current I_{c0}. From Eq. 2.20, it is easy to see $\alpha, \eta, M_s H_k \propto K_u$ are the factors determining the critical current. On the other hand, The slope of WER-V curve is a function of voltage/current but its asymptotic value at small

WER can be approximated from Sun's equation (Eq. 2.21) in the limit WER$\rightarrow 0$ where we get

$$\mathcal{S} := -\frac{d\log[\text{WER}]}{dV} \approx \frac{\mu_B}{M_s\Omega}\frac{\eta}{1+\alpha^2}\frac{t}{qR} \tag{2.29}$$

with R the junction resistance. Table 2.3 shows that Eq. 2.29 can be used to approximate the WER slope and its dependence on the physical parameters which is summarized qualitatively in Fig. 2.17.

Table 2.3: WER slope for different parameters

	parameter value	decades per 100 mV		ratio of FPE to $\mathcal{S}_0$
		FPE*	eq. 2.29	
	reference	1.26	1.56 ($\mathcal{S}_0$)	0.99
Fig. 2.13	$T = 425$ K	1.26	1.56	0.99
Fig. 2.14. left	0.6α	1.33	1.56	1.05
Fig. 2.14. left	0.3α	1.33	1.56	1.05
Fig. 2.14. right	$0.6H_k$	1.29	1.56	1.02
Fig. 2.14. right	$0.3H_k$	1.32	1.56	1.05
Fig. 2.15	$0.6M_s$	2.09	2.61	1.66
Fig. 2.15	$0.3M_s$	4.10	5.22	3.25
Fig. 2.16	$\eta = 0.9$	1.92	2.35	1.52
Fig. 2.16	$\eta = 0.99$	2.10	2.58	1.67

*slope extracted from numerical FPE simulations.

Table 2.4: Table reprinted from [1] with permission @2017 IEEE.

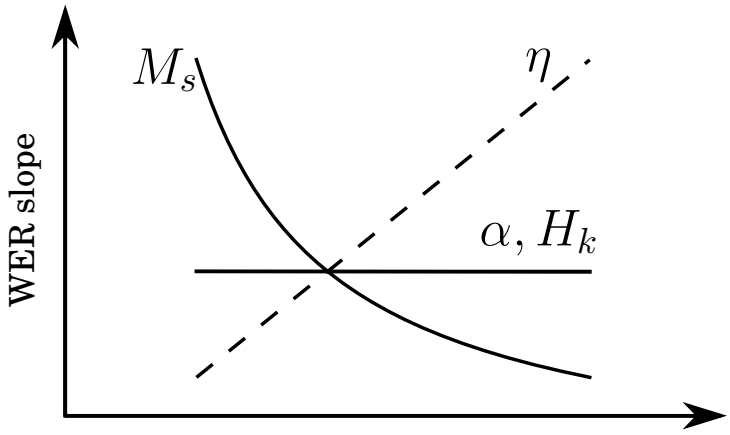


Figure 2.17: Schematic for the WER slope dependence on various physical parameters. Figure reprinted from [1] with permission @2017 IEEE.

It is instructive to look at the relation between the WER and the total charge flowing through the magnet, which can be calculated approximately from the Sun's equation 2.21 or Eq. 2.29. After simple algebra, we get the following equation

$$WER \propto e^{-2(Q - Q_C)/Q_C}, \quad Q_C = \frac{qM_S\Omega}{\mu_B}\left(\frac{1+\alpha^2}{\eta}\right) \tag{2.30}$$

which means that the efficiency of switching ultimately depends on the total accumulated charge Q, and we can ramp up the accuracy by overdriving with charge exceeding the minimum critical charge Q_C to destabilize the spins towards flipping. The critical charge is obtained from simple angular momentum conservation, trading off the spin $\mu_B Q_C/q$ with that from the magnets $M_S\Omega$, while accounting for partial polarization of the tunneling electrons (making them less efficient by a factor η), and the damping correction $1+\alpha^2$ implying that part of the injected spin leaks out into the environment (because damping is dissipative). For a magnet of size 100nm $\times$20nm with about 100,000 spins, we need about 10^6 electrons to provide the critical charge Q_C. For a 10 ns switching time, this already requires a current density of ~ 1 MA/cm^2, which more or less agrees with most experimental data.

2.3.4 Non-collinear Magnet for Fast Switching

As mentioned earlier, the thermal fluctuation can help with the writing process. The reason is that when the fixed magnet and the free magnet have strictly collinear magnetization (Such orientations are called 'stagnation points'), there is zero STT since the incoming electrons only deposit the angular momentum perpendicular to the magnetization of the free layer, proportional to $\vec{m} \times \vec{m}_s$. A strict parallel or anti-parallel configuration cannot induce a torque on the magnetization. The switching time is very long when the initial angle is small. Thermal noise can 'nudge' the magnetic moment of the free magnet out of its stagnation position and initiate the switching but it is a stochastic (unreliable) force. In the remainder of this chapter, we use the 2D FPE solver to investigate the effect of non-collinearity in STT switching.

Example 1 - non-collinear initial $\vec{m}$

In the first example, we take a closer look at how a non-collinear setup can help reduce the WER. We assume an extra force is used to nudge the initial magnetic moment out of its stagnation point to facilitate switching. There are multiple proposed schemes to achieve this step including an initial thermal torque from heating to excite the magnetization to a large angle [42], or use an in-plane spin-orbital torque arising from a Spin Hall effect to disturb the initial magnetization [43]. We will discuss the orthogonal torque in the next example. Here we focus on the second subsequent dynamics following the initial kick. Fig. 2.18 shows the WER plot of from a 'tilted' initial angle. The assumption is that the initial magnetization still obeys a Boltzmann distribution except the maximum probability density is inclined at an angle θ_0 with respect to the z axis (easy axis) shown in Fig. 2.18(b). Fig. 2.18(a) shows that the initial excitation (usually fast, $<$ 1ns, so any overhead delays can be neglected) helps reduce both the critical write current and write noise margin. Fig. 2.18(c) further illustrates the different types of switching behavior in a canted system versus a collinear system. Whereas in the collinear case, the initial magnetization distribution remains almost unchanged for some time and starts to 'diffuse' to the -z direction, in the canted case the magnetization starts to precess and move to the -z direction almost immediately after the

current is applied. Therefore the initial delay in the collinear case is avoided in a canted system.

We should caution that our assumption of the Boltzmann distribution of the tilted initial angle is likely over-simplified. Any non-symmetric force such as the stray field could break the symmetry and result in a non-Boltzmann distribution. In that case, a separate simulation can be performed to obtain the initial distribution before switching. Including different torques from external magnetic fields or spin-orbital couplings is quite straightforward in the general 2D FPE.

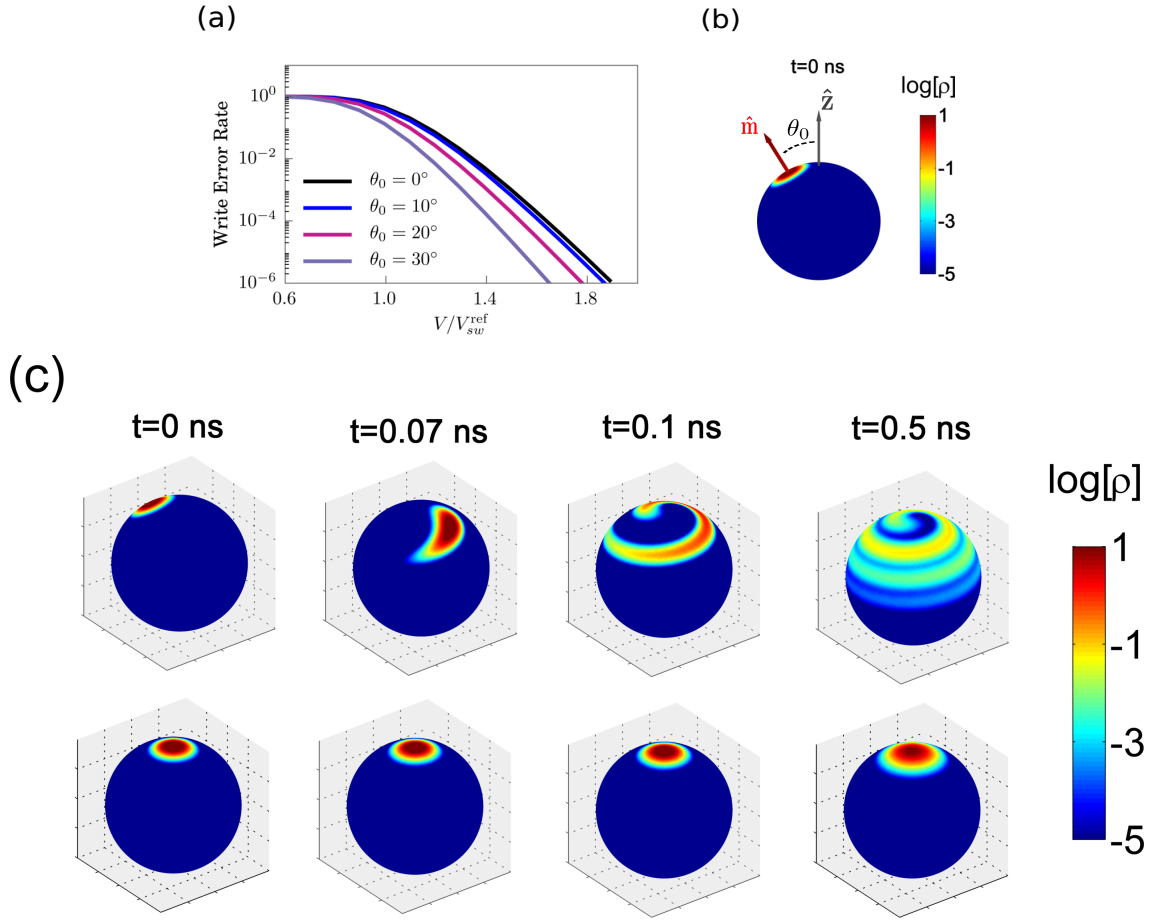


Figure 2.18: (a) The write error rate of magnetization switching from an initial angle. (b) Initial distribution of magnetization probability $\rho(\theta, \phi)$ on a unit sphere. (c) Time evolution of the probability density for canted case (top) with $\theta_0 = 30°$ and collinear case (bottom) with $\theta_0 = 0°$ at $1.25\,\mathrm{V}_{\mathrm{sw}}^{\mathrm{ref}}$ junction bias. Figure reprinted from [1] with permission @2017 IEEE.

Example 2 - orthogonal torque

One possible way to nudge the magnetization out of its stagnation point is to use an orthogonal spin polarizer, which creates a non-collinear spin injection into the free layer as shown in Fig. 2.19(a). This method was used to achieve fast switching in in-plane MTJs [44]. However, for MTJs with perpendicular magneto-anisotropy (PMA), there seems to be a trade-off between the speed and the current density. Fig. 2.19(b) shows a FP simulation of the WER as a function of switching delay for collinear $\theta = 0^\circ$ and non-collinear $\theta = 30^\circ$ spin injections. When the injected current is small, the non-collinear spin injection has a worse WER than the collinear case. The non-collinear spin injection only performs better when the applied current is much larger than the critical current, which is in the dissipative and ultrafast switching regime.

We can understand the trade-off from Fig. 2.19(c): at small applied current, the magnetization switching in PMAs is precessional. At a small angle in the initial stage, a non-collinear torque helps to switch (regime 1) in the first half of the precession cycle, but it brings the magnetic moment right back towards its initial state in the second half (regime 2). This can be avoided only when the current is large enough to allow the magnetic moment to cross the equator in regime 1 itself. Such a large current is hard to achieve in an MTJ and creates undesirable dissipation. Other possible ways to generate an orthogonal torque could include a combination of GSHE and STT where the orthogonal torque can be generated from the GSHE without passing a large current through the MTJ.

Example 3 - Easy-cone magnets

Another example of non-collinear alignment is a system with conical magnetoanisotropy. In these systems, the energy profile can be expressed as $E = -K_{\text{eff}} \cos^2\theta - 0.5K_4 \cos^4\theta$, where K_{eff} is the sum of all second order anisotropy terms including the usual interfacial perpendicular anisotropy and the demagnetization, while K_4 is a higher order anisotropy term. It has been shown that when the free layer CoFeB is within a certain thickness range, the energy minimum appears along the surface of a cone rather than along the axis [45]. Fig. 2.20 shows the FPE simulation of the WER for an easy-cone magnet compared to an

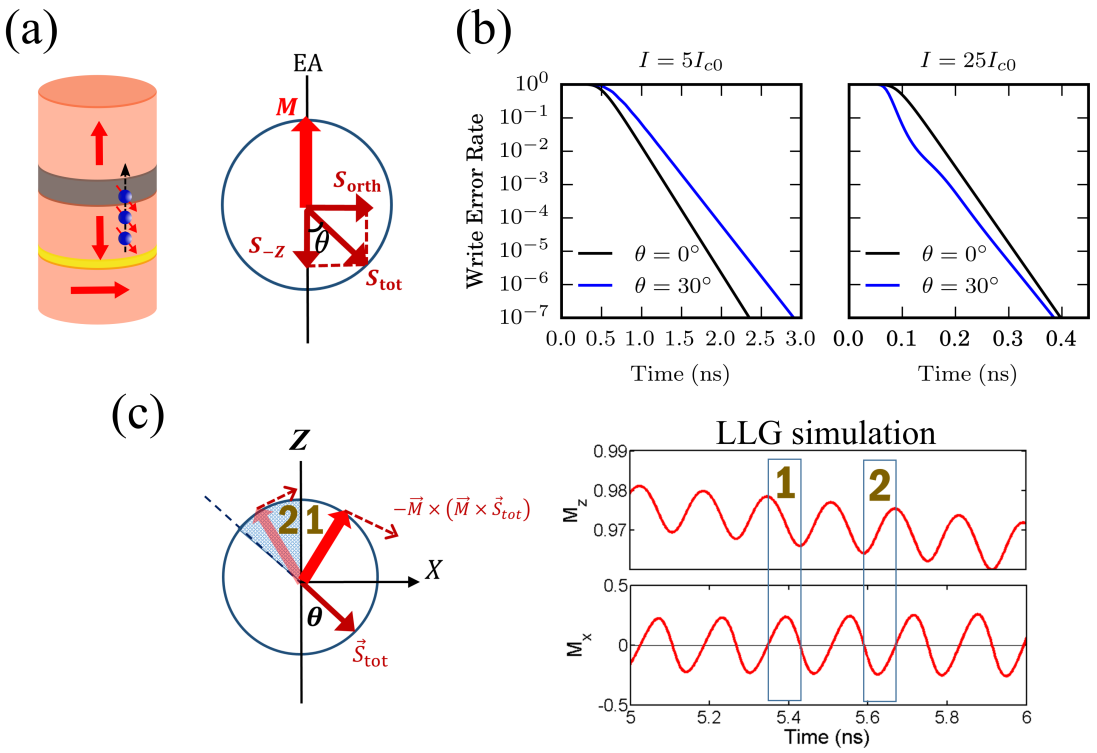


Figure 2.19: (a) Schematic of creating non-collinear spin injection through an extra magnetic layer with an in-plane magnetic moment. By vector addition, we can think of the total effect as a non-collinear torque (with angle θ to the easy axis) injected into the free magnet. (b) Fokker-Planck simulations comparing collinear and non-collinear spin injections for different applied current. (c) Schematic illustration of different regimes. The direction of the spin transfer torques $-\mathbf{M} \times (\mathbf{M} \times \mathbf{S}_{\text{tot}})$ are indicated in different regimes in dashed arrows. LLG simulations of the $m_x(t)$, $m_z(t)$ at the initial stage of switching. $m_z \to -1$ implies the STT helps to switch while $m_z \to +1$ implies the STT is against switching.

easy-axis magnet with the same thermal stability factor Δ. As expected, a 'tilted' initial magnetic moment helps reduce the switching time and error. We see a quick initial evolution since the stagnation point of zero torque has been shifted. Subsequently we reach the shifted stagnation point, and after that the evolution slows down, but since the initial magnetic moment for the easy cone sytem is not along the stagnation point, we still see a steeper slope of the WER.

2.4 Summary of Contributions

In this chapter, we discussed the fundamentals of the magnetic tunnel junction and how to go from material bandstructure to the read/write performance in MTJ-based devices. We

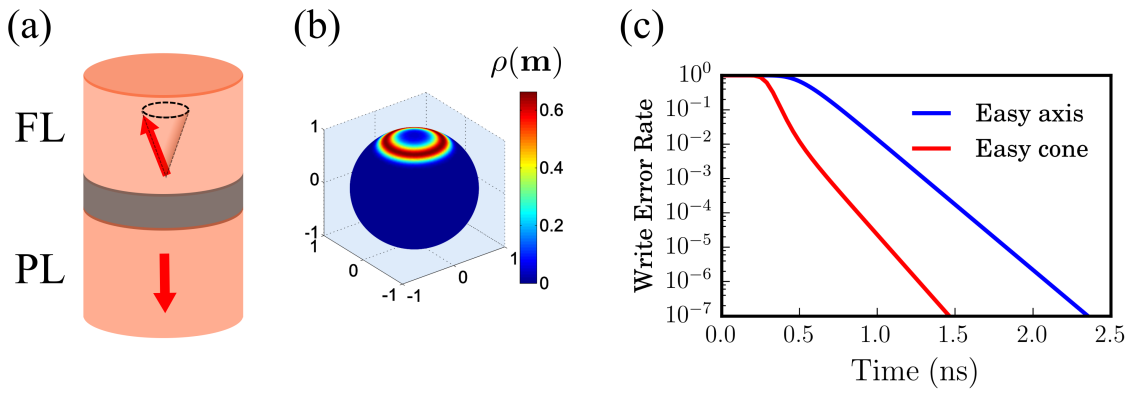


Figure 2.20: (a). Schematic of an MTJ with a free layer that has the easy-cone magnetoansiotropy. (b). The equilibrium probability distribution of the magnetic moment in an easy-cone magnetic structure. $\Delta = 43$, $K_4 = -1.25K_{\text{eff}}$ are used for the easy cone case. (c). Comparing the WER as a function of the switching time for easy-cone and easy-axis magnetic structure. The easy-axis case is set to the same energy barrier $\Delta = 43$. The applied current is set $I = 5I_c$ in both cases, where I_c is the critical current for the easy-axis device.

have also covered the thermal fluctuation induced write error in nanomagnetic switching. Here is a breakdown of our main contributions from this chapter:

1. We have worked with the *Smeagol* group to implement a first-principle NEGF-based approach to calculate the STT in a layered structure (I formulated the problem and conducted benchmark calculations with our native code while Ivan Rungger translated the code to be part of *Smeagol* codebase). While it is being demonstrated on MTJs, it can be used to calculate the spin current in any two-terminal structure.

2. We have studied the effect of different electrode materials and interfacial configurations on the STT efficiency. CoFe with Co-MgO interface has provided the largest torque.

3. We did exploratory analysis on the performance of Half-Heusler/MgO MTJs and identified NiMnSb as the most promising candidate among those we studied.

4. We have developed a general 2D Fokker-Planck solver for STT based switching. It offers a fast and more accurate evaluation of thermal noise induced WER. The code can be easily adapted to other kinds of torque if needed.

5. We have investigated how the WER slope in a perpendicular STT-MRAM depends on the magnetic properties of the MTJ. We also derived a simple analytical equation (in two different forms) to characterize the parameter dependencies of the WER.

6. We used the 2D FPE solver to examine different non-conventional switching schemes and pointed out their strengths/weakness.

7. Finally, throughout this chapter (and also the rest of the dissertation) we have shown that this multiscale approach is very flexible because different levels are loosely coupled (through extraction of few key parameters) and each level can be fine-tuned to different accuracy depending on the needs, e.g. atomistic model vs. free-electron model.

Chapter 3

Alternative Spin Source - Topological Insulator

In a magnetic tunnel junction, the charge current is polarized by the first magnetic layer it passes through. The spin-to-charge ratio (spin polarization η) is determined by the bandstructure of the magnetic layer and tunnel barrier. It is a critical factor determining both the read and the write performance of MTJ-based devices. Since the intrinsic spin polarization cannot exceed 100% in an MTJ, there is an upper bound on the energy efficiency of MTJs because flipping a magnetic requires a certain amount of angular momentum, which, at best, equals the number of electrons passing by (see Eq. 2.30). Therefore the dissipation, which is proportional to the total charge passing through the junction, is limited by the total magnetic moment of the free layer. To break this limit, one has to separate the spin current from the charge current. Recently, progress has been made in novel materials/structures that can generate spin current with efficiency higher than one. Two relevant physics are Giant Spin Hall Effect (GSHE) from heavy metal underlayer and the topological insulator surface states. The creation of spin current in both systems rely on the spin-orbit coupling. Fig. 3.1 shows a schematic of the common setup and spin flow: in-plane charge current only flows through the bottom layer while net spins accumulate on the surfaces of the bottom layer due to spin-orbit coupling. Those net spins at the interface exert a torque on the top magnet. Since the net spins flow in an orthogonal direction as the charge current,

the spin-to-charge ratio is not limited by 100%. In fact, it is determined by the strength of the spin-orbit coupling and the ratio between the magnet/heavy metal interface area and the cross-section of the heavy metal (geometrical gain). Furthermore, there is another advantage of the three-terminal setup: it separates the write current path from the read current path. The readout is still carried out by passing a small current through the MTJ but the write operation is carried out by a different current flowing in-plane of heavy metal or TI. Separating those two processes could reduce the read error because the they can be tuned separately. In this chapter, we study the spin transport on the topological insulator surface. In particular, we want to leverage a PN junction setup on the TI surface to propose a possible experimental measurement for the Klein Tunneling (KT) physics and discuss possible role of Topological Insulator PN Junction (TIPNJ) in spintronics applications. Part of this work was published in [46] coauthored with Yaohua Tan and Avik W Ghosh.

