November 2020

Reading and Programming Spintronic Devices for Biomimetic Applications and Fault-tolerant Memory Design

Kawsher Ahmed Roxy

University of South Florida

Follow this and additional works at: https://scholarcommons.usf.edu/etd

Part of the Electrical and Computer Engineering Commons

Scholar Commons Citation

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Reading and Programming Spintronic Devices for Biomimetic Applications and Fault-tolerant Memory Design

by

Kawsher Ahmed Roxy

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy
Department of Electrical Engineering
College of Engineering
University of South Florida

Major Professor: Sanjukta Bhanja, Ph.D.
Wilfrido A. Moreno, Ph.D.
Ismail Uysal, Ph.D.
Rasim Guldiken, Ph.D.
Ravi Panchumarthy, Ph.D.

Date of Approval:
November 12, 2020

Keywords: Spintronics, Transverse Read, Domain Wall Memory, Non-Volatile Memory, In-Memory Computation

Copyright © 2020, Kawsher Ahmed Roxy
Dedication

This dissertation is dedicated to

My Parents
My Wife, Arifa
My Daughter, Rayyan
My Brother
Acknowledgments

Finally, the journey is nearing its end. There are many who helped me along the way on this journey through their guidance, support and care. First, I would like to express my utmost gratitude to my graduate advisor, Dr. Sanjukta Bhanja, who was always happy to discuss research ideas and provide guidance, constant encouragement, and precious professional and personal advice. Specially at the struggling moments, her continuous support made graduate school an incredibly rewarding experience. I would like to thank my committee members, Dr. Wilfrido Moreno and Dr. Ismail Uysal, Dr. Rasim Guldiken, Dr. Ravi Panchumarthy and the chairperson of my defense, Dr. Kingsley Reeves. Without their guidance, I would not have made it.

I would like to thank several other faculty members from whom I have learned a lot and who have given me valuable academic and career advice. Those professors include Dr. Sylvia Thomas, Dr. Srinivash Kathkoori, and Dr. Jorge Galvis. I also want to thank the staff of the EE main office for their help and the entire department for their warmth and welcoming manner.

My acknowledgement would be incomplete without thanking my wonderful friends for their significant supports during my journey towards this degree. I especially want to thank Abir with whom I discussed research ideas for hours. Here comes to thank Imran, Inteha, Shiblee, Tirtha, Dewan for all the fun-filled tours and memorable moments. This is also the time to thank Hasnat, Shuvo, Adnan, Safi, Ridita, Zahangir, Arindam, Arif, Sampad, Jakir for their encouragement. I must also thank my great labmates: Ilia, Pavia, Prayash for their unconditional support.

Finally, to my wife, Arifa, and my daughter, Rayyan: your love and understanding helped me through the challenging times. For the last five years the person whose contributions are the most is my wonderful wife, Arifa. I am also blessed to have her as my colleague. Without you believing in me, I never would have made it. My achievements and this dissertation remain incomplete without acknowledging the unconditional love and support that I got from my parents and brother.
Table of Contents

List of Tables iv

List of Figures v

Abstract vii

Chapter 1: Introduction 1
  1.1 Conventional CMOS Memories 2
  1.2 Emerging Memories: Spintronic Memories 4
  1.3 Spintronic Memories in Bio-Mimetic Computing Architectures 6
  1.4 Contribution 7
  1.5 Outline of the Dissertation 9

Chapter 2: Spintronic Memories 11
  2.1 Single Domain Nanomagnet 11
  2.2 Multilayer Spintronic Devices 12
  2.3 Magnetic Tunnel Junction 13
    2.3.1 Tunnel Magnetoresistance 14
    2.3.2 Spin Transfer Torque (STT) 16
    2.3.3 CMOS Integration 19
  2.4 Spin Valve 19
  2.5 Domain Wall Memories 20
  2.6 Application of Spintronic Devices 21
    2.6.1 Memory Applications 22
    2.6.2 GMR Sensors 22
    2.6.3 Spin Torque Nano-Oscillator 23
    2.6.4 All-Spin Logic Device (ASLD) 24
    2.6.5 Nanomagnetic Energy Minimizing Co-processor 25
  2.7 Conclusion 26

Chapter 3: Reading of Magnetic Energy Minimizing Co-Processor 27
  3.1 Introduction 27
  3.2 Quadratic Optimization via Energy Minimization of Nanomagnets 28
    3.2.1 A Case Study-Perceptual Grouping in Computer Vision 28
    3.2.2 Total Magnetic Energy in the Magnetic System 29
    3.2.3 Mapping the Optimization Problem to the Magnetic System 29
    3.2.4 Translating Output of the Problem from Magnetic States 31
  3.3 Electrical Reading of the Output 31
    3.3.1 Structure and Resistance Modelling of a Cell 32
  3.4 Reading Schemes 33
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Comparison between traditional and emerging memories.</td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>Definition of parameters used in this chapter.</td>
<td>17</td>
</tr>
<tr>
<td>3.1</td>
<td>Current requirements for reading the cells.</td>
<td>39</td>
</tr>
<tr>
<td>3.2</td>
<td>Sense margin improvement by using pre-amplifier.</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>Comparison between different methods.</td>
<td>40</td>
</tr>
<tr>
<td>5.1</td>
<td>Material properties for NiFe and CoFeB-MgO used in simulation.</td>
<td>60</td>
</tr>
<tr>
<td>5.2</td>
<td>Definition of symbols used in this chapter.</td>
<td>61</td>
</tr>
<tr>
<td>5.3</td>
<td>Comparative analysis of IMA and PMA nanowires.</td>
<td>65</td>
</tr>
<tr>
<td>5.4</td>
<td>Performance comparison between conventional and transverse read.</td>
<td>67</td>
</tr>
<tr>
<td>6.1</td>
<td>Material properties used in 16-bit nanowire simulation.</td>
<td>77</td>
</tr>
</tbody>
</table>
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>A typical memory hierarchy in a conventional von Neumann architecture.</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Conventional CMOS memories.</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Emerging devices and non-Boolean frameworks.</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>Spin based devices.</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>Different layers and ground states of an MTJ.</td>
<td>13</td>
</tr>
<tr>
<td>2.3</td>
<td>Band diagram in two ground states.</td>
<td>15</td>
</tr>
<tr>
<td>2.4</td>
<td>Different torques involved in switching mechanism.</td>
<td>17</td>
</tr>
<tr>
<td>2.5</td>
<td>The STT effect caused by the electron flow.</td>
<td>18</td>
</tr>
<tr>
<td>2.6</td>
<td>An MTJ cell is integrated with an access transistor.</td>
<td>19</td>
</tr>
<tr>
<td>2.7</td>
<td>Structure and access ports of an DWM nanowire.</td>
<td>21</td>
</tr>
<tr>
<td>2.8</td>
<td>Basic configuration and output of an STNO.</td>
<td>23</td>
</tr>
<tr>
<td>2.9</td>
<td>Anatomy of a basic all spin logic Device.</td>
<td>24</td>
</tr>
<tr>
<td>2.10</td>
<td>A sketch of the nanomagnetic co-processor.</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>Steps involved in problem mapping into nanomagnetic co-processor.</td>
<td>30</td>
</tr>
<tr>
<td>3.2</td>
<td>Magnetic state variables.</td>
<td>31</td>
</tr>
<tr>
<td>3.3</td>
<td>Resistance calculation due to process variation.</td>
<td>32</td>
</tr>
<tr>
<td>3.4</td>
<td>Read circuit to read magnetic states.</td>
<td>34</td>
</tr>
<tr>
<td>3.5</td>
<td>Variation tolerance of the circuit in reference resistance mechanism.</td>
<td>35</td>
</tr>
<tr>
<td>3.6</td>
<td>Variation tolerance of the circuit in differential read mechanism.</td>
<td>37</td>
</tr>
<tr>
<td>3.7</td>
<td>Pre-amplifier stage of the comparator.</td>
<td>38</td>
</tr>
<tr>
<td>4.1</td>
<td>Extension and comparison of simulation results with MFM micrographs.</td>
<td>43</td>
</tr>
<tr>
<td>4.2</td>
<td>The disk dimensions and layouts associated with the simulations.</td>
<td>44</td>
</tr>
<tr>
<td>4.3</td>
<td>Magnetic coupling energy.</td>
<td>45</td>
</tr>
</tbody>
</table>
Figure 4.4: Phase-plot of a CoFeB nanodot.
Figure 4.5: SOT-mediated programmable array of nanomagnets.
Figure 4.6: A hypothetical magnetic layout with SOT-based reconfigurability.
Figure 5.1: Anatomy of a typical DWM nanowire.
Figure 5.2: Signals associated with read, write and shift operation.
Figure 5.3: Conduction electrons bend at the wall exerting a force on the wall.
Figure 5.4: The DWM segment used for simulation.
Figure 5.5: A transverse read in both directions
Figure 5.6: Resistance values of IMA and PMA nanowires for different combinations.
Figure 5.7: Analysis of variation of CoFeB DW nanowire.
Figure 5.8: Parallel TR study for longer nanowires
Figure 5.9: Number of ones on the sides for different positions.
Figure 6.1: Position errors during shifting.
Figure 6.2: Phase boundary between transverse and vortex wall.
Figure 6.3: A domain wall is pinned at a triangular notch.
Figure 6.4: A 4-domain nanowire with one deformed notch.
Figure 6.5: Normal distribution of notch’s width and depth.
Figure 6.6: Modeling the curvature of the deformity.
Figure 6.7: Angle and radius of notch to achieve shift current for different positions.
Figure 6.8: Pinning potential variation with the deformities in a notch.
Figure 6.9: A domain wall nanowire with 4 domains.
Figure 6.10: Domain wall displacement with applied current pulses for shifting.
Figure 6.11: Critical current variation with deformities.
Abstract

Despite the triumph of conventional computing architectures till today, there emerges a lot of computing problems that are solved poorly by them. The reason behind this are twofold: i) the computing algorithm is incompetent in solving those problems, and ii) non-ideal effects of the traditional device technologies outperforms the benefits of using them. Hence, extensive research efforts have been put to devise novel algorithms as well as new devices. Among them spintronic devices demonstrate better performance in traditional architectures as well as offers way better solution to a lot of new problems when bundled with unconventional computing algorithms. Apart form being used as data-storage, spintronic devices are also leveraged as computing elements in many recently proposed architectures as the underlying device physics can directly solve many problems. However, the benefits of using them as computing elements become reduced or diminished because the peripherals of these architectures are based on conventional technologies. Specially the reading and programming mechanisms are not straightforward. In this dissertation, we have explored the challenges and developed the three reading techniques for a new computing framework based on spintronic devices. These nano-structured devices are prone to process variations which can significantly impair the read operation. We have addressed the process variations and modified the read techniques to combat the non-ideal effects. We have also devised a spin-orbital torque-mediated programmable magnetic grid that can solve multiple instances of a problems in same hardware. One particular spintronic device is domain wall memories. Though these memories offer high density and low power operation, they suffer from inaccurate shifting due to various faults. In this dissertation, we have discussed a novel transverse read technique to generate error correction codes to mitigate different faults. Lastly, we have modeled a new type of shifting faults, known as pinning faults, in domain wall memories due to the process variations. We have developed a analytical model of geometric variation, and characterized the fault using the critical shift current.
Chapter 1: Introduction

Recent decades experience a hike in technological innovations; a simple example can be “the on-board computer in Apollo 11 has a random-access-memory (RAM) of 32kbits and read-only-memory (ROM) of 72kbits”[1], whereas a typical smartphone contains 4Gbytes of RAM and up to 512Gbytes of storage, and easily fits in a pocket. This infers the pervasive nature of modern technologies and demonstrates how these technologies widen the capability of a human. Day-by-day with the evolution of internet-of-things, every instances of life are going to be interconnected and can be controlled by a hand-held device using a thumb. However, these benefits come with increasing challenges of obtaining ultra-fast information processing, low power consumption, expanded data storage, data security with additional expectation of miniature devices. The question arises here: ‘Are the traditional architectures capable to handle the increased computational demands?’.