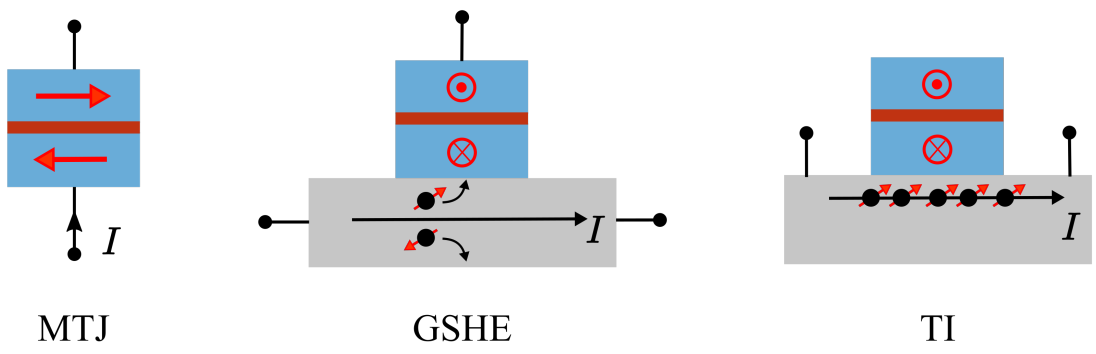


Figure 3.1: Comparison of a switch unit between isolated magnetic tunnel junction (left), the GSHE+MTJ (middle) and the TI+MTJ (right)setups.

3.1 Topological Insulator PN Junction and Klein Tunneling

Topological insulator has a unique bandstructure where the bulk is insulating while the surfaces are metallic with Dirac type bandstructure. Unlike graphene, which also has a Dirac type bandstructure, TI surface state has a distinctive feature where the electron spin and its momentum is 'locked' due to spin-orbit coupling (as shown in Fig.3.2 Left). This helical spin structure makes TI surface a natural spin polarizer. At equilibrium, there is no net spin on the TI surface due to the time-reversal symmetry of its bandstructure. Under finite bias,

there are more electrons moving forward (as from source to drain) and those unbalanced electrons are naturally spin polarized, thus creating net spins on the TI surface as shown in Fig.3.2 Right. The orientation of the net spins can be reversed when the bias is reversed.

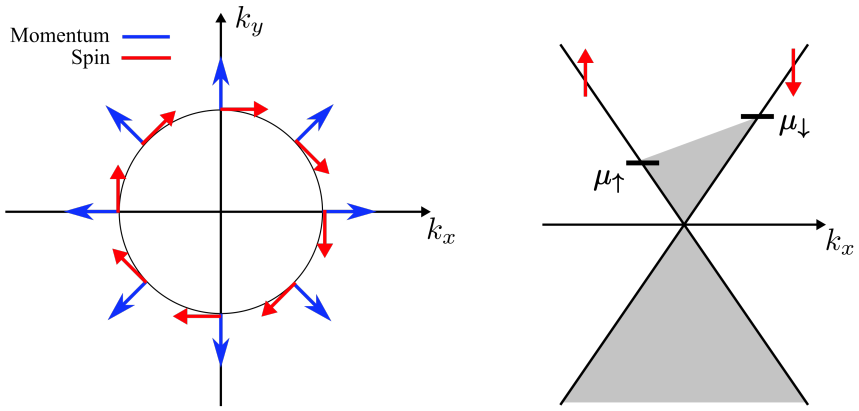


Figure 3.2: Left. Fermi surface in the conduction band of the TI surface. The electron spin and momentum are locked to each other in a chiral manner. Right. Schematics of band filling under non-zero bias. One side of the band with positive group velocity (electron flow direction) is filled higher than the other side of the band.

The similarities between TI surface states and graphene as well as their differences have inspired us to look at the TI PN junction. Several experiments have demonstrated that the TI surface can be chemically doped into P-type or N-type [47, 48]. It can also be doped electrostatically through a gate. One can combine the two approaches to create an in-plane PN junction on the TI surface as shown in Fig. 3.3. Recent experiment has already shown an innovative way to put atomically abrupt gate on TI to create in-plane PN Junction [49]. By varying the gate voltage, the TI surface can switch between a homogeneous PP surface and a NP junction.

The behavior of the Dirac-type massless electron in the presence of a potential barrier has been studied theoretically a long time ago. Klein tunneling (KT) - a consequence of quantum electrodynamics where relativistic particles pass through a high potential barrier unimpeded [50] - is an intriguing phenomenon that has yet to be directly observed in experiments. Researches closest to testing the KT phenomenon are mostly conducted in graphene, with recent progress in the demonstration of anomalous broadened quantized states in a graphene quantum dot [51] and negative index [52] in graphene. Exciting as it

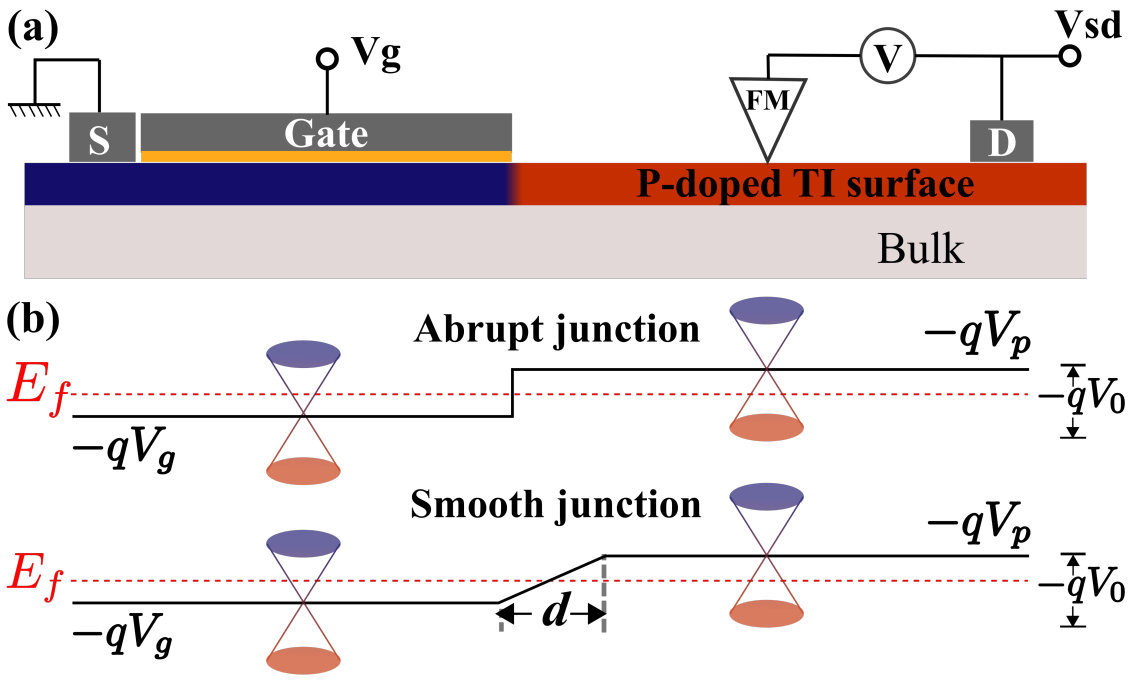


Figure 3.3: Topological insulator PN junction and corresponding potential profile. (a) One possible setup for TI pn junction. A top gate is placed on a P-doped TI surface. (b) Potential profile for a NP junction used in simulations.

is, a direct measurement of Klein tunneling in graphene is very hard because electron flow in graphene sums over all momenta equally and current measurements cannot differentiate those mixed electron momenta. One would expect the TI surface state to share similar physics as graphene due to alike bandstructure. In previous work from our group [4], Masum has shown that a potential barrier created by the PN junction acts like a collimator, filtering out most incoming electrons except for those with small incident angles. Based on those observations, Here we propose a new way to measure KT on 3D TI surface. The core of our proposed idea relies on the measurement of electron spin potential on the TI surface through a spin selective ferromagnetic probe. Since momenta couple with the spin on TI surface, the spin selective probe can also be momentum selective. The method of potentiometric measurement with a ferromagnetic probe to detect the spin structure on the TI surface has been well-established both theoretically [53, 54] and experimentally [55–58]. In this chapter, we model the potentiometric measurements on a TIPNJ and demonstrate from detailed calculations that the angle and voltage-dependent potentials measured at the probe bear direct signatures of Klein tunneling across the PN junction.

3.2 Modeling Potentiometric Measurement on TIPNJ

We will model the spin transport in a TIPNJ with the NEGF method as discussed in sec.2.2.2. Here we adopt a $k \cdot p$ Hamiltonian instead of an atomistic Hamiltonian so that we can simulate larger size structure. Fig. 3.3(a) shows a schematic structure of a TIPNJ in a potentiometric measurement setup. The P-doped TI surface with a top gate on the source side that can swing it electrostatically to N-type. The rest of the P-type TI surface is exposed and a ferromagnetic probe is placed on top of the exposed surface to monitor the voltage at different gate bias and angular orientations of the probe magnetic moment.

The $k \cdot p$ Hamiltonian describing the TI surface states close to the Dirac point [59]:

$$H = v_F \hat{\mathbf{z}} \cdot (\boldsymbol{\sigma} \times \mathbf{p}) \tag{3.1}$$

where $\hat{\mathbf{z}}$ is the normal vector of the surface and v_F is the speed of electrons near the Dirac point. $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. It should be emphasized that this parameterized surface Hamiltonian ignores any bulk leakage current that could control the strength of the measured voltage. In binary TI compounds such as BiSb, Bi_2Se_3, Bi_2Te_3, it can be challenging to separate the surface contribution from the dominant bulk contribution [60–62]. One possible solution is to use ternary compounds like Bi_2Te_2Se with low carrier density in the bulk [63]. Minimizing the leakage current into the bulk of TI is still an active research topic that is outside the scope of this dissertation. Here we only discuss the pure surface states of 3D TI. The electrostatic potential across the TI PN junction is given by:

$$\begin{aligned} V(x) &= -qV_p, \ \text{ exposed P side} \\ &= -qV_g', \ \text{ gate side} \end{aligned} \tag{3.2}$$

where $E_p = -qV_p$ is the energy difference between the local electron chemical potential and the Dirac point ($E = 0$). V_g' is the effective potential on the source side of the TI surface under the gate voltage V_g as shown in Fig. 3.3. For the simplicity of the discussion, we assume good electrostatic control of the gate on the TI surface (gate capacitance much larger than other capacitors in the system) that gives $V_g' \approx V_g$. Two potential profiles are

depicted in Fig. 3.3(b), one with an abrupt potential change at the junction interface while the other assumes a smooth transition. We will first derive the analytical results for electron transmission in abrupt junctions and then extend it to smooth junctions, which is closer to a realistic profile [51]. For smooth junctions, the transition region between N and P is set to 50 nm wide and the FM probe is placed 80 nm from the junction interface.

In the ballistic limit, the electrons only scatter at the PN junction interface. A ferromagnetic voltage probe weakly coupled to the TI surface can detect the local chemical potential of the non-equilibrium electrons with different spin orientations. To calculate the voltage measured by the FM probe, we treat it as a third contact (Büttiker probe) besides the source and the drain. From the Landauer theory [15], the exchange of electrons between the voltage probe and the TI surface follows the terminal current equation:

$$\begin{aligned} I_{in} &= \mathrm{Tr}\left[\Gamma_{\mathrm{FM}} G^n\right] = \mathrm{Tr}\left[\Gamma_{\mathrm{FM}}\left(f_s A_s + f_d A_d\right)\right] \\ I_{out} &= f_p \mathrm{Tr}\left[\Gamma_{\mathrm{FM}} A\right] = f_p \mathrm{Tr}\left[\Gamma_{\mathrm{FM}}\left(A_s + A_d\right)\right] \end{aligned} \tag{3.3}$$

where I_{in} (I_{out}) is the incoming (outgoing) currents through the probe. Γ_{FM} is the coupling between the FM probe and the TI surface. G^n is the correlation matrix (G^n is connected to $G^<$ in sec.2.2.2 by $G^n = -iG^<$) while $A_s(A_d)$ are the partial spectral functions populated by the source (drain). $A = A_s + A_d$ is the total spectral function. f_s, f_d, f_p are the Fermi-Dirac distribution functions of the source, drain and the floating probe respectively.

The coupling between the FM probe and the TI surface depends on the magnetization of the FM probe $\mathbf{m} = (m_x, m_y, m_z)$ and electron spin $\boldsymbol{\sigma}$ of the TI surface:

$$\Gamma_{\mathrm{FM}}(\mathbf{m}) = \gamma_0\left(1 + P_{\mathrm{FM}}\mathbf{m}\cdot\boldsymbol{\sigma}\right) \tag{3.4}$$

where $\gamma_0 = \frac{\gamma_p + \gamma_{ap}}{2}$ is the average coupling between the FM probe and the TI surface when the magnetization of the probe is in parallel or anti-parallel alignment with the surface electron spin. $P_{\mathrm{FM}} = (\gamma_p - \gamma_{ap})/(\gamma_p + \gamma_{ap})$ is the 'polarization' of the FM probe, representing the sensitivity of the FM probe to the electron spins.

The voltage signal measured by the FM probe is determined by its distribution function f_p, which can be solved based on the condition that a voltage probe draws zero net current

$I_{in} = I_{out}$:

$$\begin{aligned} f_p(\mathbf{m}) &= \frac{(f_s - f_d)\,\mathrm{Tr}\left[\Gamma_{\mathrm{FM}} A_s\right]}{\mathrm{Tr}\left[\Gamma_{\mathrm{FM}} A\right]} + f_d \\ &= \lambda(\mathbf{m})(f_s - f_d) + f_d \end{aligned} \tag{3.5}$$

f_p varies when the magnetization $\mathbf{m}$ points to different directions. We use the dimensionless parameter $\lambda(\mathbf{m})$ to characterize the dependence of the voltage signal on the direction of the magnetization. At low-temperature and small bias, the Fermi-Dirac distribution reduces to a step function and chemical potential of the probe can be expressed as:

$$\mu_p(\mathbf{m}) = \lambda(\mathbf{m})(\mu_s - \mu_d) + \mu_d \tag{3.6}$$

Experimentally instead of switching the magnetization of the FM probe we can drive current along two opposite directions (source to drain and vice-versa), then relate the measured voltage difference $\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})$ to $\Delta\lambda(\mathbf{m}) = \lambda(\mathbf{m}) - \lambda(-\mathbf{m})$ through the charge current and the ballistic resistance of the junction:

$$\begin{aligned} \Delta\lambda(\mathbf{m}) &= \frac{\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})}{qIR_B} \\ R_B &= \frac{h}{q^2 T(E_f)} \end{aligned} \tag{3.7}$$

where R_B is the gate voltage dependent ballistic resistance of the junction, calculated using the average transmission at the Fermi energy. We can further define a quantity $p(\mathbf{m})$ for the measured 'polarization' of the TI surface electrons along the magnetization direction $\mathbf{m}$:

$$\begin{aligned} p(\mathbf{m}) &= \frac{\lambda(\mathbf{m}) - \lambda(-\mathbf{m})}{\lambda(\mathbf{m}) + \lambda(-\mathbf{m})} \\ &= \frac{\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})}{\mu_p(\mathbf{m}) + \mu_p(-\mathbf{m}) - 2\mu_d} \end{aligned} \tag{3.8}$$

The physical interpretation of Eq. 3.8 becomes obvious when we substitute Eq. 3.4 into Eq. 3.8 and see that $\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}(\mathbf{m})A\right] = \mathrm{Tr}\left[\Gamma_{\mathrm{FM}}(-\mathbf{m})A\right]$ due to the time reversal symmetry of TI

surface states. Eq. 3.8 reduces to:

$$p(\mathbf{m}) = P_{\mathrm{FM}} \frac{\mathrm{Tr}[(\mathbf{m}\cdot\boldsymbol{\sigma})\gamma_0 A_s]}{\mathrm{Tr}[\gamma_0 A_s]} \tag{3.9}$$

when Eq. 3.9 is evaluated in the bias window, it indicates the spin polarization of the non-equilibrium electrons along direction $\mathbf{m}$. Notice that P_{FM} also appears in the equation to account for the sensitivity of FM probe. Our definition is compatible with the polarization defined in [53] for homogeneous TI surface.

3.2.1 Analytical Results

For infinitely large TI surface with an abrupt PN junction potential profile (Fig. 3.3(b)), we can derive quasi-analytical equation to characterize the potentiometric signals. The eigen-functions to the Hamiltonian 3.1 are given by:

$$\begin{aligned} |\psi\rangle_\sigma &= \frac{1}{\sqrt{2S}} \begin{pmatrix} 1 \\ -sie^{i\theta} \end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ s &= \mathrm{sgn}(E_{\mathbf{k}}) \end{aligned} \tag{3.10}$$

where $\mathrm{sgn}(E_k) = 1$ is for the N type and $\mathrm{sgn}(E_k) = -1$ for the P type. S is the surface area. At the PN junction, the transmitted and scattered electrons are connected by:

$$\begin{aligned} |\psi\rangle_\sigma &= |\psi_i\rangle_\sigma + r|\psi_r\rangle_\sigma \\ |\psi\rangle_\sigma &= t|\psi_t\rangle_\sigma \end{aligned} \tag{3.11}$$

where $|\psi_i\rangle_\sigma, |\psi_r\rangle_\sigma, |\psi_t\rangle_\sigma$ are the incoming, reflected and transmitted electron wave functions respectively (see Fig. 3.4) and r/t is the reflection/transmission coefficient. Solving Eq. 3.11 with wave function continuity condition at the junction $\mathbf{r} = 0$, we get the transmission

coefficient:

$$\text{for NP} \qquad t = \frac{e^{i\theta_i} - e^{-i\theta_t}}{e^{-i\theta_i} - e^{i\theta_t}} \tag{3.12}$$

$$\text{for PP} \qquad t = \frac{e^{i\theta_i} - e^{i\theta_t}}{e^{-i\theta_i} - e^{-i\theta_t}} \tag{3.13}$$

It is convenient to replace θ_t with $\theta_t + \pi$ in the NP case so that the expressions for t are the same in both PP and NP cases. The incident and transmitted angles are connected through the conservation of k_y across the junction: $(E - qV_n)\sin\theta_i = (E - qV_p)\sin\theta_t$, we can calculate the transimission probability $|t|^2$:

$$|t|^2 = \frac{\cos^2\theta_i}{\cos^2\left(\frac{\theta_i + \theta_t}{2}\right)} \tag{3.14}$$

Smooth PN junction. The effect of a smooth PN junction (as shown in Fig. 3.3(b)) is an additional exponential factor from the abrupt junction case (Eq.3.14). Here we borrow the result for the transmission coefficient $|t|^2_{\text{smooth}}$ from the smooth graphene PN junction (see details in [64]):

$$\begin{aligned} |t|^2_{\text{smooth}} &= \frac{\cos^2\theta_i}{\cos^2\left(\frac{\theta_i + \theta_t}{2}\right)} e^{-\pi \frac{k_i k_t}{k_i + k_t} \sin\theta_i \sin\theta_t d} \\ &= \frac{\cos^2\theta_i}{\cos^2\left(\frac{\theta_i + \theta_t}{2}\right)} e^{-\pi \frac{\hbar v_F}{qV_0} k_t^2 \sin^2\theta_t d} \end{aligned} \tag{3.15}$$

where d is the transition length between N region and P region. $V_0 = |V_p - V_n|$ is the potential difference from N region to P region. Notice that the exponential factor should only be added in cases with different types (such as PN, NP) across the junction . In other cases (such as PP', NN'), the difference between abrupt and smooth junctions is negligible, which can be seen in our comparison between analytical and numerical results later.

In the small bias window near E_f, the charge current is given by:

$$\begin{aligned} I &= \frac{q}{h}T(E_f)(\mu_s - \mu_d) \\ T(E_f) &= \frac{qV_pW}{hv_F}\int_{-\pi/2}^{\pi/2}|t|^2\cos\theta_t d\theta_t \end{aligned} \tag{3.16}$$

where $T(E_f)$ is the electron transmission across the junction and W is the width of the TI surface. Knowing the transmission coefficient allows us to calculate $\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}A_s\right]$ (Eq. 3.3-3.6):

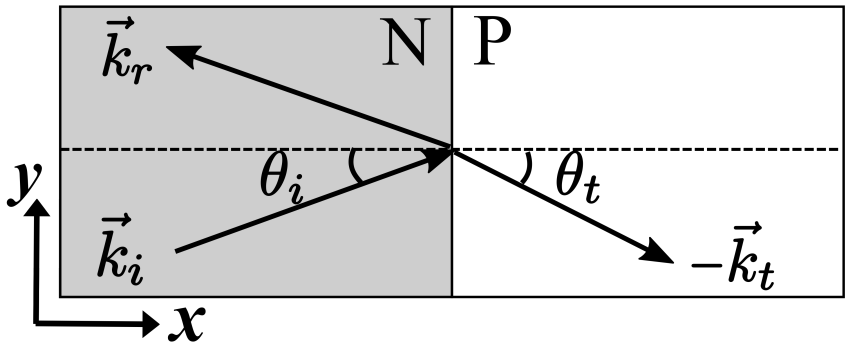


Figure 3.4: Incident, reflected and transmitted electrons waves in a TI pn junction.

$$\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}A_s\right] = W\sum_{v_x(\mathbf{k}_t)>0}[1 + P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)t(\mathbf{k}_t)]\delta(E_f - E(\mathbf{k}_t)) \tag{3.17}$$

$\mathbf{s}(\mathbf{k}_t)$ is the spin orientation of the transmitted electron with wave vector $\mathbf{k}_t$. $t(\mathbf{k}_t)$ is the transmission coeffecient given by Eq. 3.13. $\lambda(\mathbf{m})$ in Eq. 3.6 can then be calculated:

$$\begin{aligned} \lambda(\mathbf{m}) = \frac{\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}A_s\right]}{\mathrm{Tr}\left[\Gamma_{\mathrm{FM}}A\right]} &= \frac{\sum_{v_x(\mathbf{k}_t)>0}[1 + P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)t(\mathbf{k}_t)]\delta(E_f - E(\mathbf{k}_t))}{\sum_{\mathbf{k}_t}[1 + P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)]\delta(E_f - E(\mathbf{k}_t))} \\ &= \frac{\sum_{v_x(\mathbf{k}_t)>0}[1 + P_{\mathrm{FM}}\mathbf{m}\cdot\mathbf{s}(\mathbf{k}_t)t(\mathbf{k}_t)]\delta(E_f - E(\mathbf{k}_t))}{\sum_{\mathbf{k}_t}\delta(E_f - E(\mathbf{k}_t))} \end{aligned} \tag{3.18}$$

The last step in Eq. 3.18 holds because each pair of states $\mathbf{k}_t, -\mathbf{k}_t$ cancel out due to the time reversal symmetry of TI surface Hamiltonian $\mathbf{s}(\mathbf{k}_t) = -\mathbf{s}(-\mathbf{k}_t)$. Assume the ferromagnetic voltage probe has an in-plane magnetization (m_x, m_y). Substitute the transmission coefficient into Eq. 3.18 and replace $\sum$ with $\frac{S}{4\pi^2}\int d^2k$. For the denominator, notice that it is just the

density of states on the P side. Therefore $\lambda(\mathbf{m})$ can be calculated:

For PP:

$$\lambda(\mathbf{m}) = \left|\frac{E_f - qV_g}{E_f - qV_p}\right| \int_{-\pi/2}^{\pi/2} \frac{\cos^2\theta_i \left(1 + P_{\mathrm{FM}} m_x \sin\theta_t - P_{\mathrm{FM}} m_y \cos\theta_t\right)}{2\pi \cos^2\left(\frac{\theta_i+\theta_t}{2}\right)} d\theta_t \tag{3.19}$$

For NP:

$$\lambda(\mathbf{m}) = \left|\frac{E_f - qV_g}{E_f - qV_p}\right| \int_{-\pi/2}^{\pi/2} \left\{ \frac{\cos^2\theta_i \left(1 + P_{\mathrm{FM}} m_x \sin\theta_t - P_{\mathrm{FM}} m_y \cos\theta_t\right)}{2\pi \cos^2\left(\frac{\theta_i+\theta_t}{2}\right)} \times \exp\left[\frac{(E_f - qV_p)^2 \pi d}{|V_g - V_p| q\hbar v_F}\right] \right\} d\theta_t \tag{3.20}$$

where $\theta_t = \sin^{-1}[(E_f - qV_g)/(E_f - qV_p)\sin\theta_i]$.

3.2.2 Numerical Approach

To numerically simulate the ballistic electron/spin transport in TIPNJ from the $k \cdot p$ model, an artificial term $\sigma^z = \gamma\hbar v_F \sigma^z (k_x^2 + k_y^2)$ is added to the surface Hamiltonian Eq.3.1 to avoid the fermion doubling problem as have been done in the previous studies [4,53]. The modified TI surface Hamiltonian is discretized on a square lattice by the finite difference method [4]:

$$H = \sum_i \epsilon c_i^\dagger c_i + \sum_i \left(t_x c_{i,i}^\dagger c_{i,i+1} + \mathrm{H.C.}\right) + \sum_j (t_y c_{j,j}^\dagger c_{j,j+1} + \mathrm{H.C.})$$

$$\epsilon = -4\hbar v_F \frac{\alpha}{a}\sigma^z \quad t_x = \hbar v_F \left[\frac{i}{2a}\sigma^y + \frac{\alpha}{a}\sigma^z\right] \tag{3.21}$$

$$t_y = \hbar v_F \left[-\frac{i}{2a}\sigma^x + \frac{\alpha}{a}\sigma^z\right] \tag{3.22}$$

where a is the square mesh size ($a = 5\,\mathrm{nm}$ is chose for the simulations). $\alpha = \gamma/a$ is a fitting parameter and $\alpha = 1$ describes the correct bandstructure near the Dirac cone [4]. Periodic boundary condition is assumed in the transverse direction to simulate infinitely wide TI surface. The retarded green's function is given by:

$$G^R(E, \mathbf{k}_\perp) = (E + \delta - H(\mathbf{k}_\perp) - \Sigma_s(E, \mathbf{k}_\perp) - \Sigma_d(E, \mathbf{k}_\perp))^{-1} \tag{3.23}$$

where E is the energy and $\mathbf{k}_\perp$ is the transverse wavevector. $\Sigma_{s,d}$ are self-energies from the source and drain. The FM probe is assumed to be weakly coupled to the TI surface so the effect of Σ_p (assign a very small value) on electron transport is neglected when calculating $G^R(E, \mathbf{k}_\perp)$. Then the spectral functions can be calculated numerically through the NEGF formalism:

$$\begin{aligned} A_s &= G^R \Gamma_s G^{R\dagger}, \quad \Gamma_s = i(\Sigma_s - \Sigma_s^\dagger) \\ A_d &= G^R \Gamma_d G^{R\dagger}, \quad \Gamma_d = i(\Sigma_d - \Sigma_d^\dagger) \end{aligned} \tag{3.24}$$

$\lambda(\mathbf{m})$ is then calculated from the matrix forms of Γ_{FM}, A_s, A.

3.3 Expected Klein Tunneling Signals

3.3.1 Gate Voltage Dependent Signal: From PP to NP Regime.

The impact of a TIPNJ on surface electron transport is summarized schematically in Fig. 3.5(a). Consider a small source-drain bias near the Fermi energy, as shown in Fig. 3.3(b). As the gate voltage varies from $V_g = V_p$ to $V_g = -V_p$, the TI switches from a homogeneous P-doped surface to an NP junction. Electrons see a potential barrier from the N region to the P region. In a normal semiconductor, such a barrier creates decaying electron waves in the P region and results in a vanishing current. For Dirac type TI surface, however, the junction acts like a collimator for electrons, filtering out electrons with large incident angles but preserving the normally incident modes that cannot back-scatter due to spin conservation. The resulting electron transmission for various gate voltages is plotted in Fig. 3.5(a). This behavior can translate to the gate voltage dependence of $\Delta\lambda(\mathbf{m})$ defined in Eq. 3.7. Fig. 3.5(b) shows the gate voltage dependence of $\Delta\lambda = \lambda(-\hat{y}) - \lambda(\hat{y})$. $\Delta\lambda$ first goes down as we move from PP to PI (I: intrinsic), then goes up a bit and saturates in the NP region. The decrease of $\Delta\lambda(\mathbf{m})$ in the PP region is due to a mismatch of modes between the gate side and the probe side as the Fermi energy approaches the Dirac point (intrinsic doping) on the gate side. When $V_g = 0$ V the Fermi level on the gate side lies exactly on

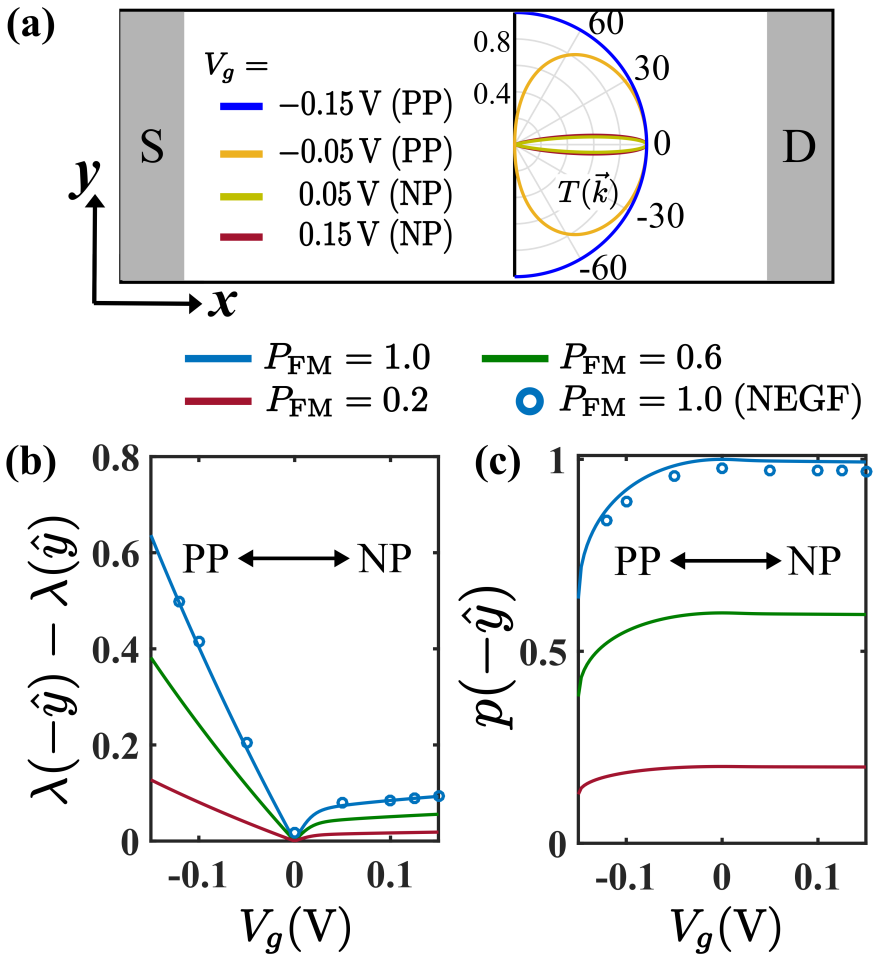


Figure 3.5: (a) Schematic plot of the electron transmission through the junction at different gate voltages. (b) Gate voltage dependence of $\Delta\lambda(-\hat{\mathbf{y}}) = \lambda(-\hat{\mathbf{y}}) - \lambda(\hat{\mathbf{y}})$ for various probe sensitivities. (c) The measurable polarization of TI surface electrons along $\hat{y}$ direction. The circles are benchmark results from NEGF simulations.

the Dirac point with zero density of states and thus $\Delta\lambda(\mathbf{m}) = 0$. It is worth mentioning that the 'zero' is an idealized simplification. A rigorous calculation involves integration over the bias window which would result in a small but non-zero value.

When the gate side is switched to the N region, the angular filtering effect shows up and results in a smaller value of $\Delta\lambda(\mathbf{m})$ compared to its symmetric point (with the same $|V_g|$) in the PP region. Since the normal incident mode is not affected by the potential barrier, a small but near constant $\Delta\lambda(\mathbf{m})$ shows up in the NP region as V_g increases. This asymmetry between PP and NP region and the non-vanishing $\Delta\lambda(\mathbf{m})$ in the NP region separates the TI surface from other 2D systems such as graphene or Rashba systems where

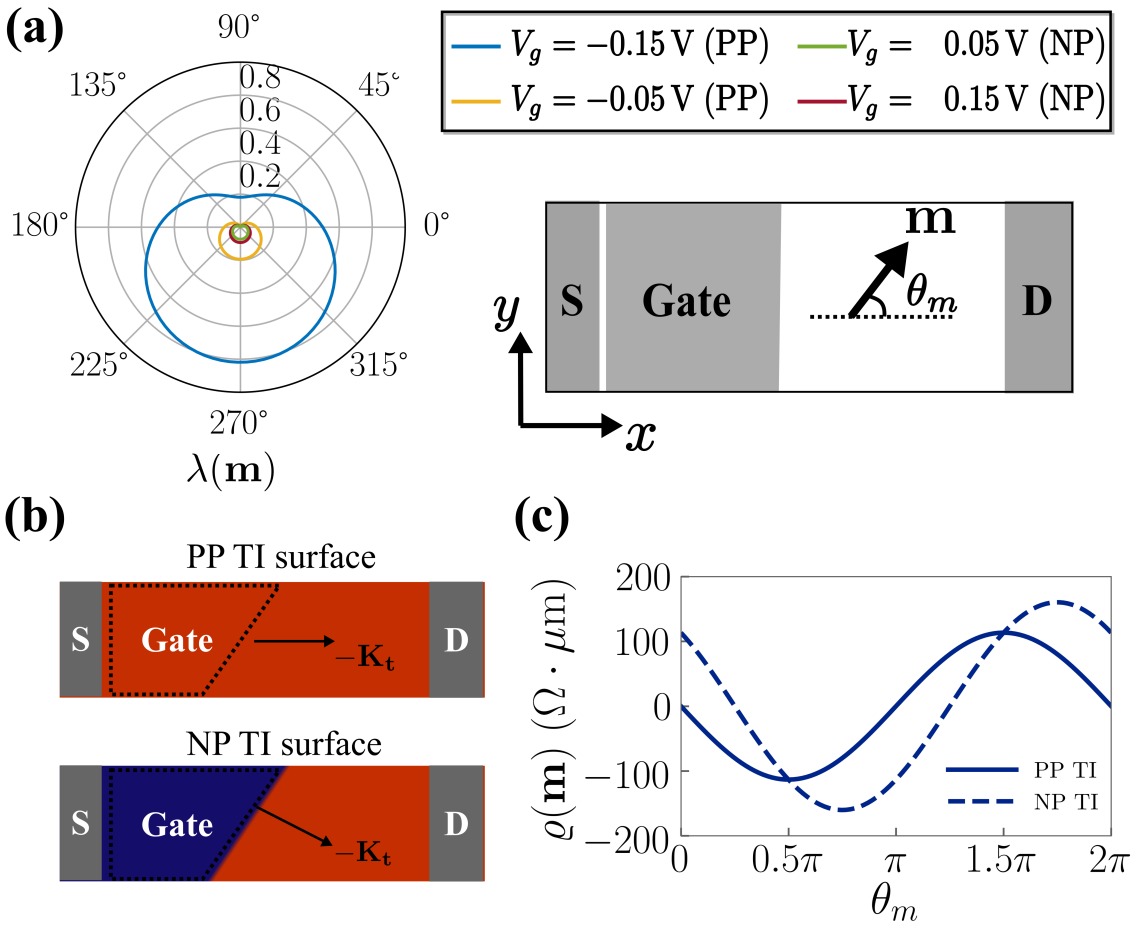


Figure 3.6: (a) Angular dependence of $\lambda(\hat{\mathbf{m}})$ for different gate voltages. (b) Schematics of a tilted gate on TI surface. (c) Compare the angular dependence of $\varrho(\mathbf{m})$ in PP and NP cases. A phase shift equal to the tilt angle is expected in the angular signals.

there is either $\Delta\lambda(\mathbf{m}) = 0$ in all regions due to spin degeneracy (graphene) or $\Delta\lambda(\mathbf{m}) = 0$ in the transmitted N region due to decaying waves in a potential barrier for massive tunneling electrons (Rashba).

We can further demonstrate collimation in TIPNJ by plotting polarization $p(-\hat{\mathbf{y}})$ as a function of the gate voltage, as shown in Fig. 3.5(c). Electrons moving along the $\hat{\mathbf{x}}$ direction carry $-\hat{\mathbf{y}}$ spin. Right across the NP junction, filtered electrons have a narrower $\mathbf{k}$ distribution compared to the homogeneous PP case, and thus higher (close to 100%) spin polarization. In reality, this kind of measurement is limited by the sensitivity of the FM probe, but a clear and significant increase of polarization should be observable as we proceed from homogeneous PP case to NP doping with reasonable P_{FM} values.

3.3.2 Angular Dependent Signal in a Tilted Junction

Our discussion so far focused on measurement along two opposite directions ($\pm\hat{\mathbf{y}}$), assumed to be orthogonal to the electron transport direction. For an arbitrary orientation of the magnetization $\mathbf{m}$, $\lambda(\mathbf{m})$ is a cosine function of the relative angle between the magnetization $\mathbf{m}$ and the spin orientation of the non-equilibrium electrons. Fig. 3.6(a) shows the angular dependence of $\lambda(\mathbf{m})$ with different gate voltages. From homogeneous PP to NP junction, apart from the change in the magnitude, $\lambda(\mathbf{m})$ remains the same cosine function. This is because the FM probe cannot isolate individual modes but measures the sum over all transport modes with different weights. In our basic setup, the PN junction filters electrons with large incident angles but the transmitted modes are still symmetrically distributed with respect to $\hat{\mathbf{x}}$. Therefore the average momenta in the PP and NP junction only differ from each other by their magnitude. To experimentally observe the normal tunneling mode, we can put a tilted gate that is not orthogonal to the transport direction (see Fig. 3.6(b)). A tilted gate will not affect the results from the homogeneous case but will collimate the electrons to a different angle for NP, thereby creating a phase shift in the angular dependence of $\lambda(\mathbf{m})$. Since we only care about the phase of $\lambda(\mathbf{m})$, we can define an angular function as:

$$\varrho(\mathbf{m}) \quad = \quad \frac{\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m})}{qJP_{\mathrm{FM}}} \tag{3.25}$$

which will scale $\Delta\mu_p(\mathbf{m})$ by the charge current density J and make the PP and NP cases easier to compare, as shown in Fig. 3.6(c).

3.3.3 Discussions on Experimental Constraints

Ballistic versus Diffusive Limit

Note that we formulated our equations Eq.3.3-3.8 assuming a ballistic channel where $\mu_p(\mathbf{m})$ can be directly related to the chemical potentials from the source and drain. However, our analysis can be easily adopted to a diffusive system with a different interpretation. μ_s and μ_d in the previous discussions should be replaced by the local chemical potential $\mu_\uparrow$ and $\mu_\downarrow$ for spin up and spin down channels, as indicated in Fig. 3.7. All of our previous discussions

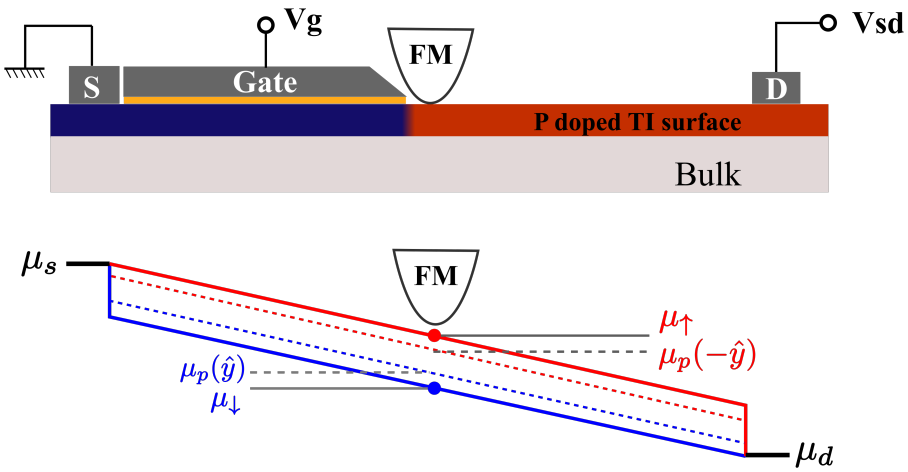


Figure 3.7: Top. A possible experimental setup for diffusive system. Bottom. A schematic chemical potential profile in a diffusive system.

are still valid given the following conditions: in a diffusive system, a momentum scattering event can disrupt the collimation effect of the NP junction. To be able to detect the Klein tunneling physics of the junction, the probe needs to be placed very close to the junction, preferably within the mean free path of the TI surface electrons ($\sim 120\,\text{nm}$ estimated in Bi_2Te_3 [65]). To place the probe in such short distance from the gate edge, it possibly requires either a very thin gate ($< 100\,\text{nm}$) or specially etched shape (as shown in Fig. 3.7) to avoid crashing with the probe. From the discussion of $p(\mathbf{m})$ earlier, we need information on μ_d (replace by $\mu_\downarrow$) at the junction. One way to do this is to use a normal voltage probe to map out the resistance from junction to the drain to extract the slope shown in Fig. 3.7, and then estimate the local electrochemical potential from the applied drain bias.

Possible Experimental Setup.

Ideally we would like to rotate the magnetization of the ferromagnetic probe to map out the angle-dependent voltage signals. To our knowledge such a reorientation of an FM probe is challenging. Even fixing the magnetization of the FM probe orthogonal to the transport direction is not straightforward. Instead, we propose placing two separate gates near the source and drain (Fig. 3.8), creating a symmetric system. Only one of the gates is used at a time to create an N region on one side. When the current direction is switched, we flip the gate polarities on both sides and the entire system is mirrored. Another possibility is to put

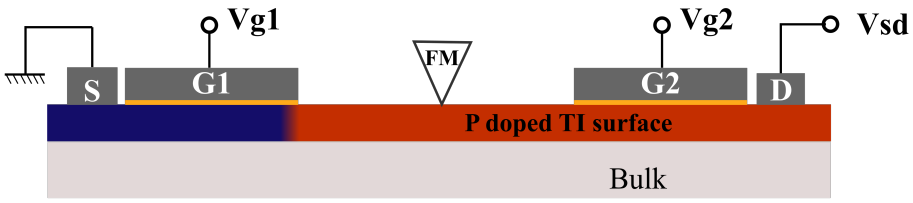


Figure 3.8: One possible experimental measurement set-up. Intrinsically P-doped topological insulator under a N type gate near the source. The FM probe is placed on the exposed P side.

two probes (one FM, one normal) close to each other and measure the voltage difference between them. It is not difficult to show that $\mu_p(\mathbf{m}) - \mu_p(-\mathbf{m}) = 2(\mu_p(\mathbf{m}) - \mu_{nm})$ where μ_{nm} is the voltage measured at the non-magnetic probe.

3.4 TIPNJ as a Spin Amplifier?

In a previous work from our group [4], it has been shown that a PN junction can tune the charge-to-spin conversion efficiency (usually referred as the Spin Hall Angle θ_{SH}) up to around 20, much higher than the spin hall angle measured in most GSHE systems such as 0.07 in Pt [66], $0.12 - 0.15$ in β-Ta [12], and 0.12 in Pt-doped Au [67]. The main idea is illustrated in Fig. 3.9. When the TI surface swings from the homogeneous NN regime to the NP regime, the overall charge current is suppressed while the spin current is amplified because of the spin-momentum locking feature of the TI bandstructure. These results seem to suggest that one can use a nanomagnet as the source contact and switch it with a TIPNJ. However, previous simulations were done on a pure TI surface that assumes the source and the drain are also extended TI surface. Therefore the spin amplification is only for the current that flows inside the TI surface. We will show that this 'intrinsic' spin amplification doesn't translate to an external gain on the magnet, at least for the particular configuration where the magnet is the source contact.

We set up a toy system by putting a nanomagnet on top of the TI surface as shown in Fig. 3.10. The source magnet is modeled by a simple free electron model mentioned in sec. 2.1 with energy split $\Delta = 2.15\,\text{eV}$ between the spin-up and the spin-down channel. Due to the small size of the magnet and the simplified single s-orbital tight-binding model, we

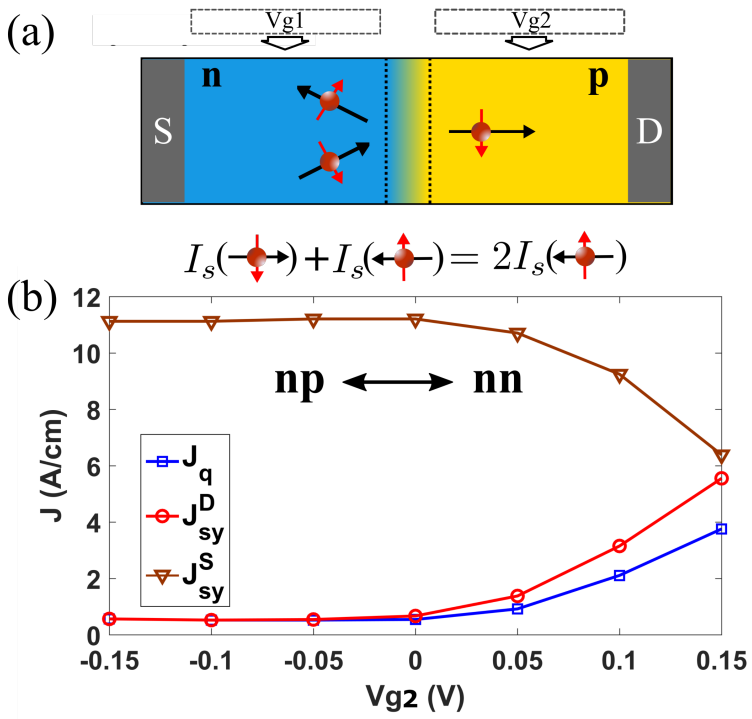


Figure 3.9: (a) Schematic electron (and spin) trajectory on the TIPNJ. The charge current is reduced when the junction is in PN mode but the spin current is multiplied on the source side due to spin flip when the electrons are reflected back to source. (b) NEGF simulation of charge and spin current as a function of drain side gate voltage V_{g2}, assuming the other gate is fixed at $V_n = 0.15\,\mathrm{V}$. The source-drain bias is set at $V_{sd} = 0.1\,\mathrm{V}$. The simulated TI surface has dimensions $200\,\mathrm{nm} \times 120\,\mathrm{nm}$ with a splitting d = 100 nm between the N regime and the P regime. Other details can be found in Ref. [4].

adjust the effective electron mass for the magnet to get a reasonable spin polarization of electrons at the Fermi level.