This question is very fundamental, and answer to this question depends on the performance of individual components of conventional architecture. In a broad view, modern electronics are built on von Neumann architecture, where a central processing unit (CPU) computes the data stored in memories at different levels. This requires back-and-forth fetching of data from memory hierarchy and CPU. The key bottleneck of this architecture is the mismatch between the processing speed and data-fetching speed. In addition to that, two most accessed semiconductor memories by CPU, cache and random access memories (RAMs), are of low capacity due to bulkiness and consumes a great deal of energy due to non-volatility. Despite the fact that high-scalability increases the density of these memories to free more real estate to accommodate additional resources for computation, ultra-miniaturization increases the non-ideal effects significantly, and combating these effects nullify the benefits.

Another major drawback of traditional Boolean architecture is synthesizing all problems into similar framework. Parallel processing in modern computers increase the computing speed highly, there exists a lot of problems that cannot be effectively parallelized without compromising the accu-
racy. These problems often can be solved with much efficiency by different algorithms. Therefore, there is a hike in research on unconventional computing algorithms e.g. non-Boolean computation, neuromorphic computing, neural networks, in-memory computation, hierarchical temporal memory, Ising model etc. in recent years to target those problems. Many of these problems are heavily data-centric and require data storage with much higher capacity, and faster accessible with less amount of power consumption.

The evolution of these algorithms raise another question: ‘Can the existing memory technologies cope with the novel algorithmic needs?’. Clearly, the conventional complimentary metal-oxide-semiconductor (CMOS) memories experience increased challenges to be compatible with innovative applications that are data-heavy, and often require centi-second processing speed. The CMOS memories achieve immense success, even till today, because there are different types of memories to be offered for target applications. To be more specific, Fig. 1.1 captures the typical memory hierarchy of a von Neumann architecture, where two basic types of memory: i) memory-class storage, ii) storage-class memory form a memory-management to process and store data.

Figure 1.1: A typical memory hierarchy in a conventional von Neumann architecture.

1.1 Conventional CMOS Memories

Two memories in the top shelf of the hierarchy i.e. central processing unit (CPU) registers, static random access memories (SRAMs) [2] are developed to contain and process instruction sets for computation. These memories are very fast in operation, but largest in size, and consume
a huge amount of power due to higher leakage currents. The main memory typically consists of dynamic random access memories (DRAMs) [3] which have much higher densities than CPU registers and SRAMs, but process data in a slower speed. The key issue with DRAMs is the non-volatile characteristic which requires a ‘refresh’ of the stored data in a regular interval. This additional requirement increases the power consumption overhead.

The main storage of this architecture is the highly dense storage-type memories e.g. NAND FLASH [4], Solid State Drives (SSD) [5], 3D XPoint [6], etc. The main feature of this class of memory is the non-volatility which means they can retain the stored data for a longer period of time without power. Despite having the advantage of non-volatility and the higher density, the write and read-speed is several orders of magnitude slower than the cache or main memory. Moreover, the power requirements for writing new data and accessing stored data are significantly high.

Fig. 1.2 portrays state-of-art CMOS memories in their basic configurations. A typical SRAM cell, shown in Fig. 1.2a, contains two cross-coupled inverters (four transistors) to store data and two access transistors to write and access the stored data. The density of SRAM memories is poor because of the 6 transistor requirement. The sub-threshold current flow through the access transistors during power-off causes charge-leakage of stored data, therefore requires a ‘refresh’ after a certain period to restore the data. Fig. 1.2b shows a standard DRAM cell of 1T-1C cell where, the capacitor stores the data as charge, and the access transistor is to write and access the data. The charge-decay from the capacitor requires re-writing the stored data periodically. DRAMs have much better density compared to SRAM, thus are utilized as main memory of a computing
architecture. Unlike SRAM and DRAM, the NAND FLASH memories, shown in Fig. 1.2c, is non-volatile, thus the periodic ‘refresh’ is unnecessary. Since a NAND FLASH memory cell is a transistor with an additional floating gate, this memory has the highest density which makes it perfect for data-storage. The writing process involves electron tunneling though a thin oxide layer from the floating gate and the bulk. The writing and reading speed of a FLASH memory are much slower than SRAM and DRAM counterpart.

From the above discussion it is evident that SRAM and DRAM can not be efficiently used as main storage due to non-volatility and lower density. On the contrary, FLASH memories or SSDs can not be used as main memory due to ultra-slow speed. Therefore, there always have been a quest to develop a utopian memory which is non-volatile, highly dense, faster like SRAM or beyond and consume less power.

1.2 Emerging Memories: Spintronic Memories

In the decade of 2000, active research paradigm was to search for suitable alternative technologies that utilized charge transfer as the primary mode of computing. A few examples are Carbon Nanotube [7, 8], Graphene FET [9, 10], Tunnel FET [11, 12], Resonant tunneling diodes [13], Spin FET [14], Piezo [15] and NEMS [16]. However, there are other non-charge-based state variables like magnetization [17].
Table 1.1: Comparison between traditional and emerging memories based on key attributes [18, 19].

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Traditional Memories</th>
<th>Emerging Memories</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SRAM</td>
<td>DRAM</td>
</tr>
<tr>
<td>Non-Volatility</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Cell Size (F²)</td>
<td>50-120</td>
<td>6-10</td>
</tr>
<tr>
<td>Read time (ns)</td>
<td>≤ 2</td>
<td>30</td>
</tr>
<tr>
<td>Write time (ns)</td>
<td>≤ 2</td>
<td>50</td>
</tr>
<tr>
<td>Write power</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Endurance (cycles)</td>
<td>10¹⁶</td>
<td>10¹⁶</td>
</tr>
<tr>
<td>Scalability</td>
<td>Good</td>
<td>Limited</td>
</tr>
</tbody>
</table>

Memory technologies have seen an unprecedented growth. While 3D FLASH is the benchmark for non-volatile memories, many potential breakthroughs have happened through phase change memory (PCM) [20], ferroelectric RAM (FeRAM) [21], Spin-transfer-torque(STT)-MRAM [22, 23, 24]) and spintronic memories (Magnetoresistive Random Access Memory (MRAM) [25]. Table 1.1 compares the benchmark attributes of conventional memories with the emerging memories. Evidently, STT-MRAM possesses the desired features from the traditional memories as well as offers solutions to the bottlenecks of them. Hence, STT-MRAM can be the utopian memory which can make the idea of ‘single-memory computing architecture’ feasible. Since spintronic devices are the basis of STT-MRAM, most parts of this manuscript focus on magnetization as the state variable.

Another promising spintronic memory is domain wall memory (DWM) which, unlike other emerging memories, can store multiple bit in a single cell. Domain wall memories consist of long nanowires with multiple magnetic domains at regular interval. Each domain stores a single bit of data, and typically a nanowire can store 64-512 bits of data. The interesting characteristic of this type of memory is the adjacent placement of bits. This enables an additional important operation besides read and write with no extra technical overhead: shifting. The stored data in a DWM nanowire can be shifted in both direction making it a perfect candidate for cache memory. Additionally, higher density ensures maximum capacity of data-storage with domain wall memories.
Unlike the conventional electronics, the core principle of spintronic devices leverages both spin and charge properties of electrons, rather than exploiting only the charge property. Spin property of electrons faced significant ignorance in conventional logic and information processing. Earlier, spin was only exploited for magnetic recording in a macroscopic way [26], where the magnetization of ferromagnet is used. The microscopic manipulation of the spin for controlling electron transport in a device became possible after the discovery of ”Giant Magnetoresistance (GMR)” [27, 28] in 1988. Subsequently, the development of spintronic devices triggered research in a broad range of application domains, such as highly sensitive magnetic-field sensors [29], magnetic read heads [30], and nonvolatile magnetic memory applications [31]. High density, radiation hardness, and long data retention make MRAM an excellent choice for data storage and main memory in applications under unfavorable conditions.

1.3 Spintronic Memories in Bio-Mimetic Computing Architectures

In a parallel endeavor, researchers have been mesmerized by the computation that occurs in nature and human brain. The energy efficiency clearly has significant supremacy over conventional Boolean processing for a multitude of complex tasks. While we do not have a clear understanding how the brain actually works, many hypotheses have emerged for bio-mimetic computing.

The spintronics research has already established various flavors of alternate non-von Neumann problem mapping like neuromorphic [32, 33, 34], and non-Boolean computing with oscillators [35, 36]. In general, coupled oscillators have shown to solve associative processes [37, 37, 38, 39]. Recently, pairwise coupling was experimentally demonstrated in [40]. Application-specific algorithms have been proposed for signal processing [41, 42, 43, 44] for a while.
In energy minimization framework, briefly discussed in chapter 2 and 3, the computational theme is mapping the quadratic energy minimization problem spaces into a set of interacting magnets. This way the energy relationship between the problem variables is proportional to that of the dipolar coupling energies between the corresponding magnets. The optimization is actually accomplished by the relaxation physics of the magnets themselves, and solutions can be read-out in parallel. In essence, given a specific instance of the problem, a specific magnetic layout can be achieved, and the relaxed state of which will be the solution to the original problem. The key idea is that the nanomagnetic disks in a critical dimension settle into two different magnetization ground states: a vortex state when weakly coupled, or a single domain state when strongly coupled.