From Fig. 3.10(b), it is clear that the in-plane spin current J^{sy}_{TI} is different from the spin current that flows from the magnet to the TI surface. Although J^{sy}_{TI} can be much larger than the charge current, the actual spin current that flows through the magnetic contact is smaller than the charge current. To understand this discrepancy, we use the simplified picture of TI mentioned earlier in Fig. 3.2 Right: we look at the two channels spin-up (correspond to $+y$) and spin-down (correspond to $-y$). The ferromagnetic contact also has those two channels. On the TI surface, the electrons in the spin-up channel can only move from right to left ($-k_x$) while the electrons in the spin-down channel move in the opposite direction like two one-way streets. This feature results in the spin amplification effect. However, when the magnet exchanges electrons with the TI surface, both spins can be injected into or taken

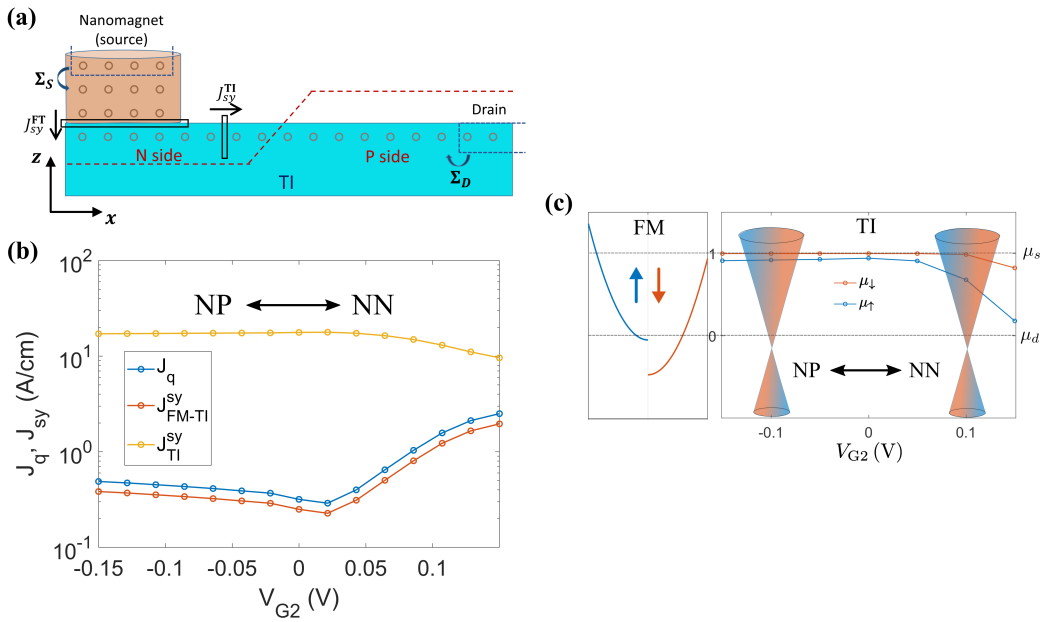


Figure 3.10: charge and spin transport in a TIPNJ with a ferromagnetic contact. (a) Discretization of the combined FM-TIPNJ setup. The left contact (source) is assumed to the extension of the ferromagnet while the right contact (drain) remains an extension of TI surface. The magnetic moment of the FM is oriented to the $-y$ direction. (b) charge and spin current density calculated at different locations. J_q is the charge current density (conserved throughout the system). $J^{sy}_{\rm TI}$ is the spin current (polarized along y direction) calculated on the TI surface at the source side. $J^{sy}_{\rm FM-TI}$ is the spin current density at the FM-TI interface. The source-drain bias V_{sd} is fixed at 0.1 V. (c) The average chemical potential spin-up/down channels in FM contact (schematics) and TI surface (simulation).

out of the magnet. What determines the influx or outflow of spins from the magnet is the chemical potential difference between the magnet and the TI surface. If the magnet is the source, then it has the highest chemical potential in both spin channels (see. Fig. 3.10(c)). Therefore there is only net outflow of spins from the magnet to the TI surface in both spin channels, albeit at different rates. Under such condition, the spin polarization cannot exceed one just like the magnetic tunnel junction. One might ask where does the 'extra' spin current go on the TI surface? Unlike charge current, the spin current doesn't need to be conserved. In our simulation with a finite TI surface, the electrons hit the left boundary of the TI surface and reflected back with flipped spin. In reality, the TI surface is a closed space and the spin-up and the spin-down channels form their individual closed loop. Therefore the extra spins can go around the surface and eventually taken out from the drain contact.

With the understanding of why the 'intrinsic gain' doesn't translate to an 'external gain',

one can further explore how to overcome it in the future work. For example, instead of using the nanomagnet as a source contact, if we use it as the third terminal with its Fermi energy lie in between the chemical potentials of the spin-up and the spin-down channel of the TI surface, then it can draw majority spins from the TI surface while deposit minority spins to the TI surface. By tuning the magnet's Fermi energy, one can, in principle, reduce the charge current to zero, creating a spin Hall angle much larger than one. Another possibility is to have a non-local setup so that the spin current is separated from the charge current. For instance, the nanomagnet can sit to the left of the source contact. Then the reflected spins from the PN junction can move past the source and diffuse into the nanomagnet. In this setup, the charge current is minimized by the PN junction while the spin current is still robust, therefore achieving a larger spin hall angle. Those topics are outside the scope of this dissertation but we believe a tunable TI surface can be very useful in future spintronics applications.

3.5 Summary of Contributions

In this chapter, we explored one of the topological systems that can be used as a potential spin current generator. We have seen there are still fundamental physics we can study with the topological insulator. A better understanding of the spin transport in a TIPNJ setup will help pave the way to future spintronics applications. The contributions from this chapter are summarized below:

1. We have explored the spin collimation feature of the TIPNJ - a.k.a the Klein Tunneling (KT) through analytical as well as numerical analysis. We have proposed possible experimental measurements to confirm the KT physics in the TIPNJ and showed the expected signal in a quasi-analytical form.

2. We have analyzed the spin transport between a nanomagnet and the TI surface. We have clarified that the intrinsic spin amplification effect can only be converted to an impressive external spin source under certain conditions.

Chapter 4

Alternative 'Nanomagnet' - Magnetic Skyrmions

4.1 Magnetic Skyrmions in Thin Films

4.1.1 Magnetic Skyrmions as Information Carrier

Magnetic skyrmions have topological spin textures that are stabilized by the antisymmetric exchange or Dzyaloshinskii-Moriya interaction (DMI). Their vortex-like spin configurations were predicted to exist stably at the nanometer scale in bulk non-centrosymmetric materials as well as thin film heterostructures [68,69]. Substantial progress has been made in observing skyrmions and skyrmion lattices in multiple systems by means of neutron scattering, Lorentz transmision electron microscopy, scanning tunneling microscopy and X-ray holography. Reported material systems supporting skyrmions include the B20 family (such as MnSi [70,71], FeGe [72], $Fe_{0.5}Co_{0.5}Si$ [73]), multiferroic materials [74], tetragonal inverse heuslers [75], thin film Fe/Ir [76, 77], FeCoB/Pt [78], and amorphous ferrimagnets such as $Gd_{44}Co_{56}$ [79]. These exciting results bring up possible applications of skyrmions in reliable high density information processing and storage, such as racetrack memory, where information is stored in magnetic domains and driven by a current, as has been demonstrated in magnetic nanowires [80].One of the challenges with racetrack memory is the pinning of domain wall at defect sites. Skyrmions can be manipulated by spin transfer torque or spin-orbit torque

much like the magnetic domains, with the added advantages of smaller size and some amount of topological protection, reducing the threshold current for activation around defects [81]. However, the reported skyrmions are either too big ($\sim 100\,\mathrm{nm}$) or exist only at low temperatures [77]. Another issue is that the magnus force (from the chiral spin structure of skyrmion) swivels the skyrmions away from a linear trajectory towards the device edges for potential annihilation.

4.1.2 Skyrmion Type

The antisymmetric Dzyaloshinskii-Moriya interaction has the following general form (in the continuous limit) [82]:

$$M_i \frac{\partial M_j}{\partial \xi} - M_j \frac{\partial M_i}{\partial \xi} \tag{4.1}$$

where $M_{i,j}$ are the components of the magnetic moment $\mathbf{M} = (M_x, M_y, M_z)$ and ξ is a spacial coordinate (x, y, z). With different crystal symmetries, various kinds of magnetic skyrmions can be stablized (Neel skyrmion, Bloch skyrmion, and Antiskyrmion).

$$\begin{aligned}
E_{\mathrm{dm}}^{\mathrm{Neel}} &= D\mathbf{m} \cdot [(\hat{\mathbf{z}} \times \vec{\nabla}) \times \mathbf{m})] = \left(m_x \frac{\partial m_z}{\partial x} - m_z \frac{\partial m_x}{\partial x}\right) + \left(m_y \frac{\partial m_z}{\partial y} - m_z \frac{\partial m_y}{\partial y}\right) \\
E_{\mathrm{dm}}^{\mathrm{Bloch}} &= \left(m_z \frac{\partial m_y}{\partial x} - m_y \frac{\partial m_z}{\partial x}\right) + \left(m_x \frac{\partial m_z}{\partial y} - m_z \frac{\partial m_x}{\partial y}\right) + \left(m_y \frac{\partial m_x}{\partial z} - m_x \frac{\partial m_y}{\partial z}\right) \\
E_{\mathrm{dm}}^{\mathrm{Anti}} &= m_z \frac{\partial m_x}{\partial y} - m_x \frac{\partial m_z}{\partial y} + m_z \frac{\partial m_y}{\partial x} - m_y \frac{\partial m_z}{\partial x}
\end{aligned}$$

Fig. 4.1 shows the spin textures of those three types of skyrmions. Neel type skyrmions commonly exist from the broken inversion symmetry at the interface in heterostructures consisting of a thin magnetic film and a heavy metal over/under-layer. Bloch type skyrmions exist in non-centrosymmetric bulk materials such as B20 compounds (MnSi, FeGe, MnGe, and etc.). Anti-skyrmions exist in systems with D_{2d} symmetry e.g. tetragonal inverse heuslers [75]. Our studies focus on the Neel type skyrmion in thin films.

4.1.3 Skyrmion Hamiltonian

Under constant external field and in the absence of current driven motion, the skyrmion state is stablized by competing interactions from the exchange, Dzyaloshinskii-Moriya, anisotropy,

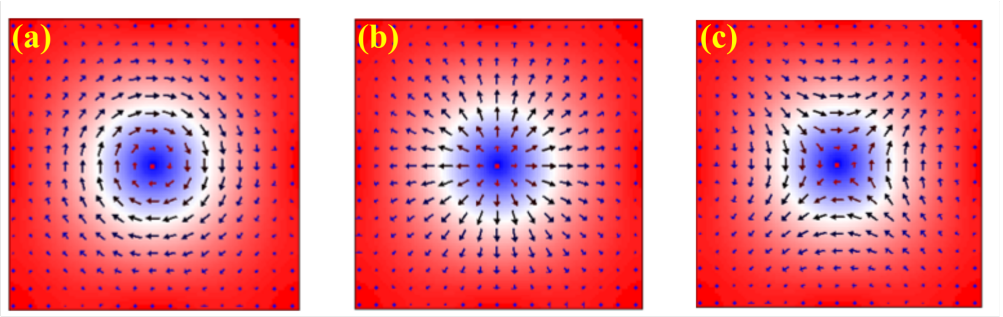


Figure 4.1: Different types of magnetic skyrmion spin texture (top view). The arrows show local spin direction and the color shows the z component of the spin. (a) Bloch type skyrmions. (b) Neel type (Hedgehog) skyrmion. (c) Anti-skyrmions. The figures are generated with public code OOMMF [5].

stray field, and external magnetic field. In the continuous limit, the energy of an isolated skyrmion in a thin film can be written as:

$$E = t \iint \Big\{ A\left(\boldsymbol{\nabla}\mathbf{m}\right)^2 - K m_z^2 - \mu_0 M_s \mathbf{m}\cdot\mathbf{H}_{\text{ext}} - \frac{1}{2}\mu_0 M_s \mathbf{m}\cdot\mathbf{H}_{\text{d}} + D\mathbf{m}\cdot[(\hat{\mathbf{z}}\times\vec{\nabla})\times\mathbf{m}] \Big\} d^2\mathbf{r} \tag{4.2}$$

with t the film thickness, A the exchange stiffness, K the uniaxial anisotropy, $\mathbf{H}_{\text{ext}}$ the external magnetic field and $\mathbf{H}_{\text{d}}$ the stray field. The last term describes the interfacial DMI characterized by the coefficient D. $\hat{\mathbf{z}}$ is the unit vector normal to the interface between the magnetic film and heavy metal underlayer/overlayer.

All types of skyrmion have an associated integer winding number $N_{\text{sk}} = \pm 1$ which is calculated from its spin texture:

$$N_{\text{sk}} = \frac{1}{4\pi}\int \mathbf{m}\cdot\left(\frac{\partial\mathbf{m}}{\partial x}\times\frac{\partial\mathbf{m}}{\partial y}\right) dxdy \tag{4.3}$$

4.1.4 Isolated Skyrmion versus Skyrmion Lattice

The solution to Eqn 4.2 can result in isolated skyrmion, skyrmion lattice, or Neel stripe phases depending on the parameter range. Energetically speaking, skyrmion lattice is the ground state when DMI is large because it is the global energy minimum compared

to the ferromagnetic state. On the other hand, the isolated skyrmion state stabilizes at a local energy minimum as we will see in the following sections. From an application perspective, isolated skyrmion is more favorable because it allows individual control (creation and annihilation) of a single bit of information.

4.2 Analytical Model for Skyrmion Size and Energy

4.2.1 2π Model for Skyrmion

To gain insights into how to create small and stable skyrmions. We derive analytical solutions to Eqn. 4.2 by adopting a parameterized skyrmion model - the skyrmion 2π model. It is a variation of the 2π domain wall model in the one-dimensional case.

One-dimensional π domain wall model

In one-dimensional Neel type domain walls (Fig. 4.2), the magnetic moments rotate in the $x-z$ plane with the rotation angle $\theta = \theta(x)$ as a function of the position along the x axis. In the absence of the external field and DMI, the total energy from Hamiltonian 4.2 gives:

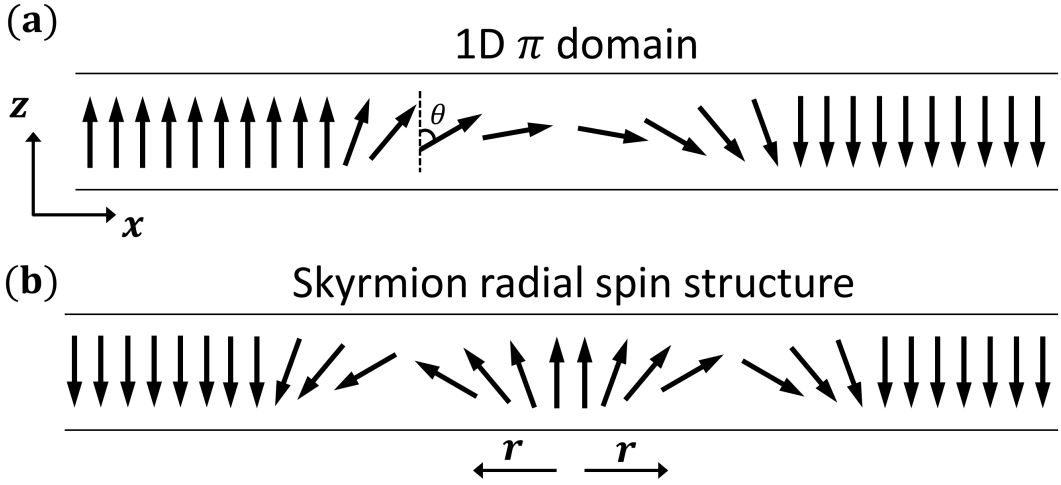


Figure 4.2: (a) The magnetic structure of the 1D Neel type π domain wall (side view). (b) The radial spin structure of a Neel type skyrmion. The arrows represent the local magnetic moments.

$$\epsilon[\theta(x)] = \int_{-\infty}^{\infty} \left[A \left(\frac{d\theta}{dx} \right)^2 - K \cos^2 \theta \right] dx \tag{4.4}$$

Using standard variation calculus to minimize the energy with respect to the shape function $\theta(x)$ gives the corresponding Euler equations:

$$\begin{aligned} \frac{d^2\theta}{dx^2} &= \frac{\sin\theta\cos\theta}{\Delta_0^2} \\ \frac{d\theta}{dx} &= 0 \qquad \text{when } x = \pm\infty \end{aligned} \tag{4.5}$$

where $\Delta_0 = \sqrt{A/K}$ is the wall width. Integrating Eq.4.5 gives us the shape function of a π domain wall:

$$\theta_\pi = 2\arctan\left[e^{\frac{x-x_0}{\Delta_0}}\right] \tag{4.6}$$

where x_0 is the center point of the 1D wall. Fig.4.3(a) shows an example solution. As shown in Fig. 4.2(b), the π domain wall spin texture resembles the radial profile of the skyrmion $\theta(r)$ if one replaces $x - x_0$ with $r - R$, where R is the radius of the skyrmion. It turns out the π model is a good approximation for larger skyrmion as long as $R \gg \Delta_0$. For $R \leq \Delta_0$ (small skyrmions), π model fails to capture the center of the skyrmion as $\theta(r = 0) \neq 0$ or π as illustrated in Fig.4.3(b). A more suitable shape function is the 2π domain wall model, which is a combination of two π models:

$$\begin{aligned} \theta_{\text{sk}}(r) &= 2\arctan\left[e^{\frac{-r+r_0}{\Delta}}\right] + 2\arctan\left[e^{\frac{-r-r_0}{\Delta}}\right] \\ &= 2\arctan\left(\frac{\cosh\frac{r_0}{\Delta}}{\sinh\frac{r}{\Delta}}\right) \rightarrow 2\arctan\left(\frac{\sinh\frac{R}{\Delta}}{\sinh\frac{r}{\Delta}}\right) \end{aligned} \tag{4.7}$$

in the last step, we changed the form a little bit by replacing $\cosh\frac{r_0}{\Delta}$ with $\sinh\frac{R}{\Delta}$ so that the parameter R agrees with the skyrmion radius definition we used $\theta(r = R) = \pi/2$. It can be shown this shape function guarantees $\theta(r = 0) = \pi,\ \theta(r = \infty) = 0$ as in Fig.4.4 (b) for both small and large skyrmions.

Adding one extra parameter - the azimuthal angle ψ (Fig.4.4 (a)(c)), one obtains a more general skyrmion shape function which can be used to describe the Neel type, Bloch type, or a mixed type skyrmion (corresponding to $\psi = 0, \pi$ for pure Neel type skyrmions, $\psi = \pm\pi/2$ for pure Bloch skyrmions, and any other value of ψ for mixed type).

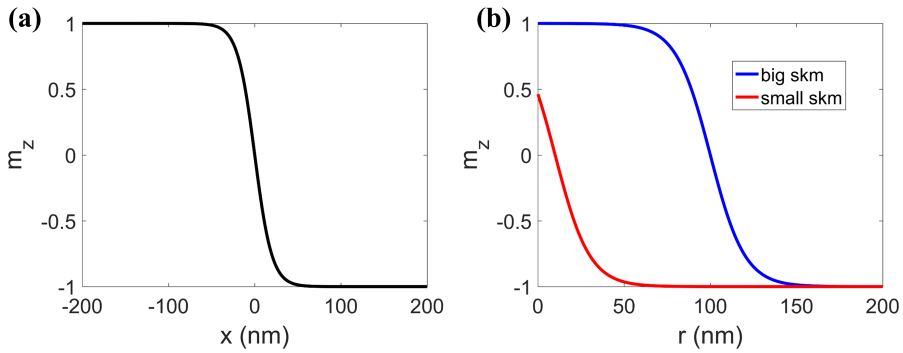


Figure 4.3: (a) 1D domain wall spin profile with $x_0 = 0$. (b) Using 1D π model as the radial function to describe different skyrmion sizes.

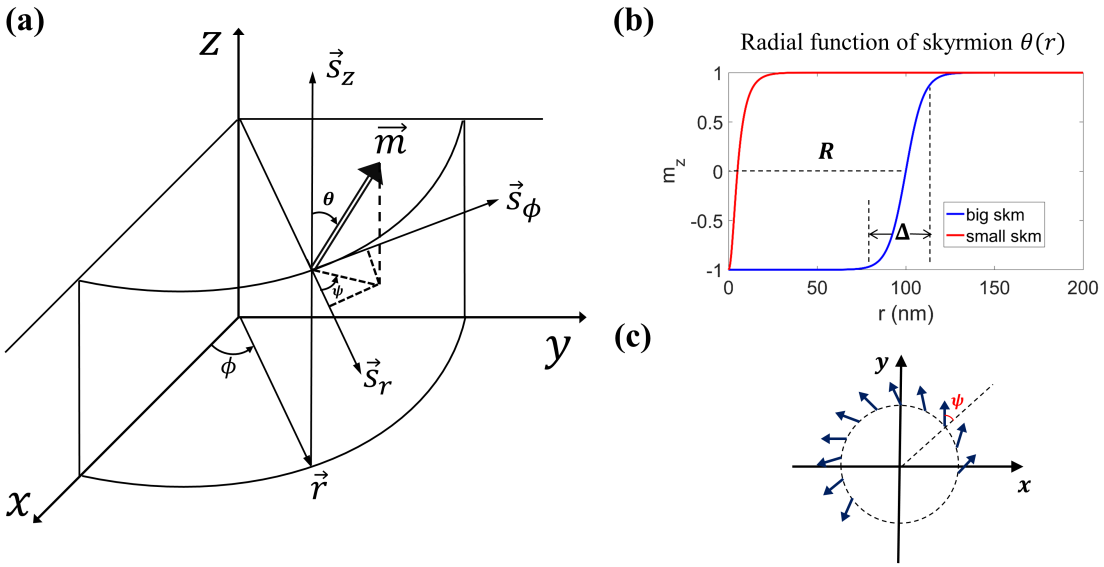


Figure 4.4: (a) Spacial coordinates in cylindrical coordinates and magnetic moment in spherical coordinates. (b) Radial function of the skyrmion with parameters R and Δ. (c) Top view of the azimuthal parameter of the spin moment.

skyrmion energetics

Unlike the π model for 1D domain wall, the shape function $\theta_{\mathrm{sk}}(r)$ is a trial function rather than a solution. To obtain the parameters R and Δ, we substitute the shape function Eq. 4.7 into Eq.4.2 to minimize the skyrmion energy $E_{\mathrm{sk}}(A, D, K, t)$ for a given set of material parameters. In the following paragraphs, we write down each energy term in the cylindrical coordinates (Fig. 4.4(a)) and then fit them with simpler functions to obtain an analytical equation for skyrmion size in the small skyrmion regime.

Exchange energy

$$
\begin{aligned}
E_{\text{ex}} &= 2\pi t A \int_0^\infty \left\{ \left(\frac{d\theta}{dr}\right)^2 + \frac{1}{r^2}\left(\frac{d\theta}{d\phi}\right)^2 + \sin^2\theta \left[\left(\frac{d(\psi+\phi)}{dr}\right)^2 + \frac{1}{r^2}\left(\frac{d(\psi+\phi)}{d\phi}\right)^2 \right] \right\} r dr \\
&= 2\pi t A \int_0^\infty \left\{ \left(\frac{d\theta}{dr}\right)^2 + \frac{\sin^2\theta}{r^2} \right\} r dr \\
&= 4\pi t A \sinh^2\left(\frac{R}{\Delta}\right) \left[\frac{1}{\Delta^2} \int_0^\infty \frac{\cosh\frac{2r}{\Delta}+1}{\left(\sinh^2\frac{r}{\Delta}+\sinh^2\frac{R}{\Delta}\right)^2} r dr \right. \\
&\qquad \left. + \int_0^\infty \frac{\cosh\frac{2r}{\Delta}-1}{\left(\sinh^2\frac{r}{\Delta}+\sinh^2\frac{R}{\Delta}\right)^2} \frac{dr}{r} \right] \\
&= 4\pi t A \sinh^2\left(\frac{R}{\Delta}\right) \left[\int_0^\infty \frac{\cosh 2x+1}{\left(\sinh^2 x+\sinh^2\frac{R}{\Delta}\right)^2} x dx + \int_0^\infty \frac{\cosh 2x-1}{\left(\sinh^2 x+\sinh^2\frac{R}{\Delta}\right)^2} \frac{dx}{x} \right]
\end{aligned}
\tag{4.8}
$$

DMI energy

$$
\begin{aligned}
E_{\text{DMI}} &= 2\pi t D \int_0^\infty \left[\cos\psi \frac{d\theta}{dr} - \sin\theta\cos\theta\sin\psi \frac{d(\psi+\phi)}{dr} \right. \\
&\qquad \left. - \frac{1}{r}\sin\psi\frac{d\theta}{d\phi} + \frac{1}{r}\sin\theta\cos\theta\cos\psi\frac{d(\psi+\phi)}{d\phi} \right] r dr \\
&= 2\pi t D \int_0^\infty \left[\cos\psi\frac{d\theta}{dr} + \frac{1}{r}\sin\theta\cos\theta\cos\psi \right] r dr \\
&= 4\pi t D \sinh\left(\frac{R}{\Delta}\right)\cos\psi \left[-\frac{1}{\Delta}\int_0^\infty \frac{\cosh\frac{r}{\Delta}}{\sinh^2\frac{r}{\Delta}+\sinh^2\frac{R}{\Delta}} r dr \right. \\
&\qquad \left. + \int_0^\infty \frac{\sinh\frac{r}{\Delta}(\sinh^2\frac{r}{\Delta}-\sinh^2\frac{R}{\Delta})}{(\sinh^2\frac{r}{\Delta}+\sinh^2\frac{R}{\Delta})^2} dr \right] \\
&= 4\pi t D \Delta \sinh\left(\frac{R}{\Delta}\right)\cos\psi \left[-\int_0^\infty \frac{\cosh x}{\sinh^2 x+\sinh^2\frac{R}{\Delta}} x dx \right. \\
&\qquad \left. + \int_0^\infty \frac{\sinh x(\sinh^2 x-\sinh^2\frac{R}{\Delta})}{(\sinh^2 x+\sinh^2\frac{R}{\Delta})^2} dx \right]
\end{aligned}
\tag{4.9}
$$

Anisotropy

The uniaxial anisotropy (assumed in the z direction perpendicular to the film) energy is given by:

$$\begin{aligned} E_{\text{an}} &= 2\pi t K \int_0^\infty \sin^2\theta r dr \\ &= 4\pi t K \sinh^2\left(\frac{R}{\Delta}\right) \int_0^\infty \frac{\cosh\left(\frac{2r}{\Delta}\right) - 1}{\left[\sinh^2\left(\frac{r}{\Delta}\right) + \sinh^2\left(\frac{R}{\Delta}\right)\right]^2} r dr \\ &= 4\pi t K \Delta^2 \sinh^2\left(\frac{R}{\Delta}\right) \int_0^\infty \frac{\cosh\left(2x\right) - 1}{\left[\sinh^2\left(x\right) + \sinh^2\left(\frac{R}{\Delta}\right)\right]^2} x dx \end{aligned}$$

Zeeman energy

The Zeeman energy from the external magnetic field perpendicular to the film is evaluated as:

$$\begin{aligned} E_{\text{hext}} &= 2\pi t \mu_0 M_s H_z \int_0^\infty (1 - \cos\theta) r dr \\ &= 4\pi t \mu_0 M_s H_z \sinh^2\left(\frac{R}{\Delta}\right) \int_0^\infty \frac{1}{\sinh^2\left(\frac{r}{\Delta}\right) + \sinh^2\left(\frac{R}{\Delta}\right)} r dr \\ &= 4\pi t \mu_0 M_s H_z \Delta^2 \sinh^2\left(\frac{R}{\Delta}\right) \int_0^\infty \frac{1}{\sinh^2\left(x\right) + \sinh^2\left(\frac{R}{\Delta}\right)} x dx \end{aligned}$$

Demagnetization energy

The demagnetization energy from the stray field is the most complex term due to the non-local nature of the interaction. It has non-trivial dependencies on the parameters t, R, Δ. In appendix A, we have derived the magnetostatic energy for isolated skyrmions and expressed it in a 2D integration form. Although the demagnetization energy can be fitted with a simpler equation as well, here we borrow the result for an infinite ferromagnetic film, where the stray field energy is only a function of the saturation magnetization M_s. With this approximation, the demagnetization energy can be included in the uniaxial anisotropy, which forms a smaller effective anisotropy. Our numerical simulations show that this is a

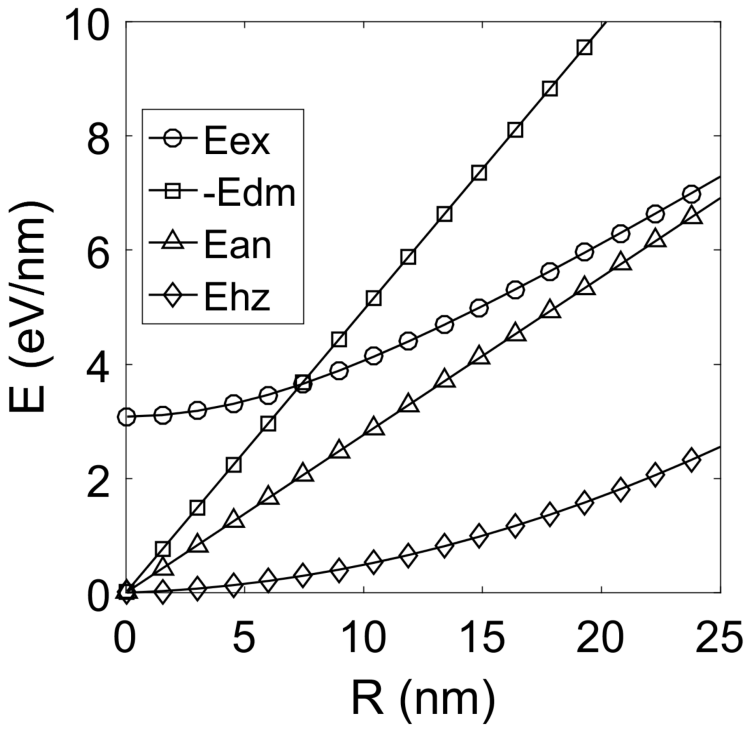


Figure 4.5: skyrmion energies as a function of R with fixed Δ. Simulation parameters $M_s = 1\text{e}6\,\text{A/m},\ K = 1.26\text{e}5\,\text{J/m}^3,\ D = 1.8\,\text{mJ/m}^2,\ A = 2\text{e} - 11\,\text{J/m},\ t = 2\,\text{nm}$.

relatively good approximation as long as the skyrmions size and the film thickness are small.

$$K_{\text{eff}} = K - \frac{\mu_0 M_s^2}{2} \tag{4.10}$$

Fig. 4.5(a) shows how each energy term depend on the skyrmion radius R when Δ is fixed. It is easy to see the role of DMI in stabilizing the skyrmion state since it is the only energy term that has a negative coefficient. Without the DMI, the skyrmion will shrink and eventually melt into the ferromagnetic state. It is worth mentioning that this statement is only true for smaller skyrmions. When the accurate demagnetization energy is included in the calculation, larger skyrmions can exist from the stray field without DMI but those skyrmions will have Bloch type or mixed type spin texture.

Analytical results for small skyrmions

In the small skyrmion regime ($R \sim \Delta$), we fit each energy term with a simple analytical function to obtain an analytical result for the skyrmion size and energy. In the absence of

any external field (Zeeman term can be fitted with a quadratic equation too):

$$\begin{aligned} E_{\text{ex}} &\approx tAC_{e1}\sqrt{C_{e2}^2 + R^2/\Delta^2} \\ E_{\text{DMI}} &\approx -tDC_dR \\ E_{\text{an}} &\approx tKC_aR\Delta \\ E_{\text{sk}} &= E_{\text{ex}} + E_{\text{ex}} + E_{\text{an}} \end{aligned} \tag{4.11}$$

where C_{e1}, C_{e2}, C_d, C_a are fitted dimensionless constants and their values are given in Table 4.1.

Table 4.1: Fitted constants

const.	C_{e1}	C_{e2}	C_d	C_a
value	11.92	2.07	19.80	12.48

Minimizing skyrmion total energy with respect to R and Δ gives the metastable skyrmion state. To make the algebra easier, we can replace the variable R, Δ with $s = R/\Delta$, Δ and make the derivatives equal to zero:

$$\begin{aligned} \frac{1}{t}\frac{\partial E_{\text{sk}}}{\partial \Delta} &= -C_dDs + 2C_aKs\Delta = 0 \\ \frac{1}{t}\frac{\partial E_{\text{sk}}}{t\partial s} &= C_{e1}A\frac{s}{\sqrt{s^2 + C_{e2}^2}} - C_dD\Delta + C_aK\Delta^2 = 0 \end{aligned}$$

which satisfies:

$$\begin{aligned} \Delta &= \frac{C_dD}{2C_aK} \\ u = \frac{R}{\sqrt{R^2 + C_{e2}^2\Delta^2}} &= \frac{C_d^2D^2}{4C_aC_{e1}KA} < 1 \\ R_{sk} &= \frac{C_{e2}\Delta}{\sqrt{u^{-2} - 1}} \end{aligned} \tag{4.12}$$

for $u \geq 1$, there is no local energy minimum and thus an isolated skyrmion is not stable. We can rewrite Eq. 4.13 to a similar form as the skyrmion radius equation from the literature [83]:

$$R_{sk}=\left(\frac{D}{D_c}\right)^3 \frac{2.63\Delta_0}{\sqrt{1-\left(\frac{D}{D_c}\right)^4}} \tag{4.13}$$

$$R_{sk}=\frac{\Delta_0}{\sqrt{2(1-D/D_c)}} \qquad \text{in literature} \tag{4.14}$$

where $\Delta_0 = \sqrt{A/K}$ is the 1D domain wall width and $D_c = 4\sqrt{AK}/\pi$ is the critical DMI. Substituting the above solutions into the fitted equations and sum over all energy terms we can get

$$E_{min}\left(R_{sk}\right) = \frac{AC_{e1}}{\Delta}\sqrt{R_{sk}^2 + C_{e2}^2\Delta^2} - \frac{C_d D R_{sk}}{2} \tag{4.15}$$

Although the Zeeman energy term was not included in the derivation, we found that simply replacing K with $K + H_z M_s$ gives good approximation for skyrmion size under small external magnetic field as shown in Fig. 4.6b.

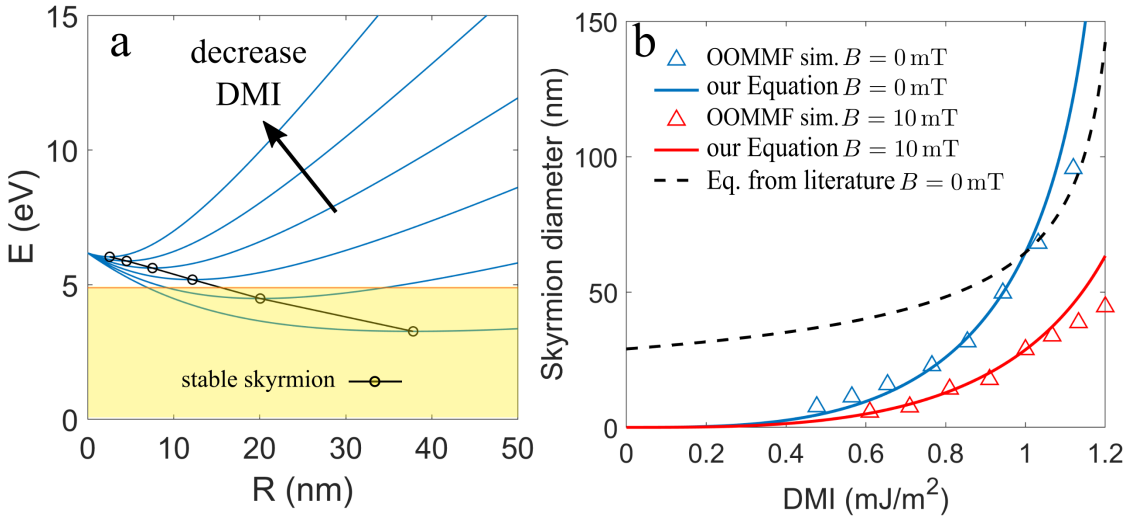


Figure 4.6: (a) Minimizing skyrmion energy for different DMI values. (b) Compare our analytical equation, numerical simulation, and the equation used in the literature for skyrmion size.

skyrmion size versus skyrmion stability

In Fig. 4.6(a), we see that the skyrmion size reduces as DMI decreases. However, as the skyrmion size reduces, the energy barrier separating the skyrmion phase and the ferromagnetic phase ($E = 0$) also diminishes, rendering the skyrmion unstable against thermal fluctuations at room temperature. Thus there is a narrow window of phase space for a stable small isolated skyrmion. The required lifetime of a skyrmion is application specific. Here we consider a minimum barrier of $50k_BT$ to be stable in the context of conventional memory/storage applications. This thermal stability constraint puts a lower limit on the skyrmion size. It is worth distinguishing the minimum skyrmion radius discussed here from the minimum radius due to the discrete nature of the atomic sites discussed in the literature [84]. The minimum skyrmion radius with adequate thermal stability can be quite large (10s of nanometers), while the minimum radius due to the discretization is very small ($< 1\,\text{nm}$). The thermal stability-constrained minimum radius depends on the material parameters as well as the film thickness. However, the discretization of spin lattice also affects the thermal stability because it modifies the energy barrier. In a continuous model, the top of the barrier ($E(R \to 0)$ in Fig. 4.6) is approximately $27.3At$. In a discrete lattice, the barrier top reduces to $\sim 23At$ in a certain discrete lattice structure [85]. This reduction of barrier can have dramatic effect on the stability of skyrmions because At can be much larger than k_BT. When we evaluate the skyrmion stability in real material systems (sec.4.3), we subtract $2At$ from the analytical model estimate to partially account for this barrier reduction due to discretization.

In Fig. 4.7, the energy barrier as a function of the skyrmion radius is plotted against different DMI and anisotropy constant K. It is important to point out that tuning individual material parameter independently might not be the optimal solution and the optimization strategy can depend on the objective: e.g. achieving the smallest skyrmion versus achieving an optimal phase space under constraints such as limited D and K. These trade-offs become clearer in sec4.3 when we examine the small skyrmion space for the Inverse Heuslers.

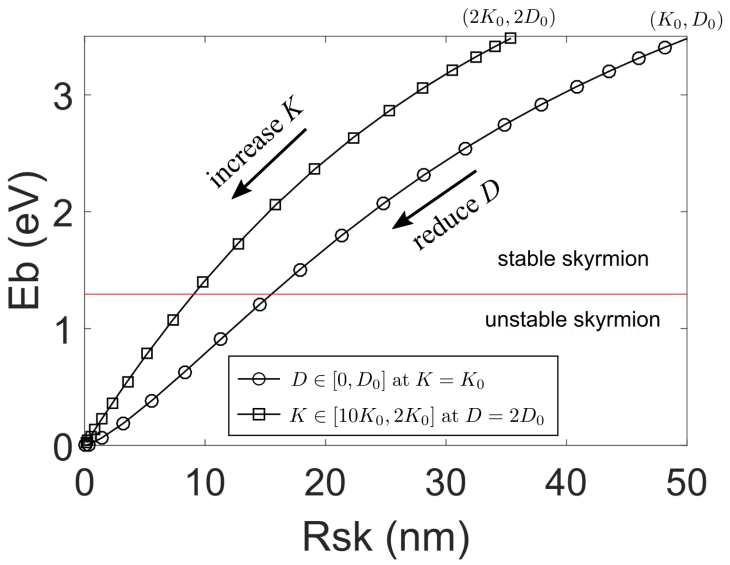


Figure 4.7: The energy barrier separating the skyrmion phase and the ferromagnetic phase as a function of skyrmion radius. The skyrmion radius is tuned by varying D or K. Reducing D alone reduces the skyrmion size. However, the minimum radius also depends on other parameters described earlier. If the objective is to reach the smallest skyrmion at a fixed energy barrier, we need to maximize (instead of minimize) D while adjusting other parameters such as K.

4.2.2 Current-Driven Skyrmion Mobility

Although our main focus is on the equilibrium properties of isolated skyrmions, it is worth reviewing some skyrmion dynamics to gain insights into what kind of materials we are looking for. Skyrmions can be driven by the spin-orbit torque created from an in-plane current through the heavy metal unlayer, as has been demonstrated in multiple experiments [78,86,87]. In the high speed regime, the skyrmion mobility Eq.4.16 is determined by the skyrmion size and net magnetization. Given the size constraint, lowering the saturation magnetization is critical in improving the skyrmion mobility. One key concern related to the current-driven skyrmion motion is the skyrmion Hall effect (SKHE). Some potential solutions based on film edge engineering or material composition engineering have been proposed [88, 89]. The skyrmion Hall angle is directly related to the skyrmion winding number (Eq. 4.17), which takes integer numbers ($N = \pm 1$ for single ferromagnetic skyrmions). One way to reduce the SKHE is to have two coupled skyrmions with opposite winding numbers. For example, in a synthetic antiferromagnetic structure, the top and bottom layers each host a skyrmion but

with mirrored spin texture [90, 91]. An alternative is to use ferrimagnet or antiferromagnet where the spin sublattices are naturally coupled. The mobility of the coupled skyrmions in a ferrimagnet can be derived from the Thiele equation [92, 93]:

$$\mu = \frac{\gamma\hbar 4\pi}{2e} \frac{1}{\sqrt{(4\pi)^2 \frac{\langle N\rangle^2}{\mathcal{D}_{xx}^2} + \alpha^2}} \frac{I}{\mathcal{D}_{xx}} \frac{\Sigma_i \theta_{sh_i}}{\Sigma_i t_i \ M_{s_i}} \tag{4.16}$$

with skyrmion hall angle:

$$\Theta = \tan^{-1}\left(\frac{4\pi\langle N\rangle}{\alpha\mathcal{D}_{xx}}\right) \tag{4.17}$$

with γ the gyromagnetic ratio, θ_{sh} the spin hall angle of the heavy metal layer, t the thickness, and e the electron charge. $\mathcal{D}_{xx} = \int(\frac{\partial m}{\partial x})^2 d^3r$, $\langle N\rangle = \Sigma_i N_i t_i M_{s_i}/\Sigma_i t_i M_{s_i}$, $I = \frac{1}{4}\int(\sin\theta\cos\theta + r\frac{d\theta}{dr})dr$. The summation is over spin sublattices and $N_i = \pm 1$ is the winding number of skyrmion for each sublattice. To increase the skyrmion mobility, one should reduce the four parameters $\langle N\rangle, \alpha, t, \overline{M}_s$, where $\overline{M}_s = \sum_i t_i M_{s_i}$ is the net magnetization. Reducing the film thickness t compromises the skyrmion stability. Reducing $\langle N\rangle$ and $\overline{M}_s$ imply that one should look for antiferromagnets or ferrimagnets with low net magnetization. At that point, low magnetic damping also becomes increasingly important when $\langle N\rangle$ becomes small. However, if $\alpha \ll \langle N\rangle$, the skyrmion Hall angle can actually increase (see Eq. 4.17). In fact, with ultra-small damping (as in the case of half-metals which we will discuss shortly) the skyrmion Hall angle is very close to 90°, meaning that the current direction and the skyrmion momentum is orthogonal. In that case, an alternative circuit where current is applied in an orthogonal direction might offer a different solution to the skyrmion Hall effect problem.