### 1.4 Contribution

Although a good number of computing architectures harness the innate physical properties of spintronic devices for storage and computation, the peripherals of the hardware e.g. control circuits, sensing mechanisms, writing techniques are designed with conventional CMOS technologies. Spintronic devices are still in development phase to offer a homogeneous all-spintronic platforms.
Therefore, the integration with CMOS is imperative, and quite challenging. The hybridization process between these two technologies poses bilateral constraints on each other while integration.

In addition to these, the state variables in spintronic systems depend on the electrical excitations i.e. current, voltage. This implies there is a possibility the current for reading the data can switch the magnetization of the device, in other words, can alter the stored data. Apart from this in nanomagnetic energy minimization hardware, nanomagnets have to placed within a certain distance, typically 30 nm, which put a upper limit of CMOS technology to be integrated with CMOS technologies. Prior efforts [45, 46, 47] analyzed the constraints and mitigated the challenges of integration.

The primary version of magnetic energy minimizing co-processor (MEMCoP) lacks the electrical reading method. In this dissertation, we have discussed how the ground states in a spintronic computing hardware can be detected without altering the original output. We have devised three reading mechanisms, analyzed the trade-offs and compared their performances. We have also detailed the pre-amplifier used to increase the sense margin between the ground states. Typically, the circular nanomagnets used in MEMCoP have diameter ranging between 110-150 nm with thickness varying from 4-8 nm. The fabrication imperfection impairs the magnetic states of these nanoscale dots. Therefore, in this dissertation we have extended our read techniques to handle process variations.

Also, the initial-proof-of-concept was demonstrated by fabricating a specific layout of nanomagnets for a specific problem. However, a more realistic architecture should be able to synthesize multiple instances of a problem in the same hardware. Parallel execution of multiple problems is desired too. That is why a developing programmable architecture is mandatory. Since the computing algorithm harnesses the dipolar coupling between the nanomagnets, programmability can be achieved by counterbalancing the dipolar coupling. We have developed spin-orbital torque (SOT)-mediated programmable grid to reuse the hardware for a different set of problems. We have studied the current requirement for programming, and discovered that the ramp-up speed of applying current to generate SOT has significant impact on relaxation process.

Domain wall memories provide high density storage solution by storing more than one bits in a nanowire. Multiple bits share the same read port. During the reading process, the desired bit is shifted to the read port, and after reading, shifted back to its original position. Hence,
the traditional reading process involves two shifting making the read power higher. In many applications, it is necessary to determine the parity of data or the number of ‘1’s (alternatively ‘0’s) in the data. In such scenarios, conventional reading process is proven inefficient, power hungry, and slow. In this dissertation, we have discussed how a novel transverse read in a DWM nanowire can provide the global information of the stored data without shifting back and forth. This reduces significant power consumption and make the read process faster.

Bi-directional shifting of the stored data in a domain wall nanowire aids in logical and arithmetic operations i.e. multiplication, convolution etc. Generally, two different bits of data stored in two domains are separated by a varying magnetic boundary, known as domain wall. Domain walls are pinned to particular locations along the nanowires by intentionally patterning pinning sites. During the shift process, a domain wall is depinned from the site and travel to the next site and stays pinned there. However, process variations in pinning sites create non-uniform pinning strength at different positions. In that case, a wall with a higher pinning strength is not be depinned from the site for the critical shift current. This phenomenon is characterized by ‘pinning fault’. Finally in this dissertation, we have analyzed how deformities in a notch can instigate pinning fault. We have also discussed a model to generate a distribution for variations of a notch. We envision that these contributions towards the use of spintronic devices in computing framework will resolve the key issues and benefit the overall nano-computing paradigm.

1.5 Outline of the Dissertation

1. Chapter 2 describes the fundamentals of key spintronic devices. The physics behind different operations are briefly discussed.

2. Chapter 3 discusses the read mechanisms developed to detect the ground states of magnetic energy minimizing co-processor. A brief discussion on the co-processor structure, and theory of operation is also included in this chapter.

3. Chapter 4 covers the SOT-medicated programmable nanomagnetic co-processor. This chapter also includes the analytical model of calculating pairwise dipolar coupling energy between two adjacent nanomagnets.
4. Chapter 5 details the transverse read technique in domain wall memories. A theoretical analysis along with the circuit level understanding have been discussed.

5. Chapter 6 contains the discussion on modeling the pinning fault occurred in domain wall memories.

6. Chapter 7 concludes the dissertation and sheds light on the future research direction.
Chapter 2: Spintronic Memories

Until the last decade of the twentieth century, the state-of-the-art mainstream electronics ignored spin, a critical degree of freedom of electrons. The superiority of the charge based technologies was beyond question until some drawbacks became apparent. These shortcomings triggered the research of different spin based emerging devices. Spintronic devices conform the group which exploits the spin of electrons for data storage or information processing. These devices utilize the interaction of spin of the carrier and the magnetic properties of the material. Combining the spin degree of freedom with the conventional charge-based devices, or the standalone spin-based devices enhance the performance of electronic systems. For data storage and information processing, magnets with ultra small dimensions are of crucial importance. The maturity of fabrication technologies to manufacture nanoscale devices has intrigued the research of magnets with deep sub-micron dimensions. Hence, a good number of spintronic devices have been proposed in recent few years. Some of them exhibit excellent performance and capability in memory and computing domain. In this chapter\(^1\), we will briefly cover some promising spintronic devices, as well as their theory of operation.

2.1 Single Domain Nanomagnet

The fundamental cell of the spintronic devices is a mono-layer magnet [48] which contains its three dimensions within nanometer scale. The Theory of Micromagnetics, developed by Brown [49], models a nanomagnet as a single layer spin device. His theory explains how the interplay between quantum mechanical exchange and magnetostatic energies restrains the development of multiple magnetic domains in nanoscale particles. This behavior made possible the use of nanomagnets as switching devices, data storage, and computing elements.

Cowburn et al. [48] demonstrated the magnetic behavior of single domain circular nanomagnets. An annular nanomagnet can stay at any of its two ground states: single or vortex domain based on its diameter and thickness. An external magnetic field or a certain amount of current can switch the magnetization of the nanomagnet. Kumari et al. [50] studied the behavior of nanomagnets in a 2-D array based on their dimensions and spacing between the cells. Zhang et al. [51] demonstrated the generation of an intrinsic spin transfer torque from spin-orbit coupling without using another ferromagnetic layer.

Often in memory applications and computing paradigm, a grid of single domain nanomagnets is used purposely. Hence, the neighbor interaction between the nanomagnets in such a system is a critical factor. In memory applications, neighbor coupling hinders the functionality of a data storage; therefore it is not desired. As a result, the minimum space between nanomagnetic cells limits the density of the memory. However, in computing frameworks, the coupling between these nanomagnets is exploited in the implementation of some algorithms. In section 2.6.5, we will present an architecture based on single layer circular nanomagnets solving a quadratic optimization problem.

### 2.2 Multilayer Spintronic Devices

Later, multi-layer stacked devices were introduced for more control and adaptability while solving different problems. In this scenario, a non-magnetic (NM) layer separates two single ferromagnetic (FM) layers. The type of the non-magnetic material divides the multi-layer devices into two groups: 1) GMR based devices such as a spin valve (SV) [52] in which the spacer layer is conductive, 2) TMR based devices such as Magnetic Tunnel Junction (MTJ) [53] which uses insulating oxide material as the spacer layer. One of the two ferromagnets acts as the fixed layer, and its magnetization is pinned to a particular direction. The other FM layer is the free layer, and its magnetization is manipulated to represent logic or solution to various problems. These multi-layer devices are magnetoresistive since their resistances rely on the relative magnetization of the fixed and free layer.
2.3 Magnetic Tunnel Junction

The core operational unit of a magnetic tunnel junction (MTJ) is a three-layer stack of materials in which the top and the bottom layers are ferromagnetic (FM), and the intermediate layer is an insulator (shown in Fig. 2.1b). The most prevailing structure (shown in Fig. 2.2a) of an MTJ is CoFeB (FM)/MgO (NM)/CoFeB (FM). The magnetization of the free layer (layer 1) can be manipulated to store data or process information while the magnetization of the fixed layer (layer 2) is kept firm. The fixed layer magnetization is pinned through an antiferromagnetic coupling with another ferromagnetic layer (layer 4) under it. A thin layer (layer 3) of Ruthenium (RU) is sandwiched between them to provide the maximum coupling.
If the free layer magnetization is parallel to that of fixed layer, known as the parallel state, the resistance of the whole device is lower than when the device is in the anti-parallel state. The resistance differential can be termed as “Tunneling Magnetoresistance (TMR)” [54].

The resistance of an MTJ is dependent on the relative angle, $\theta$, between the magnetization of the free layer and the fixed layer. The conductance of an MTJ can be measured from the following equation:

$$G(\theta) = \frac{1}{2}(G_P + G_{AP}) + \frac{1}{2}(G_P + G_{AP})\cos(\theta)$$  \hspace{1cm} (2.1)

where $G_{AP}$ and $G_P$ represent the conductance ($G_P > G_{AP}$) of an MTJ in antiparallel ($\theta = 180^0$) state and parallel ($\theta = 0^0$) state respectively. The difference between these resistances is scaled down to define Tunnel Magnetoresistance (TMR) as in Eq. 2.2

$$TMR = \frac{G_P - G_{AP}}{G_{AP}} = \frac{R_{AP} - R_P}{R_P}$$  \hspace{1cm} (2.2)

2.3.1 Tunnel Magnetoresistance

To understand the tunneling magnetoresistance (TMR) effect, at first, one has to analyze the band model that causes the variation in resistance. This difference arises when there is a change in the density of states of spin up and spin down electrons at Fermi level. In a case of a conventional tunneling process, the spin of electrons is conserved, since no spin-flip scattering mechanism is involved. This phenomenon directs the electrons of a specific spin from one electrode to tunnel into the states of another electrode. These states are unique to the spin orientation. In parallel magnetization orientation, the tunneling exchange occurs between the same band electrons. On the other hand, in antiparallel orientation, the tunneling exchange is forced to happen between the opposite band electrons. These events lead to a reduced number of states available for tunneling between the electrodes in the antiparallel configuration. They result in a higher tunneling resistance than parallel configuration.
The imbalance between majority and minority density of states at Fermi level causes the difference in TMR ratio [55]. The imbalance can be understood by the spin polarization factor,

\[ P = \frac{(N↑ - N↓)}{N↑ + N↓} \tag{2.3} \]

where \( N↑ \) and \( N↓ \) are the number of spin-up and spin-down states available to contribute to the tunneling current.

The definition of TMR ratio with polarization factor can be written as:

\[ TMR = \frac{2P_1P_2}{1 - P_1P_2} \tag{2.4} \]

where \( P_1 \) and \( P_2 \) are the polarization factors of the free layer and fixed layer.