4.3 Small and Stable Skyrmion in Inverse-Heusler

The above-mentioned conditions for hosting small and fast skyrmion has inspired us to look for ferrimagnetics/antiferromagnets in the Heusler family. The selection criteria include ferrimagnetic order (exists in Inverse Heusler), high Neel temperature, and potentially half-metallic bandstructures. Half-metals are known to have low magnetic damping due to the lack of electron states in one spin channel [94, 95]. In the remainder of this chapter,

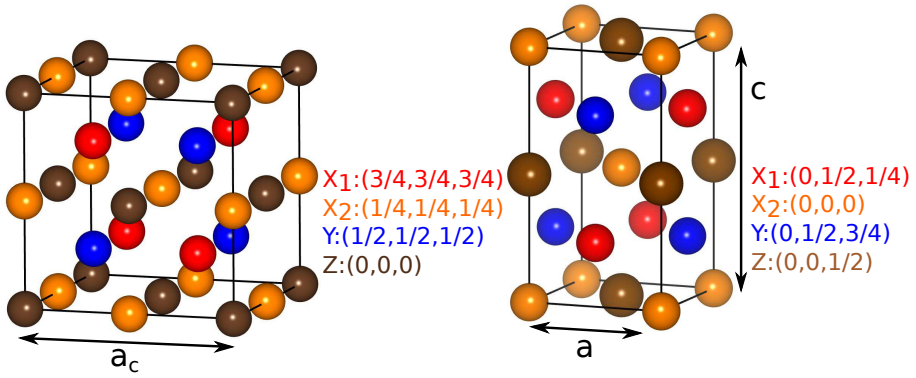


Figure 4.8: Schematic representation of (a) cubic Inverse-Heusler structure and (b) tetragonal Inverse-Heusler structure. X_1 and X_2 are the same transition metal element but they have different environments and magnetic moments. The tetragonal unit cell is rotated 45° around the z axis relative to the parent cubic structure. Figure generated by Jianhua Ma.

we present half-metallic and near half-metallic ferrimagnets from previous high-throughput studies on the inverse Heuslers [96], along with detailed calculations of their magnetic properties.

4.3.1 Ferrimagnetic Inverse Heuslers

Crystal structure

The Inverse-Heusler X_2YZ has a face-centered cubic structure with four atoms per cubic unit cell. Its crystal structure can be viewed as four interpenetrating FCC sublattices, occupied by the two X, Y and Z elements, respectively as shown in Fig. 4.8(a). We use X_1 and X_2 to distinguish these two X atoms sitting at the two nonequivalent sites. The tetragonal inverse-Heusler structure can be obtained by stretching or compressing the parent cubic structure along the z axis. We can define the tetragonality as c/a. The lattice constant a_c of the cubic structure can be obtained from a as $a_c = \sqrt{2}a$.

Atomistic exchange and Gilbert damping

My colleague Jianhua Ma has studied the fundamental electronic and magnetic properties of the Inverse Heusler familiy using the Vienna *Ab Initio* Simulation Package (VASP) [97,98] in his work [96]. We have identified multiple Inverse Heuslers with low saturation magnetization and half-metallic/near half-metallic bandstructure from the Heusler database. Jianhua has

Table 4.2: Structural and magnetic properties of Heusler compounds with low hull distance. Successive columns present: composition, calculated lattice constant, a_{Cal}, experiment lattice constant, a_{Exp}, saturation magnetization, M_S, formation energy ΔE_f, distance from the convex hull ΔE_{HD}, experiment Neel temperature, T_N(Exp), calculated Neel temperature T_N(Cal), calculated Gilbert damping at room temperature, α, and electronic ground state (Electronic ground state: M = nonmagnetic metal, HM = half-metal, NHM = Near half-metal). The last two columns show their potential tetragonal phase structure and energy difference from their cubic phase $\Delta E = E_{\mathrm{cubic}} - E_{\mathrm{tetragonal}}$.

	Cubic phase								Tetragonal phase	
XYZ	a_{Cal} (Å)	a_{Exp} (Å)	M_S (emu/cc)	ΔE_{HD} (eV/atom)	T_N(Exp) (K)	T_N(Cal) (K)	α (10^{-3})	Electronic ground state	a, c (Å)	ΔE (eV/atom)
Mn_2CoAl	5.735	5.798 [102]	393.5	0.036	720 [102]	845	4.04	HM	3.76, 6.68	-0.05
Mn_2CoGa	5.76	5.86 [103]	389.1	0	740 [104]	770	2.18	NHM	3.71, 7.13	-0.0103
Mn_2CoSi	5.621		627.3	0.018		460	3.01	HM		
Mn_2CoGe	5.75	5.80 [103]	590.6	0.03		471	4.97	HM	3.75, 6.84	0.0144
Mn_2FeAl	5.75		195.3	0.008		380	8.14	NHM	3.67, 7.28	0.0026
Mn_2FeGa	5.79		198.5	0.018		496	7.37	NHM	3.68, 7.29	0.0331
Mn_2FeSi	5.60		424	0		71	3.98	NHM	3.56, 7.26	-0.071
Mn_2FeGe	5.72		399.1	0		210	3.04	NHM	3.62, 7.45	-0.0164
Mn_2CuAl	5.89		33.4	0.042		1145	1.84	Metal		
Mn_2CuGa	5.937		57.86	0.0208		1242	1.59	Metal		

further calculated the exchange coupling and Gilbert damping via the framework of the Korringa-Kohn-Rostoker Greens function formalism, as implemented in the Munich spin-polarized relativistic Korringa-Kohn-Rostoker (SPR-KKR) package [99]. Relativistic effects were taken into account by solving the Dirac equation for the electronic states, and the atomic sphere approximation was employed for the shape of potentials. An angular momentum cutoff of $l_{max} = 3$ was used in the multiple-scattering expansion. To achieve convergence, we used the BROYDEN2 algorithm with the exchange-correlation potential of Vosko-Wilk-Nusair [100]. The atomistic exchange coupling J_{ij} based on the classical Heisenberg model are obtained from the KKR method using the Lichtenstein formula [101].

4.3.2 Skyrmion Phase Space for Ferrimagnetic Heuslers

We have down selected six compounds with relatively high Neel temperature and low saturation magnetization. Most of them have cubic structure as their ground states, which lack the perpendicular anisotropy for hosting Neel skyrmions. An interfacial anisotropy with a bottom layer is needed for interface-induced perpendicular anisotropy. We have also studied their possible tetragonal distortions that allow intrinsic anisotropy. Some of those tetragonal phases are metastable and may require special process to synthesize. Here we

treat the anisotropy as a variable and discuss the small skyrmion phase space. In Fig. 4.9, the smallest skyrmion given a fixed energy barrier E_b is plotted against different external fields and DMIs. Here we used an effective ferromagnetic model with an average exchange stiffness extrapolated from the Neel temperatures of those compounds. This approximation captures the energetics of the equilibrium skyrmion at 0K. The internal antiferromagnetic structure becomes important when spin dynamics is involved and the continuous model needs to be modified. The line separates the stable skyrmion phase (for a given energy barrier criterion) and the unstable skyrmion or ferromagnetic phase. The effects of different energy barriers (i.e. skyrmion lifetimes) and uniaxial anisotropy on the skyrmion size are also presented. It is also worth mentioning that even though the skyrmions are simulated within a continuous model (OOMMF), a small correction ($\delta E = 2At$) of the barrier top was added to account for the discrete nature of the atomic sites as mentioned in sec. 4.2.

In Fig.4.9, for a fixed barrier E_b and anisotropy Q, one common observation is that the smallest skyrmion always exists at larger B_z and DMI; increasing anisotropy helps reduce the skyrmion size but also requires a larger DMI. The need for larger DMI to achieve smaller skyrmion *for a given energy barrier* can be understood by revisiting the energy dependencies on the skyrmion radius in Fig. 4.5. Due to DMI being the only term with a negative slope with respect to R, the energy barrier is approximately associated with DMI and the skyrmion radius $dE \approx (D - K\Delta)R$. For a given dE, a larger DMI would allow smaller skyrmions if B_z or K are also increased concomitantly to push the energy minimum (i.e., skyrmion state) as close to $R = 0$ as possible. Otherwise for a fixed $K - B_z$ pair, a larger DMI would only result in a bigger skyrmion, as discussed in section 4.2. Another way to improve the skyrmion stability is to increase the film thickness, which scales all energies including the energy barrier. However, considering that DMI is an interfacial effect which decays rapidly away from the interface, it would be hard to maintain decent DMI in thicker films or to apply appreciable interfacial spin-orbit torque to drive the skyrmions. A potential solution is to use a multilayered structure to enhance the effective DMI.

4.4 Summary of Contributions

In this chapter, we looked at another type of topological excitation - magnetic skyrmions. They can be manipulated in a similar way as a nanomagnet (creation and annihilation) and magnetic domains (current-driven motion). They can be very useful for spintronics application ideas such as the racetrack memory. We have explored the conditions to achieve small and stable skyrmions. The summary of my contribution is below:

1. We have examined the energetics for isolated skyrmion phase and worked out a simple analytical equation for the skyrmion size that extends the existing equation in the literature to the ultra-small skyrmion regime.

2. We have discussed in detail about the conditions and trade-offs for achieving ultra-small and fast skyrmions. Especially, we have clarified the role of DMI in determining the size of a *stable* small skyrmion.

3. Based on our selection criteria, we have selected few Inverse-Heusler compounds for potential small skyrmion host systems. Their magnetic properties were studied and the minimum size stable skyrmion diagram is calculated for those alloys.

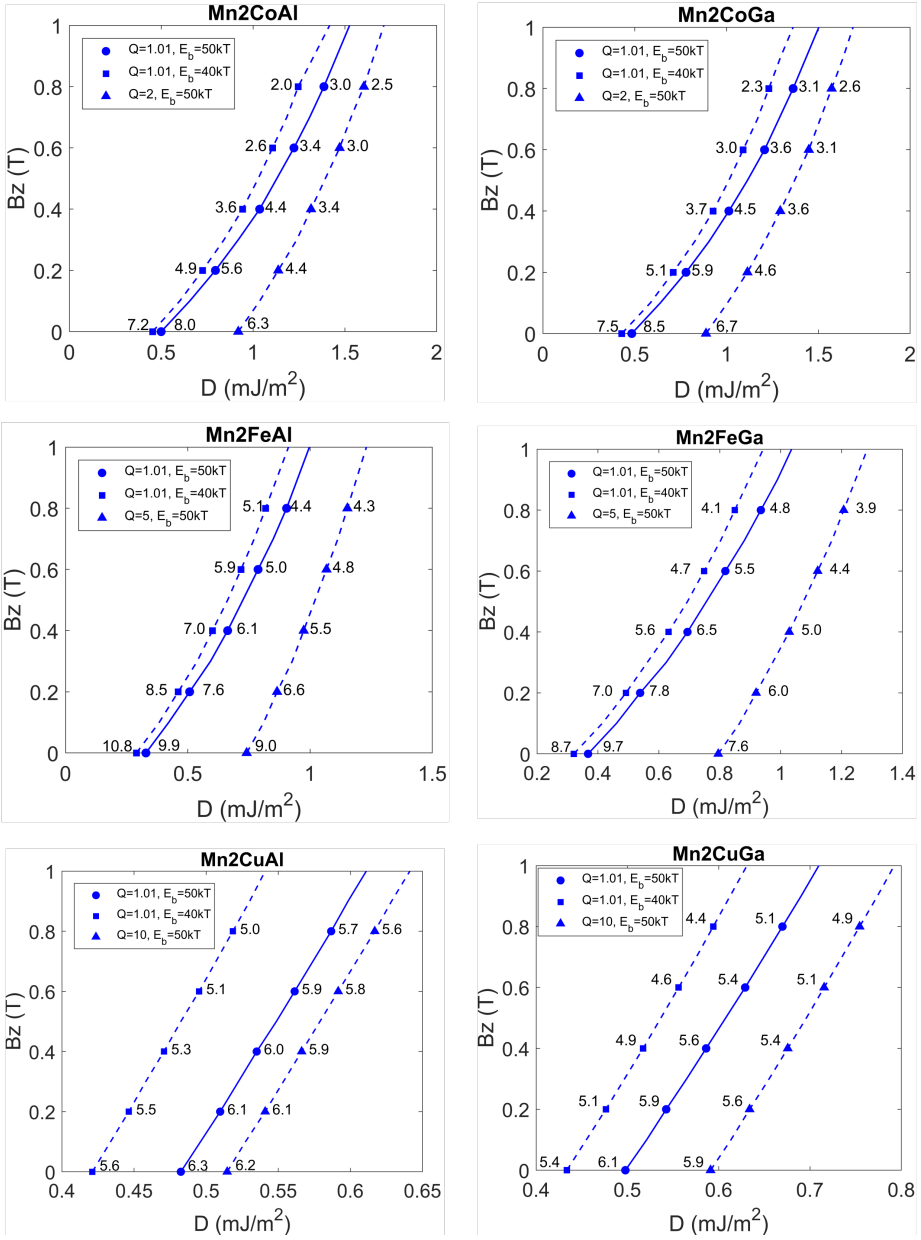


Figure 4.9: Smallest stable skyrmion boundary for Inverse Heuslers. The lines indicate the smallest stable skyrmion boundary with an energy barrier $E_b = 50k_BT$ or $E_b = 40k_BT$ between the ferromagnetic state and skyrmion state. The scatter plot samples the skyrmion size along the boundary. The film thickness is assumed 5 nm in all calculations. $Q = 2K/\mu_0 M_s^2$ defines the effective anisotropy compared to the demagnetization and $Q > 1$ is needed for a perpendicular anisotropy system. To show the effect of anisotropy on skyrmion size and stability, different Q are chosen for different materials due to varying saturation magnetization.

Chapter 5

Conclusion and Future work

The purpose of this dissertation is two-fold: First, through developing necessary tools in studying the Magnetic tunnel junction and write error in nanomagnet switching, we demonstrated the importance of a physics-based multiscale approach not only in refining our understanding of existing systems but also in exploring different new material systems and evaluating their performance limits when experiments are yet to be conducted. Second, even though the existing spintronics technology (mostly STT-MRAM) is fast approaching maturity and commercialization, it is still a long way to bring the spintronics applications to the main stage. We point to an emerging class of materials and topological excitations that could leverage their unique properties to push the boundaries of what spins can do. But a deeper understanding of them is much needed and this dissertation is just a starting point.

There are numerous challenges in exploring new materials and developing the necessary tools to study them. The challenges we encountered in our studies also translates to possible directions for future work. Here is an incomplete list of possible extensions of work:

Temperature dependent polarization of Half-metal based MTJs

Half-metals, with its theoretically 100% spin polarization, make perfect magnetic electrodes for the MTJs and our zero-temperature simulations seem to suggest the same. However, there are experimental evidence that shows the TMR reduces more dramatically with temperature in half-metal based MTJs. One possible explanation is the tilt of magnetic moment at the interface between the half-metallic layer and the tunnel barrier. The tilted moment destroys

the half-metallicity by mixing spin-up/down channels. Exactly how much does this affect the performance of an MTJ and whether it is universal in all half-metal based MTJs are unknown. One can study this effect in the first-principle NEGF framework: e.g. by tilting the magnetic moment at the interface from its equilibrium position and calculates the TMR and STT. These calculations can reveal how much degradation we can expect from the room temperature environment.

Magnet as a third terminal on TIPNJ

As discussed in sec.3.4, using a nanomagnet as the source for a TIPNJ can't utilize its impressive intrinsic gain. An alternative configuration is to have the magnet as a third contact whose chemical potential is fixed. In principle, there is a sweet spot where the charge current can be eliminated to maximize the gain. However, the drivability of this configuration doesn't only depend on the gain but also depend on the overall magnitude. How much total spin current can we expect from this configuration and whether it is difficult to control the potential of the third terminal are still unanswered.

Fokker-Planck Equation for skyrmion motion

Deriving the Fokker-Planck equation for skyrmion creation/annihilation can be difficult since those processes are highly incoherent and have a high degree of freedoms. However, after the skyrmion is created, it can be treated as a quasi-particle and its dynamics can be described by simple equation (Thiele equation) assuming its internal structure doesn't change. One could in principle derive the FPE for current drive skyrmion motion and study, for example, the pinning/unpinning probability near a defect.

Appendices

Appendix A

Demagnetization energy for thin film skyrmions

The magnetostatic energy from the demagnetization field (or stray field) for a single skyrmion in thin film comes from the dipole interaction of uncompensated magnetic charges on the surfaces as well as along the perimeter of the skyrmion in the case of a Neel type skyrmion. In a pure Bloch type skyrmion there are no net magnetic charges in the bulk. Fig. A.1 illustrates the magnetic charges for a single skyrmion in a thin film.

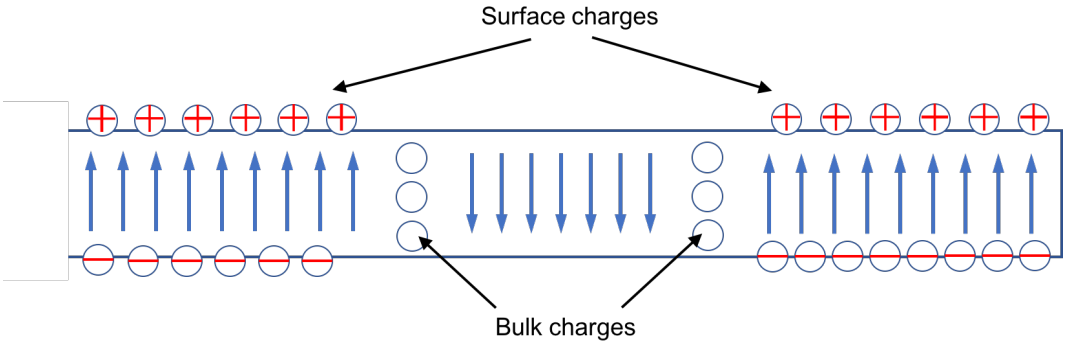


Figure A.1: schematic magnetic charges in Neel type skyrmion in thin films.

To evaluate the total magnetostatic energy E_{demag}, let's assume the thin film is infinitely wide with thickness t (from $z = -t/2$ to $z = t/2$). The total magnetostatic energy can be written as:

$$E_d = \frac{\mu_0 M_s^2}{4\pi} \int_0^\infty \int_0^\infty \int_{-t/2}^{t/2} \int_{-t/2}^{t/2} \int_0^{2\pi} \int_0^{2\pi} \frac{(\boldsymbol{\nabla} \cdot \mathbf{m})(\boldsymbol{\nabla'} \cdot \mathbf{m'})}{s} rr' d\phi d\phi' dz dz' dr dr' \tag{A.1}$$
$$s = \sqrt{r^2 + r'^2 - 2\cos\zeta + (z - z')^2}$$

where ζ is the angle between the two points $\rho = (r, \phi, z)$ and $\rho' = (r', \phi', z')$. $\boldsymbol{\nabla} \cdot \mathbf{m}$ is the magnetic charge density. We can divide the magnetic charges into bulk charges and surface charges $\mathbf{z} \cdot \mathbf{m}$ at $z = \pm t/2$ as shown in Fig.A.1. To evaluate the integral, we use the expansion of $1/|\rho - \rho'|$ in cynlindrical coordinates:

$$\frac{1}{s} = \frac{1}{|\rho - \rho'|} = \sum_{m=-\infty}^{\infty} e^{im(\phi - \phi')} \int_0^\infty \exp[-k(z_> - z_<)] J_m(kr) J_m(kr') dk \tag{A.2}$$

where $z_>/z_<$ is the larger/smaller one in z, z' and J_m is the mth Bessel function. Substitute A.2 to Eq. A.1 and integrate over ϕ, ϕ' first. we can break it down to E_{s-b}, E_{s-s}, E_{b-b} which are surface-bulk interaction, surface-surface interaction, and bulk-bulk interaction respectively. Due to the symmetry of the system, it can be shown that E_{s-b} evaluates to zero. For the bulk-bulk interaction E_{b-b}:

$$\begin{aligned}
E_d^{\mathrm{b-b}} &= \frac{\mu_0 M_s^2}{4\pi} \int_0^\infty \int_0^\infty \int_{-t/2}^{t/2} \int_{-t/2}^{t/2} \int_0^{2\pi} \int_0^{2\pi} \frac{(\boldsymbol{\nabla} \cdot \mathbf{m})(\boldsymbol{\nabla'} \cdot \mathbf{m'})}{s} rr' d\phi d\phi' dz dz' dr dr' \\
&= \pi \mu_0 M_s^2 \int_0^\infty \int_0^\infty (\boldsymbol{\nabla} \cdot \mathbf{m})(\boldsymbol{\nabla'} \cdot \mathbf{m'}) \, [\\
&\qquad \int_0^\infty \int_{-t/2}^{t/2} \int_{-t/2}^{t/2} \exp[-k(z_> - z_<)] J_0(kr) J_0(kr') dz dz' dk \Bigg] rr' dr dr' \\
&= \pi \mu_0 M_s^2 \int_0^\infty \int_0^\infty (\boldsymbol{\nabla} \cdot \mathbf{m})(\boldsymbol{\nabla'} \cdot \mathbf{m'}) \\
&\qquad \times \left[\int_0^\infty \frac{2}{k} \left[t + \frac{1}{k} e^{-kt} - \frac{1}{k} \right] J_0(kr) J_0(kr') dk \right] rr' dr dr' \\
&= 2\pi \mu_0 M_s^2 \int_0^\infty \frac{1}{k} \left[t + \frac{1}{k} e^{-kt} - \frac{1}{k} \right] \left[\int_0^\infty (\boldsymbol{\nabla} \cdot \mathbf{m}) J_0(kr) r dr \right]^2 dk \\
&= 2\pi \mu_0 M_s^2 \int_0^\infty \frac{1}{k} \left[t + \frac{1}{k} e^{-kt} - \frac{1}{k} \right] \left[\int_0^\infty \frac{\cos\psi}{r} \frac{d(r \sin\theta)}{dr} J_0(kr) r dr \right]^2 dk
\end{aligned} \tag{A.3}$$