TMR in real structure also depends on the properties of barrier material and other factors like electrode thickness, junction and electrode interface quality, bias voltage, etc. Depending on the crystal orientation (amorphous or crystalline) of the barrier layer, the device will behave differently. Two most favored tunnel barrier materials are AlOx and MgO.
2.3.2 Spin Transfer Torque (STT)

As the transportation of electrons in transition metal ferromagnets is spin-polarized, spin-transfer torque (STT) originates from this transfer process [56]. When a current flows through a magnetic nanostructure, there is an exchange interaction between the spin of conduction band electrons and the local moments. STT emerges when there is a non-collinearity between the spin current polarization and the local magnetization [56]. This torque can instigate several physical processes in the absence of applied magnetic field. Some of these are: domain wall movement along magnetic wires [57], magnetization reversal in nanopillars [58], and vortex displacement in magnetic dots [59].

Based on the direction of the current flow and the micromagnetic state, the spin-transfer torque can either expedite (leading to a stronger damping in the free layer) or offset (reducing the damping in the free layer) the relaxation processes. The magnetization of the free layer can be changed by either (1) external magnetic field, or (2) spin torque (generated by current passing through it). The free layer magnetization can be regulated by the following Landau-Lifshitz-Gilbert (LLG) equation with the STT term (Eq. 4.2) [60]:

\[
\frac{d\vec{M}}{dt} = -\gamma M_s \vec{M} \times (\vec{H}_{eff} - \alpha \frac{d\vec{M}}{dt} + \frac{J_e g h}{\mu_0 M_s^2 c t_F} \vec{M} \times \vec{M}_2) \tag{2.5}
\]

where,

\[
\vec{H}_{eff} = \vec{H}_{ext} + \vec{H}_k + \vec{H}_{demag} + \vec{H}_{ex} \tag{2.6}
\]

and

\[
g = [-4 + (1 + P)^3 \frac{(3 + \hat{s}_1 \cdot \hat{s}_2)}{4P^{3/2}}].
\]

The parameters in the above equations are defined in Table 2.1. The first term on the right side of the Eq. 2.5 formulates how the resultant magnetic field (in Eq. 2.6) affects the free layer precessional dynamics. The second one addresses the damping in the free layer. The last part accounts for the impact of spin transfer torque (STT) on the free layer magnetization.
Figure 2.4: Different torques in the free layer of an MTJ cell during switching; (a) STT expedites the relaxation process, (b) STT opposes the relaxation process.

Table 2.1: Definition of parameters [46].

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\vec{M}_1, \vec{M}_2$</td>
<td>Unit vectors of fixed and free layer magnetization.</td>
</tr>
<tr>
<td>$M_s$</td>
<td>Saturation magnetization.</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Gyromagnetic ratio.</td>
</tr>
<tr>
<td>$\vec{H}_{eff}$</td>
<td>Unit vector in the direction of effective magnetic field.</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Gilbert damping constant.</td>
</tr>
<tr>
<td>P</td>
<td>Spin polarizing factor.</td>
</tr>
<tr>
<td>$n_{+/-}$</td>
<td>Majority/Minority-state Fermi level spin densities [56].</td>
</tr>
<tr>
<td>$\hat{s}_1, \hat{s}_2$</td>
<td>Unit vectors along the global spin orientation of the free and fixed layers.</td>
</tr>
<tr>
<td>$t_F$</td>
<td>Thickness of free layer.</td>
</tr>
<tr>
<td>e</td>
<td>Electron charge.</td>
</tr>
<tr>
<td>$h$</td>
<td>Weighted Planck’s constant.</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>Permeability of free space.</td>
</tr>
</tbody>
</table>
When the spin torque is significant enough, it can influence the free layer magnetization. When a current flows from free to fixed layer, the damping effect and the STT torque join to switch the magnetization of the free layer to parallel state \[46\]. The reverse current makes damping and STT work against each other. If the torque is more influential than the damping, the free layer will be switched to the antiparallel state (shown in Fig. 2.4).

In this process, the torque generated by the tunneling current is used to change the magnetization. The spin transfer model states that the electrode with larger thickness polarizes the electrons of incoming current. As a result, the spin moment increases if the current density and the degree of polarization increase, producing torque. This spin torque can change the magnetization of the free layer in parallel or antiparallel direction on the reference layer (shown in Fig. 2.5). The direction of the STT reverses with reversing the current direction \[55\]. The critical current for switching the magnetization can be analytically expressed as \[46\]

\[
J_{sw} = \frac{2e\alpha M_s t_F(H_k \pm H_{ext}) + 2\pi M_s}{\hbar} \tag{2.7}
\]

STT switching is promising, but the current requirement is still high. Therefore, some other techniques are still being researched to assist the STT switching mechanism, for example, thermally assisted switching \[61\], strain-assisted switching \[62\], etc.

![Figure 2.5: The STT effect caused by the electron flow. The direction of the STT reverses with reversing the current direction \[46\].](image)
2.3.3 CMOS Integration

To write and read an MTJ cell electrically, and achieve the selectivity for low power operation, nowadays MTJs are integrated with the state of the art CMOS technology. The most well known STT-MRAM architecture is a 2-D crossbar array of 1T-1C. Fig. 2.6 shows a primary MTJ cell with an access transistor as well as different control signal lines (SL, BL, WL). Applying suitable biases in these signals, an MTJ cell is written and read. As MTJ cells are thermally robust, they can be monolithically fabricated by using standard back-end-of-line (BEOL) techniques [63].

Figure 2.6: An MTJ cell is integrated with an access transistor. Different control signal lines (SL, BL, WL) are shown. Applying suitable biases in these signals, an MTJ cell is written and read [46].

2.4 Spin Valve

A spin valve (SV) (shown in Fig. 2.1(a)) has a structure analogous to an MTJ, but the spacer layer consists of a conductive material. The basic operation of a spin valve is almost similar to an MTJ’s except the electron transport mechanism. The electrons are transported through this device by scattering. Electrons with a particular spin get accelerated while passing through a ferromagnetic layer; consequently, they scatter from it. These scattered electrons pass through the conductive spacer layer to the free layer to change their magnetization. Like an MTJ, a spin valve also exhibits two different resistances for parallel and antiparallel configuration. However, resistance difference, in this case, is termed as "Giant Magnetoresistance (GMR).” As the conductive ferromagnetic layers are separated by the conductive spacer layer, the GMR value is expected to be low. Also,
the critical switching current density for a spin valve is relatively high ($\sim 10^7 - 10^8 \text{A/cm}^2$ [64, 65]) which hinders the integration with current CMOS technology.

2.5 Domain Wall Memories

Spintronic domain-wall memory (DWM)—also referred to as “Racetrack” memory—recently proposed and demonstrated by IBM [66], is a promising candidate to overcome density limitations while retaining the static energy benefits of STT-MRAM. DWM is constructed from ferromagnetic nanowires—also referred to as “tapes” or “racetracks”—separated into domains and connected to a single (or possibly a few) access transistor(s) to create access ports, much like STT-MRAM. DWM conserves the advantages of STT-MRAM while increasing the storage density by up to $10 \times$ [67] and has a theoretical area per data bit as small as $2F^2$ [68], where $F$ is the technology feature size. Data access is obtained by shifting the magnetic domains along the nanowire and aligning the target domain to an access device. After alignment, data access is identical to STT-MRAM. DWM demonstrations of memory array structures [69] and content addressable memories (CAMs) [70] show fabrication feasibility with great potential for density, performance, and power consumption.

Unlike the single domain circular nano-magnetic disk, in ferromagnetic nanowires, there exist multiple magnetic domains. The magnetic states of adjacent domains are not parallel to balance the exchange and anisotropic energies [71]. In between two adjacent domains, a mobile non-magnetic barrier, known as domain wall (DW), is present. Mobilization of a domain wall present in a magnetic nanowire, makes them promising as a futuristic device for data storage [72] and logic applications [73]. Under the influence of applied magnetic field, domain walls propagate through the magnetic nanowire. Later, current-driven propagation has been proposed making this system useful for computation. Fig. 2.7 shows a schematic of a magnetic nanowire with domain wall. The micromagnetics of a domain wall propagation can be estimated by the Landau-Lifshitz-Gilbert (LLG) equation [74]:

$$\dot{\mathbf{m}} = \lambda_0 \mathbf{H} \times \mathbf{m} + \alpha \mathbf{m} \times \dot{\mathbf{m}} - (\mathbf{u} \cdot \Delta) \mathbf{m}$$

The velocity of propagation of domain walls in a nanowire is an important figure of merit. In the simplest model [75, 76] for a DW motion, the velocity is a function of applied magnetic field
as $u = \mu H$, where the mobility, $\mu = \frac{\gamma \Delta}{\alpha}$, depends on the gyromagnetic ratio ($\gamma$), Gilbert Damping factor ($\alpha$), and the width of the domain wall ($\Delta$).

Current-driven movement of domain walls is theoretically developed by Kohno et al. [77]. In their research work, they explained the effects of current injection in a magnetic nanowire: spin transfer and momentum transfer. Parkin et al. [78] experimentally demonstrated the current-controlled domain wall motion. The magnetization direction (arrows) of a domain is programmed to store either bit ‘1’ or bit ‘0’. Several transistors are connected to the stripe to perform read, write, and shift operations, respectively. They are called read access port.

**Figure 2.7:** (a) Structure of the strip containing 7 bits; (b) write port of DWM nanowire; (c) an MTJ as read-port.

### 2.6 Application of Spintronic Devices

Different manipulation of that three terms in Eqn. 4.2 has made possible to use MTJs in a wide range of applications, for example, in memory applications as Magnetoresistive Random Access Memory (MRAM) [79], in oscillator domain as Spin Torque Nano-Oscillator (STNO) [80], as
computing elements in memory paradigm [81], and in Non-Boolean Computing domain [82, 83, 84] in many ways.

When the damping in free layer magnetization dynamics is more significant, the magnetization can be in two ground states, parallel, or anti-parallel to the direction of the fixed layer magnetization; this leads to the use of MTJ for memory applications. If a sufficient amount of bias current is passed through an MTJ, the damping will be nullified, and the free layer magnetization will not be settled to any of the ground states. Rather, it exhibits an oscillatory behavior [85]. This phenomenon is exploited while MTJs are being used as nano-oscillator.

2.6.1 Memory Applications

In recent times, spin based memories (MRAM, STT-MRAM) have gained a lot of attention among the researchers and memory industries. The salient features of spin memories such as non-volatility, low switching energy, fast operation, high endurance, and unlimited retention make them promising candidates not only for data storage but also for main memory. Moreover, these memories outperform the other memories while operating in unfavorable conditions. Memory requirement in embedded systems can leverage this quality for operating in an adverse environment. As we are focusing on the non-Boolean computation by spintronic devices in this book, detailed description of the reading and writing operation in memory is beyond our scope. However, interested readers are encouraged to explore these articles [47, 86] for an in-depth insight.

2.6.2 GMR Sensors

In earlier days, GMR based nanosensors were used as the read head of computer hard drives [27, 87]. They can sense the change of local magnetic field by changing their electrical resistances. Later, these nanosensors have been exploited in different areas especially in biosensing. A GMR based biosensor was proposed in [88] for DNA sensing. A biochip based on GMR was proposed for DNA detection and HPV genotyping in [89]. Apart from this, in [90], GMR based biosensors exhibit better performance than the conventional probe in sensing protein for the diagnosis of heart diseases.
2.6.3 Spin Torque Nano-Oscillator

The operation of a Spin Torque Nano-Oscillators (STNO) relies on the current-driven oscillation of magnetization resulting self-sustaining oscillatory changes in the resistance of the device [80]. The oscillation of these devices evolves from the interaction between intensive nonlinear magnetization process and the spin dependent transport through the magnetic heterostructures. The simple configuration of an STNO consists of a relatively thick fixed magnetic layer, which serves as a polarizer, a non-magnetic spacer, and a relatively thin magnetic free layer as shown in Fig. 3.5a.

Figure 2.8: (a) Basic configuration of an STNO with a dc bias current. (b) Microwave voltage is generated due to the precessional dynamics in the free layer.

A dc current gets spin-polarized while passing through the polarizer. When this spin polarized current is large enough to transfer sufficient STT to cancel out the intrinsic damping losses of the free layer, a steady-state magnetization precession occurs. The magnetoresistive (MR) effect converts the magnetization oscillation to a microwave voltage (in Fig. 3.5b).

The STNOs have attained a significant attention in recent past for computing and microwave applications [37]. In computing paradigm, the information is coded as the frequency of oscillation of the STNOs. Hence, quantification of the frequency of an STNO is of critical importance. The frequency of oscillation of the free layer of an STNO is approximated by the following equation [91]:

\[ f_N(\theta, \zeta) = \frac{|\gamma|}{2\pi} \sqrt{H[H + \left( \frac{M_0}{\zeta} \right) \cos^2 \theta]} \]  

(2.8)
where $H$ and $\theta$ represents internal magnetization field magnitude and direction. $M_0$ is the free layer magnetization vector magnitude and $\zeta$, called supercriticality, can be expressed by

$$\zeta = \frac{I}{I_{th}}$$

where $I$ and $I_{th}$ are the applied and threshold currents respectively.

2.6.4 All-Spin Logic Device (ASLD)

Though the spintronic devices discussed above, considers the neglected electron spin as a state variable besides charge, still these devices require spin-to-charge conversion. Thus they demand extra circuitry, and the performance degrades due to the conversion efficiency. Behtash Behin-Aein et al. proposed a novel device, know as all-spin logic device [92], which claims to be completely spin based. Unlike the domain wall logic and quantum cellular automata (MQCA), this novel scheme promises selectivity during information propagation even though DW and MQCA do not require spin-to-charge conversion.

![Figure 2.9: Anatomy of a basic all spin logic Device.](image)

The very first schematic of the proposed device is depicted in Fig. 2.9 that shows two single layer nanomagnets (one input magnet and another one is for output) are mounted on a metallic conductor. The magnetization of the output magnet can be switched depending on the polarized current coming from the input magnet through the metallic channel. It is worth mentioning that, the current is polarized according to the information stored while passing through the input magnet.
Sharad et al. proposed an almost similar device, termed as Lateral Spin Valve (LSV) [81]. The device operation is almost similar to that of ASLD. Like ASLD, LSV also uses non-local spins and uses Bennett-clocking during operation.

2.6.5 Nanomagnetic Energy Minimizing Co-processor

Recently Bhanja et al. proposed a novel nanomagnetic co-processor [82] (Fig. 2.10), to solve the quadratic optimization problems arising in computer vision paradigm. The working principle of this co-processor is based on “let physics do the computation” as it uses the energy minimization framework of a 2-D grid of nanomagnets to solve the optimization problems.

Figure 2.10: A sketch of the nanomagnetic co-processor.

A set of omnipresent and non-parallelizable quadratic optimization problems may benefit from harnessing the energy minimization framework of a 2-D array of nanomagnets, named as “Magnetic Energy Minimization Co-Processor (MEMCoP)” [82, 93]. The original problem is mapped into a 2-D grid of interacting magnets where the optimization is accomplished by the relaxation physics of the magnets themselves. In essence, given a specific instance of the problem, we will arrive at a particular magnetic layout (Fig. 2.10), the relaxed magnetic state of which will be the solution to the original problem. The output magnetization of each cell can be in two states: Single Domain state, in which the remanence is the same as the saturation magnetization and Vortex Domain state, which has no zero in-plane remanence rather shows a curling in-plane magnetic configuration [48].
In computer vision domain, there are three steps regarding the object recognition: 1) feature extraction, 2) perceptual grouping, and 3) object matching. The first step is done by extraction of local features i.e. edges, points and regions using different algorithms for edge detection. Recognizing the salient edge segments of the significant objects is the second phase, perceptual organization. The last one is the object matching between the output of the second step and the database. The quadratic optimization problem in these steps, having a non-convex objective function, is an unconstrained optimization problem. Though feature extraction, in some cases, leverages some hardware platform to be faster, perceptual grouping still relies on the software for computation which is computationally expensive [94].

2.7 Conclusion

This chapter discussed the spintronic devices for storage and computation and the key physical properties of them. We reviewed the multi-layer structures, especially MTJs, and the concept of TMR. Later we detailed the generation of STT and STT-based writing. Thereafter we discussed domain wall memories as well its basic operations. We then discussed the use of MTJs as spin-torque-nano-oscillator. The all spin logic device is also briefly discussed in this chapter. Finally, we have covered the basic structure and working principle of MEMCoP. In the next chapters, we will focus on the read techniques, and the SOT-mediated reconfiguration of MEMCoP.
Chapter 3: Reading of Magnetic Energy Minimizing Co-Processor

3.1 Introduction

Alternative computing paradigms targeting heavy data applications is benefited from computing in memory framework [95, 96, 97]. In these applications, the advantage of using spin based memories is twofold: non-volatile storage, and in-memory computation. In this chapter, we discuss a novel computing paradigm that demonstrates computationally complex quadratic optimization problems might have a better solution regarding energy efficiency, and speed through a magnet based hardware [98] than the conventional software based approach. Additionally, the output of that framework is independent of the problem size. The key principle of this hardware is “Let Physics do the computation” as it harnesses the innate energy minimization phenomenon of a grid of nanomagnets to solve quadratic optimization problems, often arose in computer vision domain.

This form of non-Boolean computing impacts a wide range of critical and pervasive application domains based on the Energy Minimization-based optimization framework. Most conventional approaches to solve such problems rely on search-based simulated annealing, neural networks, and genetic algorithms. Apart from a few algorithmic approaches like graph-cuts, these methods are computationally hard. A few example application domains where this work will accelerate computation are: (a) protein-folding problems and drug discovery [99], in social media [100, 101], in error correcting codes [102], in support vector machines [103] and in inferencing Markov Random Fields (graphical Probabilistic Models) [104, 105, 106, 107, 108]. In this chapter\(^{3,1}\), we discusses the read mechanisms of a magnetic hardware solving non-Boolean problems.

3.2 Quadratic Optimization via Energy Minimization of Nanomagnets

Quadratic optimization is a classical combinatorial optimization problem, where we want to minimize a quadratic function over the unknown variables. Let us denote the \( n \) unknown variables by a vector \( \mathbf{x} \) of dimension \( n \). The solution is constrained such that the magnitude of the vector is 1, i.e., it lies on a \( n \)-dimensional hypersphere. The problem is formally expressed as \( \min \mathbf{x}^T \mathbf{A} \mathbf{x} = \min \sum_i \sum_j a_{ij} x_i x_j \) such that \( \mathbf{x}^T \mathbf{x} = \sum_i x_i^2 = 1 \) where \( \mathbf{A} \) is a real symmetric matrix of size \( n \) by \( n \). In the discrete version of the problem, the components of the vector \( \mathbf{x} \), \( x_i \)'s, are constrained to be \( \pm 1 \).

It is well-known that this optimization is NP-hard [109, 110]. Different heuristic or approximation algorithms are usually based on continuous relaxations of the original discrete problems in which vector-valued binary variables are replaced by continuous variables. From the solution of the continuous problem, the discrete solution is derived. We outline one such solution that we can implement using magnets.

3.2.1 A Case Study-Perceptual Grouping in Computer Vision

There are many problems in computer vision that require binary quadratic optimization such as motion segmentation [111], correspondences [112], figure/ground segmentation [113], clustering [114], grouping [115], subgraph matching [116], and digital matting [117]. For demonstration purposes, we focus on one such vision problem, namely that of feature grouping for object recognition [115].

Given a set of \( n \) features such as straight lines, a type of low-level feature, the task is to find the subject of feature that could possibly come from the same object. Let \( x_i \) denote the importance or saliency of the \( i \)-th line; larger values denotes importance. Between every pair of edge line we associate affinity energies, \( A_{ij} \) to capture the perceptual saliency of the relationship between them [115].

\[
A_{ij} = \sqrt{l_i l_j} e^{-\frac{o_{ij}}{\max(l_i, l_j)}} e^{-\frac{d_{min}}{\max(l_i, l_j)}} \sin^2(2\theta_{ij}) \quad (3.1)
\]

where \( l_i \) and \( l_j \) are the lengths of the \( i \)-th and \( j \)-th features, \( o_{ij} \) is the overlap, \( \theta_{ij} \) is the angle, and \( d_{min} \) is the minimum distance between the two straight lines. If two straight lines are parallel to each other then they are likely to belong to one object and hence the affinity should be high. Similarly, lines that are close together are more likely to be associated together. The quantitative
forms of the affinity function vary in different implementation, but qualitatively they capture similar aspects.

\[
\sum_{i=1}^{N} \sum_{j=i+1}^{N} a_{ij}x_i x_j + \lambda \sum_{i=1}^{N} x_i + \kappa
\]  

(3.2)

The goal is to find a group, \( x \), such that total affinity energy in Eq. 3.2 is maximized. \( N \) is the total number of edge segments. \( \lambda \) takes the value of \(-1\). \( \kappa \) is the number of edge segments in the salient group. This is a quadratic optimization problem.