Now consider the surface-surface ineraction, we replace $\boldsymbol{\nabla}' \cdot \mathbf{m}'$ with $\mathbf{n} \cdot \mathbf{m}' = \pm m'_z$ at $z = \pm t/2$ where $\mathbf{n}$ is the unit vector along z direction:

$$\begin{aligned} E_d^{\mathrm{s-s}} &= \frac{\mu_0 M_s^2}{4\pi} \int_0^\infty \int_0^\infty \int_{-t/2}^{t/2} \int_{-t/2}^{t/2} \int_0^{2\pi} \int_0^{2\pi} \frac{(\mathbf{n}\cdot\mathbf{m})(\mathbf{n}\cdot\mathbf{m})\delta(z \pm t/2)\delta(z' \pm t/2)}{s} \\ &\qquad \times rr' d\phi d\phi' dz dz' dr dr' \\ &= 2\pi\mu_0 M_s^2 \int_0^\infty \int_0^\infty m_z m'_z \int_0^\infty [1 - \exp(-kt)] J_0(kr) J_0(kr') dk rr' dr dr' \end{aligned} \tag{A.4}$$

It would be convenient to substract the demagnetization energy from the uniform FM state $m_z \equiv 1$. Homogeneous ferromagnetic film only has the surface part.

$$\begin{aligned} E_d^{\mathrm{surf}} &= 2\pi\mu_0 M_s^2 \int_0^\infty \int_0^\infty (1 - \cos\theta\cos\theta') \int_0^\infty [1 - \exp(-kt)] J_0(kr) J_0(kr') dk rr' dr dr' \\ &= 2\mu_0 M_s^2 \int_0^\infty \int_0^\infty (1 - \cos\theta\cos\theta') \sqrt{rr'}\, [k_0 K(k_0) - k_1 K(k_1)]\, dr dr' \end{aligned} \tag{A.5}$$

where K, E are complete elliptic integrals of the first/second kind and

$$k_0 = 2\frac{\sqrt{rr'}}{r + r'}, \qquad k_1 = 2\sqrt{\frac{rr'}{t^2 + (r + r')^2}} \tag{A.6}$$

and we used the following identities:

$$\begin{aligned} \int_0^\infty (1 - e^{-kt}) J_0(kr) J_0(kr') dk &= \frac{1}{\pi\sqrt{rr'}} [k_0 K(k_0) - k_1 K(k_1)] \\ \int_0^\infty J_1(kr) J_1(kr') k dk &= \frac{1}{r}\delta(r - r') \\ \int_0^\infty \int_0^\infty [1 - \exp(-kt)] J_0(kr) J_0(kr') r dk dr &= 0 \end{aligned} \tag{A.7}$$

Appendix B

Matlab code for Fokker-Planck solver

This is a starter code for running 2D Fokker-Planck simulation of nanomagnetic switching under spin transfer torque. Seven files are included here:

- **Main.m** - main file to run the simulation
- **Galerin.m** - construct coefficient matrix for the PDE based on the Galerkin's method
- **Torque.m** - specify the torque (anisotropy, external, STT, and so on) exerted on the magnetic moment. It can be modified to include other types of torque
- **parameter.m** - define all the physical paramters and constants
- **Initialize.m** - specifies the initial condition of the probability distribution
- **NonSwitchP.m** - calculates the time-dependent non-switching probability (or WER)
- **integral_W.m** - integrate the total probability on a triangular patch

Main.m:

```
% This program numerically solves the Fokker-Planck equation
% for nanomagnet under spin-torque.
% Author: Yunkun Xie
% Email: yx3ga@virginia.edu


%% Using the Distmesh (https://github.com/ionhandshaker/distmesh)
    program
```

```
% to generate triangle meshes triangle patches. This part of the code
    only
% needs to run once and can be reused if the mesh remains unchanged.
% fd=@(p) dsphere(p,0,0,0,1);
% [p,t]=distmeshsurface(fd,@huniform,0.1,1.1*[-1,-1,-1;1,1,1]);
% save('Data_p','p');
% save('Data_t','t');

function  Main(input)
tic;
t_record = cputime;
global glb_var;
glb_var = input;
Total_t =1e-9;
Delta_t =0.01e-9;
SAVE_DIR='./Data/';
disp('********************************');
disp('Start Fokker-Planck calculation.');
disp('********************************');

% Load the mesh
load('Data_p.mat');
load('Data_t.mat');

% Decide the number of nodes and triangles used in the calculation
Size_p = size(p);
Size_t = size(t);
num_node = Size_p(1);
num_tri = Size_t(1);

disp('************** Print out system info *****************')
disp(['Simulation time : ', num2str(Total_t),' s']);
[diffusion,~,~,~,~,~,~,~,~,~,~] = parameter(1);
disp('***************** End of system info  *****************')
% Calculate the coefficient matrix A ans B
disp('Creating coefficient matrix...');
[Matr_A,Matr_B] = Galerkin(num_node,num_tri,p,t);

% Initialize the probability vector W
e_file = exist([SAVE_DIR,'XXX.mat'],'file');
if e_file == 0
    W = Initialize(p,t);
else
    disp('start from existing file...')
    file = load([SAVE_DIR,'XXX.mat']);
    W = file.W_end;
end
disp('Probability distribution initialized.')
Num_step = round(Total_t / Delta_t);
Tau=0;

% Only store the chosen time points
Tau_total = Total_t * (0.0:0.02:1);
Data = zeros(num_node,size(Tau_total,2));
disp('Start time integration:')
Data(:,1) = W;
Init = 2;
for i = 1 : 1 : Num_step
    W = (Matr_B - Matr_A * Delta_t / 2) ...
        \ ((Matr_B + Matr_A * Delta_t / 2) * W);
    Tau=Tau+Delta_t;
    if abs(Tau - Tau_total(Init)) < Delta_t * 0.1;
        disp(['Data point:',num2str(Tau),' s']);
```

```
            Tau_total(Init)=Tau;
            Data(:,Init) = W;
            Init = Init + 1;
                    W_end = W;
        end
    end
    disp('Time integration finished.')

    %% Calculate the error rate
    NonSP=NonSwitchP(Data,t,p);
    spin = toc;
    save([SAVE_DIR,'test_perp_Isx_',num2str(input,'%1.2f'),'Ic0.mat'],'
        Tau_total','NonSP','W_end','Data');
    e_record = cputime - t_record;
    disp(e_record);
```

Galerkin.m:

```
function [A,B] = Galerkin( num_node,num_tri,p,t)
% Galerkin function calculates the PDE coefficient
% matrix A and B

A = sparse(num_node,num_node);
B = sparse(num_node,num_node);
[diffusion,~,~,~,~,~,~,~,~,~,~,~]=parameter(0);
for k = 1 : 1 : num_tri
    t_k = t(k,:);
    P1 = ( p(t_k(1),:) + p(t_k(2),:) + p(t_k(3),:))/3.;
    Xu = p(t_k(3),:) - p(t_k(1),:);
    Xv = p(t_k(2),:) - p(t_k(1),:);
    L = Torque( P1,Xu,Xv );
    gA = Xu * Xu';
    gB = Xu * Xv';
    gC = Xv * Xv';
    g_ij = [ gA , gB ; gB , gC ];
     for ith = 1 : 1 : 3
        for jth = 1 : 1 : 3
            i = t_k(ith);
            j = t_k(jth);
            [Sum_A,Sum_B] = Integral_A_B(i,j,t_k,L,g_ij,diffusion);
            A(i,j) = A(i,j) + Sum_A;
            B(i,j) = B(i,j) + Sum_B;
        end
    end
end
end

function [Sum_A,Sum_B] = Integral_A_B( i,j,t_k,L,g_ij,k)
% This function integrates A(i,j) over the triangle
% defined by t_k. The intergration is realized by sum
% over seven points on the triangle
G_ij = sqrt(det(g_ij));
g_IJ = inv(g_ij);
Sum_A = 0.;
Sum_B = 0.;

% evaluate the functions on the seven integration points
k1      = [ 1/3, 0.0597158717, 0.4701420641, 0.4701420641...
            , 0.7974269853, 0.1012865073, 0.1012865073 ];
k2      = [ 1/3, 0.4701420641, 0.0597158717, 0.4701420641...
            , 0.1012865073, 0.7974269853, 0.1012865073 ];
k3      = [ 1/3, 0.4701420641, 0.4701420641, 0.0597158717...
```

```
            , 0.1012865073, 0.1012865073, 0.7974269853 ];
weight = [ 0.225, 0.1323941527, 0.1323941527, 0.1323941527...
            , 0.1259391805, 0.1259391805, 0.1259391805 ];

for pp = 1 : 7
    Phi_i = fun_phi(i, t_k, k3(pp), k2(pp));
    Phi_j = fun_phi(j, t_k, k3(pp), k2(pp));
    Sum_B = Sum_B + weight(pp) * G_ij * Phi_i(3) * Phi_j(3);
    Sum_A = Sum_A + weight(pp) * G_ij *( Phi_i(1,1:2) * L' * Phi_j(1,3)
        ...
              - k * Phi_i(1,1:2) * g_IJ * Phi_j(1,1:2)');
end
end

function Phi = fun_phi(i,t_k,u,v)
% This function decides which Phi function should be used
for n = 1 : 1 : 3
    if i == t_k(n)
        Point_No = n;
    end
end

% Phi(1) : Differential of Phi with respect to u at (u,v)
% Phi(2) : Differential of Phi with respect to v at (u,v)
% Phi(3) : Phi at (u,v)
if Point_No == 1
    Phi(1) = -1.;
    Phi(2) = -1.;
    Phi(3) = 1 - u - v;
elseif Point_No == 2
    Phi(1) = 0;
    Phi(2) = 1.;
    Phi(3) = v;
elseif Point_No == 3
    Phi(1) = 1.;
    Phi(2) = 0;
    Phi(3) = u;
end
end
```

Torque.m:

```
function L = Torque( P1,Xu,Xv )
% This function calculates the torque vector L
% when M=P1

R = sqrt( P1(1)^2 + P1(2)^2 + P1(3)^2 ); % Scale the length to 1
P1 = P1 / R;

% Define constants
C=1.602e-19; % [coulombs] electron charge
u0=4e-7*pi; % N/A^2
hbar=1.054571628e-34; % Planck's constant
gamma=2.21e5;
mu_B=9.274e-24;

[~,~, alpha, V, Ms, s1, s2, H_ext, H_coeff, torque,~] = parameter(0);


M_Heff=H_ext + H_coeff .* P1;

L_inplane = 1/(1+alpha^2)*(-(alpha*gamma) * cross(P1,cross(P1,M_Heff))
    ... % Damping term
```

```
21          - (gamma) * cross(P1,M_Heff) ... % precession term
22           -(mu_B*s1/(V*Ms*C)+alpha*mu_B*s2/(V*Ms*C)) * cross(P1,cross(P1,
                 torque))... % Slonczski torque
23          +(alpha*mu_B*s1/(V*Ms*C)-mu_B*s2/(V*Ms*C)) * cross(P1,torque));%
                 Field-like torque
24
25
26 L = zeros(1,2);
27 [ u,v ] = Local_coordinate( Xu,Xv,L_inplane);
28 L(1) = u;
29 L(2) = v;
30 end
31
32 function [ u,v ] = Local_coordinate( Xu,Xv,P_uv )
33 % convert the torque vector into local coordinate (u, v)
34 % of the triangular patch
35
36 A = Xu * Xu';
37 B = Xu * Xv';
38 C = Xv * Xv';
39
40 fu = C / (A * C - B^2) * Xu - B / (A * C - B^2) * Xv;
41 fv = -B / (A * C - B^2) * Xu + A / (A * C - B^2) * Xv;
42
43 u = P_uv * fu';
44 v = P_uv * fv';
45
46 end
```

parameter.m:

```
1  function [diffusion, delta, alpha, V, Ms, s1, s2, H_ext, H_coeff, torque
       , M_init] = parameter(info)
2  % This file sets the parameters/configurations for
3  % the nanomagnet and the injecting current
4  global glb_var;
5  C=1.602e-19; % [coulombs] electron charge
6  u0=4e-7*pi; % N/A^2
7  hbar=1.054571628e-34; % Planck's constant
8  gamma=2.21e5;
9  mu_B=9.274e-24;
10 alpha = 0.01; % damping coefficient
11
12 % disk dimensions (in nm)
13 a=46;
14 b=46;
15 t=2;
16 V= (pi/4)*a*b*t*1e-27; % volume in m^3
17 area=(pi/4)*a*b*1e-18; % cross area
18 Ms=1.00e6; %saturation magnetization
19 kT=1.3806504e-23*300;
20 delta=40;
21 Hk=delta*2*kT/(u0*Ms*V);
22 K=(u0*Hk*Ms)/2;
23 pol=1.0;
24 Ic0=alpha*2*C/(pol*hbar)*(u0*Ms*V*Hk);
25 Current = Ic0 * glb_var;
26 s1 = glb_var * Ic0 * pol;
27 s2=0;
28
29 torque = [1,0,0]; % direction of the spin-polarized current
30 M_init = [0,0,1]; % Initial magnetization
31
32 % specify external magnetic field
```

```
H=400 * 10^3/(4*pi);
H_ext=[0,1,0]*H;

% For a uni-axial magnetic
Ndxx = 0; Ndyy = 0;
Ndzz = -2*K*V / (u0*Ms^2*V);

H_coeff = zeros(1,3);
H_coeff(1) = -Ms * Ndxx;
H_coeff(2) = -Ms * Ndyy;
H_coeff(3) = -Ms * Ndzz;

diffusion=alpha*kT*gamma/((1+alpha^2)*u0*Ms*V);

if info == 1
    disp(['Damping coefficient : ', num2str(alpha)]);
    disp(['Disk dimension : ', num2str(a),' by ', num2str(b),' by ',
        num2str(t),' nm']);
    disp(['Saturation magnetization : ', num2str(Ms), ' A/m']);
    disp(['Anisotropy : ',num2str(K),' J/m^3'])
    disp(['Anisotropy barrier(delta) : ', num2str(delta)]);
    disp(['Hk : ',num2str(Hk),' A/m']);
    disp(['TauD : ', num2str((1+alpha^2)/(alpha*Hk*gamma)),' s']);
    disp(['External magnetic field : ', '[ ',num2str(H_ext),' ]',' A/m'])
        ;
    disp(['Current density: ', num2str(Current/area),' A/m2 ']);
    disp(['Critical current density: ', num2str(Ic0/area),' A/m2 ']);
    disp(['Effective current: i=', num2str(Current/Ic0)]);
end
```

Initialize.m:

```
function W= Initialize( p , t )
% Initialize the probability vector W
% at equilibrium condition from Boltzmann
% distribution according to parameter Delta

Total_size = size(p);
W= zeros(Total_size(1),1);

[~,Delta,~,~,~,~,~,~,~,~,M_init]=parameter(0);

for i = 1 : 1 : Total_size(1)
    Angle_cos = p(i,:)*M_init';
    if p(i,3) < 0
        W(i,1)= exp(-Delta) ;
    else
        W(i,1) = exp(-Delta*(1-Angle_cos^2));
    end
end

Dim = size(t);
num_tri = Dim(1);

sum = 0;
for k = 1 : 1 : num_tri
    t_k = t(k,:);
    sum = sum + integral_W(p,t_k,W);
end

W = W / sum; % normalize the probability

end
```

NonSwitchP.m:

```
function [NonSP] = NonSwitchP(Data,t,p)
% This funtion integrate the probability
% over the upper hemisphere - the non-switching
% probability

Dim = size(t);
num_tri = Dim(1);
Dim_data = size(Data);
NonSP = zeros(2,Dim_data(2));

EA=[0,0,1];

for n = 1 : 1 : Dim_data(2)
    sum = 0;
   for k = 1 : 1 : num_tri
      t_k = t(k,:);
      P1 = ( p(t_k(1),:) + p(t_k(2),:) + p(t_k(3),:))/3.;
      if (P1(1)*EA(1)+P1(2)*EA(2)+P1(3)*EA(3)) > 0
      sum = sum + integral_W(p,t_k,Data(:,n));
   end
end
    NonSP(1,n) = n;
    NonSP(2,n) = sum;
end

end
```

integral_W.m

```
function [ Sum ] = integral_W( p,t_k,W )
% Integrate the probability over the
% triangular patch t_k
Sum = 0;

% evaluate the functions on the seven integration points
k1     = [ 1/3, 0.0597158717, 0.4701420641, 0.4701420641...
         , 0.7974269853, 0.1012865073, 0.1012865073 ];
k2     = [ 1/3, 0.4701420641, 0.0597158717, 0.4701420641...
         , 0.1012865073, 0.7974269853, 0.1012865073 ];
k3     = [ 1/3, 0.4701420641, 0.4701420641, 0.0597158717...
         , 0.1012865073, 0.1012865073, 0.7974269853 ];
weight = [ 0.225, 0.1323941527, 0.1323941527, 0.1323941527...
         , 0.1259391805, 0.1259391805, 0.1259391805 ];

for pp = 1 : 7
    Xu = p(t_k(3),:) - p(t_k(1),:);
    Xv = p(t_k(2),:) - p(t_k(1),:);
    A = Xu * Xu';
    B = Xu * Xv';
    C = Xv * Xv';
    Phi_1 = fun_phi2(1,k3(pp),k2(pp));
    Phi_2 = fun_phi2(2,k3(pp),k2(pp));
    Phi_3 = fun_phi2(3,k3(pp),k2(pp));
    Sum = Sum + sqrt(A * C - B^2) * weight(pp) ...
            * (  Phi_1 * W(t_k(1),1) ...
               + Phi_2 * W(t_k(2),1) ...
               + Phi_3 * W(t_k(3),1));
end
end
```

Appendix C

Publications

C.1 Journals

1. Xie, Yunkun, Yaohua Tan, and Avik W. Ghosh. " spintronic signatures of Klein tunneling in topological insulators." *Physical Review B* 96.20 (2017): 205151.

2. Xie, Yunkun, Jianhua Ma, Samiran Ganguly, and Avik W. Ghosh. "From materials to systems: a multiscale analysis of nanomagnetic switching." *Journal of Computational Electronics* 16, no. 4 (2017): 1201-1226.

3. Xie, Yunkun, Behtash Behin-Aein, and Avik W. Ghosh. "Fokker-Planck study of parameter dependence on write error slope in spin-torque switching." *IEEE Transactions on Electron Devices* 64, no. 1 (2017): 319-324.

4. Ma, Jianhua, Vinay I. Hegde, Kamaram Munira, Yunkun Xie, Sahar Keshavarz, David T. Mildebrath, C. Wolverton, Avik W. Ghosh, and W. H. Butler. "Computational investigation of half-Heusler compounds for spintronics applications." *Physical Review B* 95, no. 2 (2017): 024411.

5. Munira, Kamaram, Yunkun Xie, Souheil Nadri, Mark B. Forgues, Mohammad Salehi Fashami, Jayasimha Atulasimha, Supriyo Bandyopadhyay, and Avik W. Ghosh. "Reducing error rates in straintronic multiferroic nanomagnetic logic by pulse shaping." *Nanotechnology* 26, no. 24 (2015): 245202.

6. Ma, Marco Chung Ting, Yunkun Xie, Howard Sheng, S. Joseph Poon, and Avik Ghosh. "Skyrmion Formation Induced by Antiferromagnetic-enhanced Interfacial Dzyaloshinskii Moriya Interaction." *preprint arXiv*:1806.06334 (2018).

7. Xie, Yunkun, Jianhua Ma, Hamed Vakilitaleghani, Yaohua Tan, and Avik W. Ghosh. "Computational search for ultrasmall and fast skyrmions in the Inverse Heusler family." *in preparation.*

C.2 Book Chapter

1. Xie, Yunkun, Ivan Rungger, Kamaram Munira, Maria Stamenova, A. Sanvito, and Avik W. Ghosh. "Spin transfer torque: a multiscale picture." *Nanomagnetic and Spintronic Devices for Energy-Efficient Memory and Computing* (2016): 91.

C.3 Conferences

1. Xie, Yunkun, Jianhua Ma, Hamed Vakilitaleghani, Yaohua Tan, and Avik W. Ghosh. "Search for ultrasmall&fast skyrmions in Heuslers." *International Conference on Magnetism* (2018)

2. Ganguly, Samiran, Yunkun Xie, and Avik Ghosh. "Energy-delay-reliability of present and next generation STT-RAM technology." *In Nanotechnology (IEEE-NANO), IEEE 17th International Conference on*, pp. 1010-1013. IEEE, (2017).

3. Xie, Yunkun, Behtash Behin-Aein, and Avik Ghosh. "Numerical Fokker-Planck simulation of stochastic write error in spin torque switching with thermal noise." *In Device Research Conference (DRC), 74th Annual,* pp. 1-2. IEEE, (2016).

4. Jianhua Ma, Yunkun Xie, Kamaram Munira, Avik W. Ghosh, William H. Butler. "Computational Investigation of Half-Heusler based Magnetic Tunnel Junctions." *The CNMS User Meeting and Featured Workshops* (2016)

Bibliography

[1] Yunkun Xie, Behtash Behin-Aein, and Avik W Ghosh. Fokkerplanck study of parameter dependence on write error slope in spin-torque switching. *IEEE Transactions on Electron Devices*, 64(1):319–324, 2017. [2017] IEEE. Reprinted, with permission.

[2] H. Liu, D. Bedau, J. Z. Sun, S. Mangin, E. E. Fullerton, J. a. Katine, and a. D. Kent. Dynamics of spin torque switching in all-perpendicular spin valve nanopillars. *Journal of Magnetism and Magnetic Materials*, 358-359:233–258, 2014.

[3] Z. Li and S. Zhang. Thermally assisted magnetization reversal in the presence of a spin-transfer torque. *Physical Review B - Condensed Matter and Materials Physics*, 69(13):1–6, 2004.

[4] KM Masum Habib, Redwan N Sajjad, and Avik W Ghosh. Chiral tunneling of topological states: Towards the efficient generation of spin current using spin-momentum locking. *Physical review letters*, 114(17):176801, 2015.

[5] Michael J Donahue. Oommf user's guide, version 1.0. Technical report, 1999.

[6] Gordon E Moore et al. Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1):82–85, 1998.

[7] John C Slonczewski. Conductance and exchange coupling of two ferromagnets separated by a tunneling barrier. *Physical Review B*, 39(10):6995, 1989.

[8] L Berger. Emission of spin waves by a magnetic multilayer traversed by a current. *Physical Review B*, 54(13):9353, 1996.

[9] D Pantel, S Goetze, D Hesse, and M Alexe. Reversible electrical switching of spin polarization in multiferroic tunnel junctions. *Nature materials*, 11(4):289–293, 2012.

[10] JT Heron, JL Bosse, Q He, Y Gao, M Trassin, L Ye, JD Clarkson, C Wang, Jian Liu, S Salahuddin, et al. Deterministic switching of ferromagnetism at room temperature using an electric field. *Nature*, 516(7531):370–373, 2014.

[11] Wei-Gang Wang, Mingen Li, Stephen Hageman, and CL Chien. Electric-field-assisted switching in magnetic tunnel junctions. *Nature materials*, 11(1):64–68, 2012.