3.2.2 Total Magnetic Energy in the Magnetic System

The total magnetic energy in the magnetic system, can be calculated from the summation of all the magnetic coupling energies between each other and summation of the internal magnetic energy of all the nanomagnetic disks. The total magnetic energy of the magnetic system with \( N \) nanomagnetic disks can be expressed as:

\[
E_{\text{total}} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} E_{ij} + \sum_{i=1}^{N} E_i
\]  

(3.3)

where \( E_{ij} \) is the magnetic coupling energy between the \( i^{th} \) and \( j^{th} \) nanomagnetic disk and \( E_i \) is the internal magnetic energy of the \( i^{th} \) nanomagnetic disk.

\[
E_{ij} = \gamma e^{-\sigma_{ij}} \vec{S}_i \cdot \vec{S}_j
\]  

(3.4)

\[
E_i = \beta |\vec{S}_i| + \omega
\]  

(3.5)

\[
E_{\text{total}} = \gamma \sum_{i=1}^{N} \sum_{j=i+1}^{N} e^{-\sigma_{ij}} |\vec{S}_i| \cdot |\vec{S}_j| + \beta \sum_{i=1}^{N} |\vec{S}_i| + N \omega
\]  

(3.6)

3.2.3 Mapping the Optimization Problem to the Magnetic System

The similarity between Eq. 3.2 and Eq. 3.6, which is exploited to map quadratic optimization problems into magnetic layouts using 2-D visualization techniques e.g. multi-dimensional scaling (MDS). The affinity between two line segments is inversely proportional to the distance between
the nanomagnets representing them. Hence, mapping put magnetic dots representing line segments from same objects close to each other, and different object far away. Fig.3.1 depicts the steps involved in determining the salient edges using the nanomagnetic co-processor. The compiler maps each edge segment to a single nanomagnet in the nanomagnetic co-processor grid. The steps of mapping are:

1. Step 1: Edge detection, affinity matrix calculation, multidimensional scaling and mapping of features (edge segments) to nanomagnets.


4. Step 4: Determine the salient features (edge segments) by back tracing the mapping of the single-domain computing nanomagnets with features (edge segments) [82].

After computation, the final magnetization states of all the computing nanomagnets are identified. As each computing nanomagnet represent an edge, the single domain nanomagnets are back traced to corresponding edges and are identified as salient edge segments.

Figure 3.1: Steps involved in problem mapping into nanomagnetic co-processor.
3.2.4 Translating Output of the Problem from Magnetic States

In case of a circular magnetic nanodisk, the diameter and thickness are the key factors that determine the magnetic state. Literature [48] has established the dependency of magnetization on these parameters. One can understand this relation by considering D as the diameter, and t as the thickness of a circular nanomagnet. Formation of a single domain or multidomain can be explicitly interpreted by the aspect ratio, \( m = \frac{D}{t} \). \( m < 1 \) ensures that the magnetic disk only has two (single domain or vortex) states. A phase boundary between the single domain and vortex domain was outlined in [118] for a 2D grid of closely spaced nanomagnets as shown in Fig. 6.2a. It is evident experimentally from Fig. 6.2c that, one nanomagnetic disk can exhibit in-plane single domain behavior when interacting with neighboring magnetic cells, and vortex domain behavior when isolated [119].

![Figure 3.2: (a) Phase plot; (b) Cell dimension and distance between them; (c) Closely placed nanomagnets behave as single domain, otherwise vortex domain.](image)

3.3 Electrical Reading of the Output

The resistance differential between these states aids in detecting the magnetization of the cells, in other words, the output of the problem. However, absence of shape anisotropy in the circular nanomagnetic cells substantially lowers the resistance values, thus making the read process more challenging. We investigated three different methods for finding an efficient mechanism to read the magnetic cells, and studied the variation of resistances of the cells due to the fabrication imperfection. Since the resistance difference between the magnetic states, and consequently the sense margin is low, maintaining a higher accuracy is challenging. Moreover, non-destructive reading,
and the smaller resistances potentially lower the sense margin making the reading mechanisms vulnerable to the process variations. To make the process more resilient, we proposed an additional pre-amplifier, and with this we demonstrated that the sense margin can be improved by at least 73%.

3.3.1 Structure and Resistance Modelling of a Cell

Each cell of the coprocessor has a trilayer configuration: free layer, NiFe(10nm)/insulator, MgO (1nm)/fixed layer, Co/Pd (6nm). These cells are engineered in a way so that the fixed layer will have tilted anisotropy. The magnetic polarization of the fixed layer was chosen to be tilted by 45° as when clocked, the free layer magnetization can unwind at the saddle point, y-axis. The tilted fixed layer polarization has both inplane and perpendicular components which yield the effective neighboring interaction and ease the TMR based read. The fabrication was carried out by depositing permalloy as circular nanomagnetic disks with critical dimension on Si wafer by a standard photo-lithography process.

![Figure 3.3: (a) Shape variation during fabrication; (b, c) height variation;(d) resistance variation due to the shape and geometry.](image-url)
Resistance Modeling: The resistance of a voxel of dimension of dx, dy and dz is \( r = \rho dz / dy dx \) when the current flows into z direction. The AMR is given by \( AMR_c \ast \vec{M}_{z1} \ast \vec{M}_{z1} \) and the GMR is given by \(-1/2GMRC \ast (\vec{M}_1 . \vec{M}_2)\). The resistance is calculated as \( R = r(1 + AMR + GMR) \). The whole resistance calculation is done in LLG micromagnetic simulator [120]. The proof of concept is experimentally demonstrated by fabricating a magnetic layout[82] in local fabrication facility.

We analyzed the Magnetic Force Microscopy (MFM) images of the layout to get the geometric profile of the nanomagnets. We recalculated the resistance variation of the cells by plugging the imperfections into the simulator to have the deviated resistances. The standard deviations are 7.81% and 5.93% in case of single domain and vortex domain state.

Fig. 3.3(a) shows the MFM picture of a distorted cell along the the bitmap equivalent of the cell as the input to the simulator. Fig. 3.3(b,c) refers the height variation of the cell. The summary of the resultant change in the resistance of the cells is depicted in Fig. 3.3(d). Therefore, the read mechanism is designed in such way so that it can handle the process variation effectively.

### 3.4 Reading Schemes

We experimented on three different reading mechanisms to detect the magnetic states of the cells: i) Reference resistor based reading, ii) Reference cell based reading, and iii) Differential reading. We simulated the circuits using 22 nm PTM models [121]. Both mechanisms share a general framework showed in Fig. 3.4 depicting the circuitry for reading the magnetic state of the cells. \( M_p \) is the access transistor of the computing nanomagnet, labeled as \( NM_1 \), whose magnetic state is to be detected. As discussed earlier, differential reading scheme exploits a non-participating nanomagnet, \( NM_2 \), which is always in vortex state. This reference nanomagnet has the access transistor labeled as \( M_n \). This vortex state reference nanomagnet along with its access transistor is replaced by a reference resistor, \( R_{ref} \) (shown in Fig. 3.4) in reference resistor based reading scheme. Intuitively, the value of this reference resistor should be the average of the vortex domain state resistance and the single domain state resistance in order to achieve highest sense margin. The access transistors are kept always ON by keeping signal \( \phi_1 \) always high (1V in our simulation) during the entire read operation. In general, the whole reading process is accomplished in two consecutive phases: a Pre-charge phase followed by a Sensing phase. In order to achieve robustness, a symmetry
between the arms of the read circuit is maintained so that equal amount of current passes through the computing cell and the reference cell/resistor during the pre-charge phase. During the Pre-charge phase, the $\phi_2$ signal is pulled down to low voltage to turn off transistors, $M_3$ and $M_4$. The active low signal, $\phi_3$, is pulled down to assist in fast pre-charge of nodes, X and Y, to $V_{DD}$. Signal $E_q$ is raised high to equalize the voltage at nodes X and Y through transistor $M_{eq}$. During the Sensing phase, $\phi_2$ is raised to a low voltage, say $V_{\text{read}}$ (0.5V in our simulation), for applying a low voltage bias on the output MTJs; thus it does not affect the magnetoresistance of the computing cell [122].

Figure 3.4: Basic read circuit. Left arm contains the participating cell (single domain or vortex domain) while the right arm is connected to a resistor (in reference read), or a non-participating cell (in differential read).

With $E_q = 0$ and $\phi_3 = 1$, voltage differences based on the magnetic states start to grow at nodes X and Y due to differential current, from Eqn. 3.7, through $M_5$ and $M_6$ as
\[ \Delta I = \frac{V(R_1 - R_0)}{R_0 R_1} \]  

(3.7)

Afterwards, this differential voltage is sensed by a comparator to detect the magnetic state of the computing cell by setting the output "High" or "Low". We also studied the efficacy and robustness of these read mechanisms. We evaluate power dissipation, read circuit delay and degree of variation tolerance. We also considered the worst case scenario, when the \( R_0 \) increases and \( R_1 \) decreases. In worst condition, \( R_0 \) and \( R_1 \) will be \( R_0(1 + \sigma_0) \) and \( R_1(1 - \sigma_1) \) respectively where \( \sigma_0 \) and \( \sigma_1 \) are the standard deviation of the resistances, the circuit can handle accurately. In this condition, Eqn. 3.7 will be transformed into

\[ \Delta I' = V\left(\frac{1}{R_0(1 + \sigma_0)} - \frac{1}{R_1(1 - \sigma_1)}\right) \]  

(3.8)

Figure 3.5: Variation tolerance of the circuit in reference resistance mechanism.

3.4.1 Reference Resistor Based Reading

In the scheme, a predefined resistor (marked with dashed line in Fig. 3.4) having average resistance of single domain state and vortex domain state resistances is utilized to compare with the computing magnet’s resistance. The circuit set up in reference resistor based reading is as below:

1. Resistance of \( NM_1 \): Computing nanomagnet, \( R_0 \) or \( R_1 \)
2. Resistance at reference arm: \( R_{ref} = \frac{(R_0 + R_1)}{2} \)

Since the resistances of vortex state and single domain state are found 2850\( \Omega \) and 3200\( \Omega \) respectively, in simulation we set the resistance of the reference resistor is 3000\( \Omega \).

The response of this reading framework based on the sense margin and output accuracy is depicted in Fig. 3.5. Though this scheme is a very simplistic approach, due to very low sense margin its capability to cope with resistance variation is limited. Our simulation indicated approximately 6.23\% variation tolerance of this reading scheme. In addition to that, requirement of an external resistance often compromises the thermal robustness of the circuit.