[12] Luqiao Liu, Chi-Feng Pai, Y Li, HW Tseng, DC Ralph, and RA Buhrman. Spin-torque switching with the giant spin hall effect of tantalum. *Science*, 336(6081):555–558, 2012.

[13] WH Butler, X-G Zhang, TC Schulthess, and JM MacLaren. Spin-dependent tunneling conductance of fe— mgo— fe sandwiches. *Physical Review B*, 63(5):054416, 2001.

[14] Yunkun Xie, Ivan Rungger, Kamaram Munira, Maria Stamenova, A Sanvito, and Avik W Ghosh. Spin transfer torque: a multiscale picture. *Nanomagnetic and Spintronic Devices for Energy-Efficient Memory and Computing*, page 91, 2016.

[15] Supriyo Datta. *Electronic transport in mesoscopic systems*. Cambridge university press, 1997.

[16] Tchavdar N Todorov. Tight-binding simulation of current-carrying nanostructures. *Journal of Physics: Condensed Matter*, 14(11):3049, 2002.

[17] Ruoxing Zhang, Ivan Rungger, Stefano Sanvito, and Shimin Hou. Current-induced energy barrier suppression for electromigration from first principles. *Physical Review B*, 84(8):085445, 2011.

[18] Xingtao Jia, Ke Xia, Youqi Ke, and Hong Guo. Nonlinear bias dependence of spin-transfer torque from atomic first principles. *Physical Review B*, 84(1):014401, 2011.

[19] Ioannis Theodonis, Nicholas Kioussis, Alan Kalitsov, Mairbek Chshiev, and WH Butler. Anomalous bias dependence of spin torque in magnetic tunnel junctions. *Physical review letters*, 97(23):237205, 2006.

[20] JC Slonczewski. Currents, torques, and polarization factors in magnetic tunnel junctions. *Physical Review B*, 71(2):024411, 2005.

[21] A Reily Rocha, Víctor M García-Suárez, S Bailey, C Lambert, J Ferrer, and Stefano Sanvito. Spin and molecular electronics in atomically generated orbital landscapes. *Physical Review B*, 73(8):085414, 2006.

[22] Jack C Sankey, Yong-Tao Cui, Jonathan Z Sun, John C Slonczewski, Robert A Buhrman, and Daniel C Ralph. Measurement of the spin-transfer-torque vector in magnetic tunnel junctions. *Nature Physics*, 4(1):67, 2008.

[23] Jianhua Ma, Y Xie, and A. W Ghosh. Ab-initio transport study of half-heusler/mgo magnetic tunnel junction for spintronics application. *upcoming*.

[24] Shinji Yuasa, Taro Nagahama, Akio Fukushima, Yoshishige Suzuki, and Koji Ando. Giant room-temperature magnetoresistance in single-crystal fe/mgo/fe magnetic tunnel junctions. *Nature materials*, 3(12):868–871, 2004.

[25] Stuart SP Parkin, Christian Kaiser, Alex Panchula, Philip M Rice, Brian Hughes, Mahesh Samant, and See-Hun Yang. Giant tunnelling magnetoresistance at room temperature with mgo (100) tunnel barriers. *Nature materials*, 3(12):862, 2004.

[26] S Ikeda, J Hayakawa, Y Ashizawa, YM Lee, K Miura, H Hasegawa, M Tsunoda, F Matsukura, and H Ohno. Tunnel magnetoresistance of 604% at 300 k by suppression of ta diffusion in co fe b/ mg o/ co fe b pseudo-spin-valves annealed at high temperature. *Applied Physics Letters*, 93(8):082508, 2008.

[27] YM Lee, J Hayakawa, S Ikeda, F Matsukura, and H Ohno. Effect of electrode composition on the tunnel magnetoresistance of pseudo-spin-valve magnetic tunnel junction with a mgo tunnel barrier. *Applied physics letters*, 90(21):212507, 2007.

[28] J. Z. Sun, R. P. Robertazzi, J. Nowak, P. L. Trouilloud, G. Hu, D. W. Abraham, M. C. Gaidis, S. L. Brown, E. J. O'Sullivan, W. J. Gallagher, and D. C. Worledge. Effect of subvolume excitation and spin-torque efficiency on magnetic switching. *Physical Review B - Condensed Matter and Materials Physics*, 84(6), 2011.

[29] Yunkun Xie, Jianhua Ma, Samiran Ganguly, and Avik W Ghosh. From materials to systems: a multiscale analysis of nanomagnetic switching. *Journal of Computational Electronics*, 16(4):1201–1226, 2017.

[30] Adrien F. Vincent, Nicolas Locatelli, Jacques-Olivier Klein, Weisheng S. Zhao, Sylvie Galdin-Retailleau, and Damien Querlioz. Analytical Macrospin Modeling of the Stochastic Switching Time of Spin-Transfer Torque Devices. *IEEE Transactions on Electron Devices*, 62(1):164–170, jan 2015.

[31] Kyungmi Song and Kyung Jin Lee. Spin-transfer-torque efficiency enhanced by edge-damage of perpendicular magnetic random access memories. *Journal of Applied Physics*, 118(5):17–21, 2015.

[32] William Fuller Brown. Thermal fluctuations of a single-domain particle. *Phys. Rev.*, 130:1677–1686, Jun 1963.

[33] Kezhao Zhang and Donald R Fredkin. Thermal switching of a stoner-wohlfarth particle: Numerical solution of fokker-planck equation without symmetry. Ph.D. dissertation, Dept. Phys., Univ. of California, San Diego., La Jolla, CA, 2001.

[34] Per-Olof Persson and Gilbert Strang. A simple mesh generator in matlab. *SIAM review*, 46(2):329–345, 2004.

[35] W. H. Butler, Tim Mewes, Claudia K a Mewes, P. B. Visscher, William H. Rippard, Stephen E. Russek, and Ranko Heindl. Switching distributions for perpendicular spin-torque devices within the macrospin approximation. *IEEE Transactions on Magnetics*, 48(12):4684–4700, 2012.

[36] S Ikeda, K Miura, H Yamamoto, K Mizunuma, HD Gan, M Endo, S Kanai, J Hayakawa, F Matsukura, and H Ohno. A perpendicular-anisotropy cofeb–mgo magnetic tunnel junction. *Nature materials*, 9(9):721–724, 2010.

[37] M. Gajek, J. J. Nowak, J. Z. Sun, P. L. Trouilloud, E. J. O'Sullivan, D. W. Abraham, M. C. Gaidis, G. Hu, S. Brown, Y. Zhu, R. P. Robertazzi, W. J. Gallagher, and D. C. Worledge. Spin torque switching of 20 nm magnetic tunnel junctions with perpendicular anisotropy. *Applied Physics Letters*, 100(13):2012–2015, 2012.

[38] R. C. Sousa. Tunneling hot spots and heating in magnetic tunnel junctions. *Journal of Applied Physics*, 95(11):6783, 2004.

[39] J. Sun. Spin-current interaction with a monodomain magnetic body: A model study. *Physical Review B*, 62(1):570–578, 2000.

[40] P. G. Mather, J. C. Read, and R. a. Buhrman. Disorder, defects, and band gaps in ultrathin (001) MgO tunnel barrier layers. *Physical Review B*, 73(20):205412, 2006.

[41] G. X. Miao, Y. J. Park, J. S. Moodera, M. Seibt, G. Eilers, and M. Münzenberg. Disturbance of Tunneling Coherence by Oxygen Vacancy in Epitaxial Fe/MgO/Fe Magnetic Tunnel Junctions. *Physical Review Letters*, 100(24):246803, 2008.

[42] Niladri N Mojumder, David W Abraham, Kaushik Roy, and DC Worledge. Magnonic spin-transfer torque mram with low power, high speed, and error-free switching. *Magnetics, IEEE Transactions on*, 48(6):2016–2024, 2012.

[43] Arno van den Brink, Stefan Cosemans, Sven Cornelissen, Mauricio Manfrini, Adrien Vaysset, Wim Van Roy, Tai Min, HJM Swagten, and Bert Koopmans. Spin-hall-assisted magnetic random access memory. *Applied Physics Letters*, 104(1):012403, 2014.

[44] Huanlong Liu, Daniel Bedau, Dirk Backes, JA Katine, Jürgen Langer, and AD Kent. Ultrafast switching in magnetic tunnel junction based orthogonal spin transfer devices. *Applied Physics Letters*, 97(24):242510, 2010.

[45] Justin M Shaw, Hans T Nembach, Mathias Weiler, TJ Silva, Martin Schoen, Jonathan Z Sun, and Daniel C Worledge. Perpendicular magnetic anisotropy and easy cone state in ta/co 60 fe 20 b 20/mgo. *IEEE Magnetics Letters*, 6:1–4, 2015.

[46] Yunkun Xie, Yaohua Tan, and Avik W Ghosh. Spintronic signatures of klein tunneling in topological insulators. *Physical Review B*, 96(20):205151, 2017.

[47] Bo Zhou, ZK Liu, JG Analytis, K Igarashi, SK Mo, DH Lu, RG Moore, IR Fisher, T Sasagawa, ZX Shen, et al. Controlling the carriers of topological insulators by bulk and surface doping. *Semiconductor Science and Technology*, 27(12):124002, 2012.

[48] Ngoc Han Tu, Yoichi Tanabe, Yosuke Satake, Khuong Kim Huynh, and Katsumi Tanigaki. In-plane topological pn junction in the three-dimensional topological insulator $Bi_{2-x}Sb_xTe_{3-y}Se_y$. *Nature communications*, 7:13763, 2016.

[49] Sung Hwan Kim, Kyung-Hwan Jin, Byung Woo Kho, Byeong-Gyu Park, Feng Liu, Jun Sung Kim, and Han Woong Yeom. Atomically abrupt topological p–n junction. *ACS nano*, 11(10):9671–9677, 2017.

[50] Oskar Klein. Die reflexion von elektronen an einem potentialsprung nach der relativistischen dynamik von dirac. *Zeitschrift für Physik*, 53(3-4):157–165, 1929.

[51] Christopher Gutiérrez, Lola Brown, Cheol-Joo Kim, Jiwoong Park, and Abhay N Pasupathy. Klein tunnelling and electron trapping in nanometre-scale graphene quantum dots. *Nature Physics*, 12(11):1069, 2016.

[52] Shaowen Chen, Zheng Han, Mirza M Elahi, KM Masum Habib, Lei Wang, Bo Wen, Yuanda Gao, Takashi Taniguchi, Kenji Watanabe, James Hone, et al. Electron optics with pn junctions in ballistic graphene. *Science*, 353(6307):1522–1525, 2016.

[53] Seokmin Hong, Vinh Diep, Supriyo Datta, and Yong P Chen. Modeling potentiometric measurements in topological insulators including parallel channels. *Physical Review B*, 86(8):085131, 2012.

[54] Shehrin Sayed, Seokmin Hong, and Supriyo Datta. Multi-terminal spin valve on channels with spin-momentum locking. *Scientific reports*, 6:35658, 2016.

[55] CH Li, OMJ Vant Erve, JT Robinson, Y Liu, L Li, and BT Jonker. Electrical detection of charge-current-induced spin polarization due to spin-momentum locking in bi 2 se 3. *Nature nanotechnology*, 9(3):218, 2014.

[56] Joon Sue Lee, Anthony Richardella, Danielle Reifsnyder Hickey, K Andre Mkhoyan, and Nitin Samarth. Mapping the chemical potential dependence of current-induced spin polarization in a topological insulator. *Physical Review B*, 92(15):155312, 2015.

[57] Jifa Tian, Ireneusz Miotkowski, Seokmin Hong, and Yong P Chen. Electrical injection and detection of spin-polarized currents in topological insulator bi 2 te 2 se. *Scientific reports*, 5:14293, 2015.

[58] CH Li, OMJ vant Erve, YY Li, L Li, and BT Jonker. Electrical detection of the helical spin texture in a p-type topological insulator sb 2 te 3. *Scientific reports*, 6:29533, 2016.

[59] Xiao-Liang Qi and Shou-Cheng Zhang. Topological insulators and superconductors. *Reviews of Modern Physics*, 83(4):1057, 2011.

[60] Joseph G Checkelsky, Yew San Hor, M-H Liu, D-X Qu, Robert Joseph Cava, and NP Ong. Quantum interference in macroscopic crystals of nonmetallic bi 2 se 3. *Physical Review Letters*, 103(24):246601, 2009.

[61] AA Taskin and Yoichi Ando. Quantum oscillations in a topological insulator bi 1- x sb x. *Physical Review B*, 80(8):085303, 2009.

[62] Nicholas P Butch, Kevin Kirshenbaum, Paul Syers, Andrei B Sushkov, Gregory S Jenkins, H Dennis Drew, and Johnpierre Paglione. Strong surface scattering in ultrahigh-mobility bi 2 se 3 topological insulator crystals. *Physical Review B*, 81(24):241301, 2010.

[63] Shuang Jia, Huiwen Ji, E Climent-Pascual, MK Fuccillo, ME Charles, Jun Xiong, NP Ong, and RJ Cava. Low-carrier-concentration crystals of the topological insulator bi 2 te 2 se. *Physical Review B*, 84(23):235206, 2011.

[64] Redwan N Sajjad and Avik W Ghosh. Manipulating chiral transmission by gate geometry: switching in graphene with transmission gaps. *ACS nano*, 7(11):9808–9813, 2013.

[65] Faxian Xiu, Liang He, Yong Wang, Lina Cheng, Li-Te Chang, Murong Lang, Guan Huang, Xufeng Kou, Yi Zhou, Xiaowei Jiang, et al. Manipulating surface states in topological insulator nanoribbons. *Nature nanotechnology*, 6(4):216, 2011.

[66] Luqiao Liu, Takahiro Moriyama, DC Ralph, and RA Buhrman. Spin-torque ferromagnetic resonance induced by the spin hall effect. *Physical review letters*, 106(3):036601, 2011.

[67] B Gu, I Sugai, T Ziman, GY Guo, N Nagaosa, T Seki, K Takanashi, and S Maekawa. Surface-assisted spin hall effect in au films with pt impurities. *Physical review letters*, 105(21):216401, 2010.

[68] AN Bogdanov and DA Yablonskii. Thermodynamically stable" vortices" in magnetically ordered crystals. the mixed state of magnets. *Zh. Eksp. Teor. Fiz*, 95:182, 1989.

[69] AN Bogdanov and UK Rößler. Chiral symmetry breaking in magnetic thin films and multilayers. *Physical review letters*, 87(3):037203, 2001.

[70] A Neubauer, C Pfleiderer, B Binz, A Rosch, R Ritz, PG Niklowitz, and P Böni. Topological hall effect in the a phase of mnsi. *Physical review letters*, 102(18):186602, 2009.

[71] C Pappas, E Lelièvre-Berna, P Falus, PM Bentley, E Moskvin, S Grigoriev, P Fouquet, and B Farago. Chiral paramagnetic skyrmion-like phase in mnsi. *Physical review letters*, 102(19):197202, 2009.

[72] SX Huang and CL Chien. Extended skyrmion phase in epitaxial fege (111) thin films. *Physical review letters*, 108(26):267201, 2012.

[73] XZ Yu, Yoshinori Onose, Naoya Kanazawa, JH Park, JH Han, Yoshio Matsui, Naoto Nagaosa, and Yoshinori Tokura. Real-space observation of a two-dimensional skyrmion crystal. *Nature*, 465(7300):901, 2010.

[74] Shinichiro Seki, XZ Yu, S Ishiwata, and Y Tokura. Observation of skyrmions in a multiferroic material. *Science*, 336(6078):198–201, 2012.

[75] Ajaya K Nayak, Vivek Kumar, Tianping Ma, Peter Werner, Eckhard Pippel, Roshnee Sahoo, Franoise Damay, Ulrich K Rößler, Claudia Felser, and Stuart SP Parkin. Magnetic antiskyrmions above room temperature in tetragonal heusler materials. *Nature*, 548(7669):561, 2017.

[76] Stefan Heinze, Kirsten Von Bergmann, Matthias Menzel, Jens Brede, André Kubetzka, Roland Wiesendanger, Gustav Bihlmayer, and Stefan Blügel. Spontaneous atomic-scale magnetic skyrmion lattice in two dimensions. *Nature Physics*, 7(9):713, 2011.

[77] Niklas Romming, Christian Hanneken, Matthias Menzel, Jessica E Bickel, Boris Wolter, Kirsten von Bergmann, André Kubetzka, and Roland Wiesendanger. Writing and deleting single magnetic skyrmions. *Science*, 341(6146):636–639, 2013.

[78] Seonghoon Woo, Kai Litzius, Benjamin Krüger, Mi-Young Im, Lucas Caretta, Kornel Richter, Maxwell Mann, Andrea Krone, Robert M Reeve, Markus Weigand, et al. Observation of room-temperature magnetic skyrmions and their current-driven dynamics in ultrathin metallic ferromagnets. *Nature materials*, 15(5):501, 2016.

[79] Lucas Caretta, Maxwell Mann, Felix Büttner, Kohei Ueda, Bastian Pfau, Christian M Günther, Piet Hessing, Alexandra Churikova, Christopher Klose, Michael Schneider, et al. Fast current-driven domain walls and small skyrmions in a compensated ferrimagnet. *Nature nanotechnology*, page 1, 2018.

[80] Stuart Parkin and See-Hun Yang. Memory on the racetrack. *Nature nanotechnology*, 10(3):195–198, 2015.

[81] João Sampaio, Vincent Cros, Stanislas Rohart, André Thiaville, and Albert Fert. Nucleation, stability and current-induced motion of isolated magnetic skyrmions in nanostructures. *Nature nanotechnology*, 8(11):839, 2013.

[82] IE Dzyaloshinskii. Theory of helicoidal structures in antiferromagnets. 1. nonmetals. *Sov. Phys. JETP*, 19:960–971, 1964.

[83] S Rohart and A Thiaville. Skyrmion confinement in ultrathin film nanostructures in the presence of dzyaloshinskii-moriya interaction. *Physical Review B*, 88(18):184422, 2013.

[84] A Siemens, Y Zhang, J Hagemeister, EY Vedmedenko, and Roland Wiesendanger. Minimal radius of magnetic skyrmions: statics and dynamics. *New Journal of Physics*, 18(4):045021, 2016.

[85] S Rohart, J Miltat, and A Thiaville. Path to collapse for an isolated néel skyrmion. *Physical Review B*, 93(21):214412, 2016.

[86] Kai Litzius, Ivan Lemesh, Benjamin Krüger, Pedram Bassirian, Lucas Caretta, Kornel Richter, Felix Büttner, Koji Sato, Oleg A Tretiakov, Johannes Förster, et al. Skyrmion hall effect revealed by direct time-resolved x-ray microscopy. *Nature Physics*, 13(2):170, 2017.

[87] Wanjun Jiang, Xichao Zhang, Guoqiang Yu, Wei Zhang, Xiao Wang, M Benjamin Jungfleisch, John E Pearson, Xuemei Cheng, Olle Heinonen, Kang L Wang, et al. Direct observation of the skyrmion hall effect. *Nature Physics*, 13(2):162, 2017.

[88] Hiu Tung Fook, Wei Liang Gan, Indra Purnama, and Wen Siang Lew. Mitigation of magnus force in current-induced skyrmion dynamics. *IEEE Transactions on Magnetics*, 51(11):1–4, 2015.

[89] Se Kwon Kim, Kyung-Jin Lee, and Yaroslav Tserkovnyak. Self-focusing skyrmion racetracks in ferrimagnets. *Physical Review B*, 95(14):140404, 2017.

[90] Xichao Zhang, Motohiko Ezawa, and Yan Zhou. Thermally stable magnetic skyrmions in multilayer synthetic antiferromagnetic racetracks. *Physical Review B*, 94(6):064406, 2016.

[91] Xichao Zhang, Yan Zhou, and Motohiko Ezawa. Magnetic bilayer-skyrmions without skyrmion hall effect. *Nature communications*, 7:10293, 2016.

[92] Felix Büttner, Ivan Lemesh, and Geoffrey SD Beach. Theory of isolated magnetic skyrmions: From fundamentals to room temperature applications. *Scientific reports*, 8(1):4464, 2018.

[93] AA Thiele. Steady-state motion of magnetic domains. *Physical Review Letters*, 30(6):230, 1973.

[94] Chunsheng Liu, Claudia KA Mewes, Mairbek Chshiev, Tim Mewes, and William H Butler. Origin of low gilbert damping in half metals. *Applied Physics Letters*, 95(2):022509, 2009.

[95] Ankit Kumar, Fan Pan, Sajid Husain, Serkan Akansel, Rimantas Brucas, Lars Bergqvist, Sujeet Chaudhary, and Peter Svedlindh. Temperature-dependent gilbert damping of co 2 feal thin films with different degree of atomic order. *Physical Review B*, 96(22):224425, 2017.

[96] Jianhua Ma, Jiangang He, Dipanjan Mazumdar, Kamaram Munira, Sahar Keshavarz, Tim Lovorn, C Wolverton, Avik W Ghosh, and William H Butler. Computational investigation of inverse heusler compounds for spintronics applications. *Physical Review B*, 98(9):094410, 2018.

[97] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169–11186, Oct 1996.

[98] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758–1775, Jan 1999.

[99] H Ebert, D Kdderitzsch, and J Minr. Calculating condensed matter properties using the kkr-green's function methodrecent developments and applications. *Reports on Progress in Physics*, 74(9):096501, 2011.

[100] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.

[101] A.I. Liechtenstein, M.I. Katsnelson, V.P. Antropov, and V.A. Gubanov. Local spin density functional approach to the theory of exchange interactions in ferromagnetic metals and alloys. *Journal of Magnetism and Magnetic Materials*, 67(1):65 – 74, 1987.

[102] Siham Ouardi, Gerhard H. Fecher, Claudia Felser, and Jürgen Kübler. Realization of spin gapless semiconductors: The heusler compound mn_2CoAl. *Phys. Rev. Lett.*, 110:100401, Mar 2013.

[103] G. D. Liu, X. F. Dai, H. Y. Liu, J. L. Chen, Y. X. Li, Gang Xiao, and G. H. Wu. mn_2Coz ($z = \mathrm{Al, Ga, In, Si, Ge, Sn, Sb}$) compounds: Structural, electronic, and magnetic properties. *Phys. Rev. B*, 77:014424, Jan 2008.

[104] Rie Y. Umetsu and Takeshi Kanomata. Spin stiffness constant of half-metallic ferrimagnet in mn-based heusler alloys. *Physics Procedia*, 75:890 – 897, 2015. 20th International Conference on Magnetism, ICM 2015.