3.4.2 Reference Cell Based Reading

This reading technique shares the same concept of previous method of having an average resistance at the computing arm except, instead of an external resistance this mechanism uses a parallel-series network of prefabricated non-computing nanomagnet. The network has the structure, \( 2n \times 2n \), where \( n \epsilon N \) determines the number of non-computing nanomagnets used in the network. In each column of this configuration, a single domain nanomagnet and a vortex domain nanomagnet are placed one after another intentionally to have average resistance of them. This type of configuration ensures three major advantages: i) exclusion of the reference resistor making the read process less heat susceptible, ii) making the reading process more robust by including the effect of process variation in the reference arm, and iii) the actual value of single domain resistance and vortex domain resistance is unnecessary to determine. The resistance profile in both arms looks like

1. Resistance of \( NM_1 \): Computing nanomagnet, \( R_0 \) or \( R_1 \)

2. Resistance at reference arm: \( R_{ref} = \frac{(nR_0 + nR_1)}{nN} \)

Despite its advantage in handling the process variation more efficiently, the major drawback of this framework is the number of non-computing nanomagnets increases squarely. Consequently the power consumption and delay increase non-linearly. In our study, a network having \( n = 1 \) exhibits slightest better performance in terms of accuracy to the networks having \( n \geq 2 \) in handling the process variation.
3.4.3 Differential Reading

In differential reading, our objective is primarily twofold: i) improving the sense margin, and ii) replacing the external resistance required in reference resistance based reading with a non-participating cell. Our second objective can be achieved by harnessing the inherent property of an isolated nanomagnet.

Figure 3.6: Variation tolerance of the circuit in differential read mechanism for different $R_0$ due to process variation.

Since an isolated nanomagnet with a threshold dimension stays always in vortex state, this phenomenon is exploited in differential reading. Consequently, differential reading can exclude the thermal management of the external resistor. Differential Reading mechanism has the following setup:

1. Resistance of $NM_1$: $(R_0)$ or $R_1$
2. Resistance of $NM_2$: ($R_1$)

The behavior of circuit for different values of resistances, in differential read mechanism, is delineated in Fig. 3.6. Despite the fact that it excludes the external resistance and promises better sense margin, this mechanism fails to detect the magnetic state accurately if the computing nanomagnet has a resistance variation more than 7.31%.

### 3.5 Improving the Sense Margin

Improving the sense margin to make the system more robust is a well defined problem for the memory community. During our experiment, we considered several techniques [123, 124, 125] proposed in recent years to have better sense margin.

![Bias Circuit](image)

**Bias Circuit**

Figure 3.7: Pre-amplifier stage of the comparator; a bias circuit is designed to provide biasing in amplifier circuit [126].

However, a large number of these techniques enact methodologies [124] that increase the current flow through the nanomagnets which is undesirable in our case. Some of the methods use destructive reading [123] which in turn increases the power and the delay, and were thus not considered. To improve the sense margin, we proposed an additional pre-amplifier stage (shown in Fig. 3.7) that will feed the comparator with larger sense margin. Bias circuit is designed to provide biasing in amplifier circuit [126]. However, this additional stage will increase the power dissipation as well
as the delay. The pre-amplifier is a differential amplifier with two NMOS transistors ($M_{10}$ and $M_{11}$) as the load. These load transistors are fed by the $V_x$ and $V_y$. These two signals are low enough to keep the load transistors in triode region to get different resistances, thus different gain. $M_{12}$ is used for biasing the amplifier circuit. Since the problem is mapped into circular nanomagnets, there will be no shape anisotropy, which makes the resistances susceptible to process variation. Eventually, the sense margin will not be that high like STT-MRAM, thus needs comparator with very high sensitivity.

Table 3.1: Current requirements for reading the cells.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Current ($\mu$A)</th>
<th>Current($\mu$A)</th>
<th>Current($\mu$A)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reference Reading</td>
<td>Differential Reading</td>
<td>Reference Cell Reading</td>
</tr>
<tr>
<td>Pre-charge</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>Sensing</td>
<td>Vortex</td>
<td>28.7</td>
<td>22.2</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>33.1</td>
<td>31.8</td>
</tr>
</tbody>
</table>

The additional circuitry (in the blue box in Fig. 3.7) is designed to provide a self-bias for the amplifier. We have analyzed a good number of different implementations of the pre-amplifier [126]. For this case, this design served the best interest of ours. This additional amplifier stage will lessen the necessity of using bulky PMOS transistors $M_5$ and $M_6$ of the actual reading circuit in Fig. 3.4, the providers of pre-charge and sense current. Table 3.1 shows the current requirement for the reading schemes. These currents are low enough to change the magnetization state of the nanomagnets, hence the reading mechanism does not affect the output magnetization state.

The differential reading mechanism requires comparatively less current while sensing the magnets due to the lower resistance in the reference arm. Table 3.2 shows a comparative analysis of the reading mechanisms with and without the pre-amplifier. The decision regarding the inclusion of the pre-amplifier will be determined by the problem specifications. There will be a trade-off between the power and delay with the output accuracy, though the additional amplifier increases power consumption by at most 9%. The output delays by at most 4ps while the signal $\phi_2$ and $\phi_3$ have been changing. Thus, there should be at least 4ps delay between the $\phi_2$ and $\phi_3$. As the pulse duration is in nanosecond, pre-amplifier will not affect the speed of the circuit by great means.
Table 3.2: Sense margin improvement by using pre-amplifier.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Reference Resistance Read</th>
<th>Differential Read</th>
<th>Reference Cell Read, n=1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o Pre-amp</td>
<td>w/o Pre-amp</td>
<td>w/o Pre-amp</td>
</tr>
<tr>
<td>Sense Margin</td>
<td>100%</td>
<td>173%</td>
<td>114%</td>
</tr>
<tr>
<td>Variation</td>
<td>100%</td>
<td>112%</td>
<td>100%</td>
</tr>
<tr>
<td>Variation tolerance</td>
<td></td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Power</td>
<td></td>
<td>109%</td>
<td>95%</td>
</tr>
<tr>
<td>Delay</td>
<td></td>
<td>129%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 3.3 compares the methods considered in this context for a resilient reading scheme. Self-Reference Sensing Scheme[123] proposed by Jeong et al. improves the sense margin significantly, but the reading reading mechanism is destructive, and the delay is unexpectedly high. On the other hand, Trinh et al. devised a Boosted-Voltage Sensing Scheme[125] as opposed to the self-reference sensing to achieve the nos-destructibility. However, this method does not amplify the sense margin, rather boosts the voltage levels of nodes X and Y, and the accuracy improvement is one third of what we have achieved by pre-amplifier. In this context, use of the pre-amplifier outperforms the other two methods in terms of improving sense margin, and accuracy.

Table 3.3: Comparison between different methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Type</th>
<th>Sense Margin</th>
<th>Accuracy</th>
<th>Power</th>
<th>Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-amplifier</td>
<td>ND</td>
<td>73%</td>
<td>13%</td>
<td>11%</td>
<td>31%</td>
</tr>
<tr>
<td>BVS[125]</td>
<td>ND</td>
<td>0%</td>
<td>4%</td>
<td>7%</td>
<td>30%</td>
</tr>
<tr>
<td>Self-reference[123]</td>
<td>D</td>
<td>53%</td>
<td>9%</td>
<td>21%</td>
<td>113%</td>
</tr>
</tbody>
</table>

3.6 Conclusion

This work discusses the challenging problem of reading the results and presents a proof of concept that the results are readable. This is an important problem that plagues the entire nanocomputing field, and will need solutions like those addressed in this work. The next chapter details the process of developing a programmable magnetic grid using spin-orbital torque (SOT).
Chapter 4: Programmable Magnetic Grid

4.1 Introduction

An efficient way to compute logic in memory or solve a set of non-Boolean problems is to use dipolar coupling of nanomagnets. MEMCoP is an experimentally demonstrated co-processor that can effectively solve perceptual grouping problem in computer vision as a proof of concept by harnessing the neighbor interaction. To achieve parallelism in the layout and to employ the hardware for different applications require a programmable grid of nanomagnets in a sense that any disk if not participating does not influence the magnetization of the computing nanomagnets. The key idea here is to counterbalance the dipolar field between the neighboring cells. We explored different techniques to achieve the programmability which includes applying spin orbital torque (SOT). Due to the spin Hall effect, polarized spins accumulate at the interface between the heavy metal and the ferromagnet, exerting a torque on the magnetization [127]. SOT is preferred over other magnetization control methods, such as an external Oersted magnetic field or spin transfer torque (STT), due to its higher energy efficiency [128, 129].

In this chapter, we investigate selectively switching nanomagnetic elements in a dipole-coupled array from computing to non-computing states using SOT-induced magnetic dynamics [130]. A dipole-coupled array of nanomagnets is characterized, and the dynamic response of these nanomagnets to an applied SOT current with varying ramp rates is analyzed, with the goal of achieving programmability of individual elements within the array.

4.1 Parts of this chapter were published in IEEE Transactions on Magnetics (2020), "Spin-Orbit Torque and Dipole Coupling for Nanomagnetic Array Programmability.", Nance, John A., Kawsher A. Roxy, Sanjukta Bhanja, and Greg P. Carman. Permission attached in Appendix A.
4.2 Theoretical Analysis

We are modeling pseudo-amorphous CoFeB, in which the exchange length (3-4 nm) is larger than the as-deposited grain size [131, 132]. As such, we assume that the exchange interaction overrides the magnetocrystalline anisotropy (MCA) yielding a magnetic element without significant long range MCA. Additionally, we assume absence of an external magnetic field and neglect thermal noise. It is important to address the effects of thermal noise in magnetic systems, especially those relying on dipole coupling, but the aim of this work is to serve as a proof of concept for SOT as a programming method under near ideal conditions. Future studies incorporating a stochastic thermal field will be necessary, but those efforts are beyond the scope of the current work. In this study, we focus on circular nanodisks to avoid any shape anisotropy effects. The micromagnetic model is represented by the Landau Lifshitz Gilbert (LLG) equations with an SOT term added:

$$\frac{\partial \vec{m}}{\partial t} = -\mu_0 \gamma (\vec{m} \times \vec{H}_{\text{eff}}) + \alpha (\vec{m} \times \frac{\partial \vec{m}}{\partial t}) + \frac{j_c \theta_{SH}}{2eM_s d} (\vec{m} \times (\vec{\sigma} \times \vec{m}))$$

(4.1)

where $\vec{m}$ is the normalized magnetization, $\mu_0$ is the vacuum permeability, $\gamma$ is the gyromagnetic ratio, $\hbar$ is the reduced Planck constant, $e$ is the elementary charge, $j_c$ is the SOT current density, $\vec{\sigma}$ is the direction of polarized spin accumulation, $\alpha$ is the Gilbert damping factor, $d$ is the thickness of the magnetic layer, $M_s$ is the saturation magnetization, and $\theta_{SH}$ is the spin hall angle. The material used in all studies is pseudo-amorphous CoFeB, with a room temperature exchange stiffness of $2.0 \times 10^{11}$ J/m, saturation magnetization of $1.2 \times 10^6$ A/m, and a Gilbert damping of 0.01.

4.3 Coupling Energy Estimation

The magnetic states of two-coupled nanomagnets are governed by the following Landau-Lifshitz-Gilbert (LLG) equation (Eq. 4.2) [60, 133]:

$$\frac{d\vec{M}_1}{dt} = -\gamma M_s \vec{M}_1 \times (\vec{H}_{\text{eff}} - \frac{\alpha}{\gamma M_s} \frac{d\vec{M}_1}{dt})$$

(4.2)

where $\vec{M}_1, \vec{M}_2, M_s, \alpha, \gamma$, are the magnetization of the adjacent nanomagnets, saturation magnetization, Gilbert damping coefficient and gyromagnetic ratio. The effective field, $\vec{H}_{\text{eff}}$, on $i^{th}$ magnetic disk can be defined as the sum of applied external magnetic field, demagnetization field evolved
from itself and field originating from the interaction with neighbor disks \[134\]

\[
\vec{H}_{eff} = \vec{H}_{ext} + \vec{H}_{demag} + \vec{H}_{inter}
\]

where

\[
\vec{H}_{inter} = \sum_{i,j} \vec{C}_{ij} \vec{M}_j
\]

The coupling coefficient \(\vec{C}_{ij}\) between \(i^{th}\) and \(j^{th}\) magnets can be calculated by dipole approximation \[134\] as

\[
\vec{C}_{ij} = \frac{V_i}{4\pi r_{ij}^3} \begin{bmatrix}
3\hat{r}_z^2 - 1 & 3\hat{r}_x\hat{r}_y & 3\hat{r}_x\hat{r}_z \\
3\hat{r}_y\hat{r}_x & 3\hat{r}_y^2 - 1 & 3\hat{r}_y\hat{r}_z \\
3\hat{r}_z\hat{r}_x & 3\hat{r}_z\hat{r}_y & 3\hat{r}_z^2 - 1
\end{bmatrix} \tag{4.3}
\]

where \(V_i, r_{ij}\) are the volume of the magnets and the displacement vector from \(i^{th}\) magnet towards \(j^{th}\) magnet, and denoted by \(\hat{r}_{ij} = (\hat{r}_x, \hat{r}_y, \hat{r}_z)\). From Eq. 4.3, one can easily relate the proportional relation between the volume of the magnets and the coupling coefficient. Fig. 4.1 depicts the magnetic state of a system of two nanomagnets with 110 nm diameter and 10 nm thickness with different spacing. Energy of a such system can be approximated by the following equation \[135\].

\[
E_{ij} = \frac{1}{2}\mu_r\mu_0 V_i M_i + \frac{1}{2}\mu_r\mu_0 V_j M_j - \frac{1}{2}\mu_r\mu_0 V_i C_{ji} M_i - \frac{1}{2}\mu_r\mu_0 V_j C_{ij} M_j \tag{4.4}
\]

The four terms of the right side of the above equation correspond to the demagnetization energies of two magnets, and the coupling energies between them.

![Figure 4.1: Extension and comparison of simulation results with MFM micrographs.](image-url)
It is worth noting that the coupling energy depends on the relative permeability ($\mu_r$) and the coupling coefficient which in turn depends on the volume of the nanomagnets and the distance between them as stated above. To understand the relation between the coupling energy with these parameters, we simulated multiple grid of nanomagnets with a wide range of diameter, thickness and interelement spacing.

![Diagram of disk dimensions and layouts](image)

**Figure 4.2:** The disk dimensions and layouts associated with the simulations.

The first study illustrated in Fig. 4.2a evaluates the dipole coupling energy of a pair of neighboring disks as a function of spacing and disk size. Coupling energy is defined as the difference between the total energy of the two disk system (Fig. 4.2a), and the total energy of two isolated
disks of the same dimensions and in the same magnetic states. This is mathematically represented as:

\[ E_c = E_{\text{pair}}^{\text{tot}} - 2E_{\text{isolated}}^{\text{tot}} \]  

(4.5)

where the superscript “pair” denotes the two neighboring disks, “isolated” denotes a single disk, and the total energy \( E_{\text{tot}} \) is the integral of total energy density (2) over the entire volume. Eq. 4.5 represents the amount of dipole coupling energy between two neighboring disks as a function of the spacing between the disks [119]. In this study, disk diameters are varied from 110 nm to 150 nm, thicknesses from 4 nm to 15 nm, and disk spacings from 0 nm to 100 nm using a finite-difference cell size of 3 x 3 x 1 nm³.

![Figure 4.3: CoFeB dipole coupling energy as a function of (a) disk spacing and (b) thickness.](image)

Fig. 4.3a shows a plot of coupling energy \( E_c \) as a function of disk spacing for two aligned CoFeB disks of thickness 4 nm, for four different disk diameters from 110 nm to 150 nm. The coupling energy decreases as spacing increases, for all diameters, and approaches zero as spacing increases to 100 nm. At smaller spacings, the effect of disk diameter becomes more apparent as larger diameters result in larger coupling energy. Fig. 4.3b shows a plot of \( E_c \) as a function of disk thickness for two aligned disks with 30 nm spacing and four different diameters from 110 nm to 150 nm. The coupling energy increases with disk thickness for all diameters, and larger disk diameter results in larger coupling energy. These plots provide information regarding the
relationship between dipole coupling energy and disk dimension and spacing that is used in choosing a disk size and spacing for subsequent studies and future computational platforms.

Figure 4.4: The phase boundary between single domain and vortex states as a function of disk diameter and thickness for both an isolated CoFeB disk, and for neighboring CoFeB disks is shown.

4.4 CoFeB Phaseplot

A parametric study conducted on isolated disks and a pair of disks spaced 30 nm apart is illustrated in Fig. 4.2b. This study is used to determine the dimensions at which the difference between the energy $E_{tot}$ of in-plane single domain states and vortex states is minimized, such that both are potentially stable states [48]. The disk thicknesses are varied from 1 nm to 12 nm and diameters ranged from 60 nm to 320 nm. In these studies, each disk is initialized with an out-of-plane magnetization then allowed to relax for 0.7 ns to determine its stable state.

A phase boundary between the single domain and vortex domain was outlined for a 2D grid of closely spaced CoFeB disks as shown Fig. 4.4. Any disk with the dimension lying on the line has the vortex state as the energy minimum state, but the energy difference with single domain state is the minimum here. This small energy barrier can be effectively manipulated by dipolar coupling.
energy to switch between the ground states. A nano-magnet with thickness of 4 nm and diameter of 150 nm can be in both single domain or vortex domain state based on the external parameters. It is evident experimentally from Fig. 4.4 that, one nanomagnetic disk can exhibit in-plane single domain behavior when interacting with neighboring magnetic cells, and vortex domain behavior when isolated.

4.5 SOT-based Reconfiguration in a Magnetic Array

Figure 4.5: SOT current density applied to disk 3 is plotted over time for the (a) low slope case, and (b) the high slope case; (c) The x-component of magnetization for each disk in the five-disk arrangement is plotted for the low slope case and, (d) high slope case; Snapshots of magnetization at 0.3, 0.7, and 4.0 ns are shown for the (e) low slope case and, (f) high slope case.
Fig. 4.5 compares the outcomes of applying two different SOT current density ramps to disk 3 in a five-disk arrangement while simultaneously clocking disks 2 and 4. Fig. 4.5a and Fig. 4.5b plot the SOT current density over time for two cases, a $1.33 \times 10^8$ A/cm$^2$ns (low) slope case to reach $4.00 \times 10^8$ A/cm$^2$ at 3.3 ns in Fig. 4.5a, and a $40.00 \times 10^8$ A/cm$^2$ns (high) slope case to reach $4.00 \times 10^8$ A/cm$^2$ at 0.4 ns in Fig. 4.5b.

Fig. 4.5c and Fig. 4.5d plot the x-component of magnetization for each disk in the five-disk arrangement as a function of time, for the low slope (Fig. 4.5a) and high slope (Fig. 4.5b) cases respectively. The solid line represents disk 3, the dashed line represents disks 2 and 4, and the dotted line represents disks 1 and 5. In Fig. 4.5c, each disk is initialized in a single domain state pointing in the positive x-direction. At 0.3 ns, the x-component of disks 2 and disk 4’s magnetizations drop to zero as they are clocked to the z-direction. Following the clocking of disks 2 and 4, all five disks switch to a single domain state pointing in the negative x-direction by $t = 1.5$ ns, representing their stable state. In Fig. 4.5d, each disk is initialized in the same manner as described for Fig. 4.5c. However, following clocking, disk 3 enters an oscillating state while disks 1 and 5 remain in the positive x-direction. Then, from $t = 2.0$ to $t = 2.5$ ns, disks 1, 3, and 5 switch to the negative x-direction and remain in this direction while the x-components of disks 2 and 4 oscillate around zero. The difference between Fig. 4.5c and Fig. 4.5d cannot be related to current density amplitude since both are the same. The differences between the two cases must be related to the high SOT current density slope used, inducing a dynamic magnetic response. In the high slope case, the current density reaches $4 \times 10^8$ A/cm$^2$ quickly, which brings disk 3 into an oscillating state because the SOT counteracts the effective dipole field from the other disks. In this state, disk 3 is “decoupled” and moves relatively independently of the other disks. On the other hand, in the low slope case, the slowly increasing SOT current density switches disk 3 to the negative x-direction and disks 1, 2, 4, and 5 follow shortly after due to dipole coupling in the system.

Fig. 4.5e and Fig. 4.5f show snapshots of the x-component of magnetization of each disk at times 0.3 ns, 0.7 ns, and 4.0 ns for the low slope and high slopes cases respectively. In these snapshots, the color white represents a magnetization in the positive x-direction, and black represents a magnetization in the negative x-direction. The arrows correspond to the average magnetization direction in each disk, where a straight arrow indicates a single domain in-plane state, and a circular arrow represents a vortex state. Finally, the snapshots at 4.0 ns show that in the low slope
case, disks 2 and 4 end up in single domain in-plane states, whereas in the high slope case, disks 2 and 4 end up in vortex states. In the low slope case, disks 1, 3, and 5 are pointing the same direction and move together, such that in-plane states pointing in that same direction are stable for disks 2 and 4. Vortex states are more stable in the high slope case because while disk 3 is oscillating independently of disks 1 and 5, its dipole field is continuously changing direction such that an in-plane state in any single direction is not stable. These results show that the final states following clocking in a dipole-coupled array depend not only on the amplitude of SOT applied, but also the slope with which it is applied.

4.6 A Programmable Magnetic Grid

Though this work focuses on SOT-based programmability in a 1D array of nanomagnets, as a proof of concept, the same mechanism is valid in 2D grids. This is because the non-computing cells relax from the z-axis in the presence of SOT, while the computation is done in the xy-plane. Therefore, the neighbors of non-computing cells are not influenced by those cells. This is true whether the array extends in 1D or 2D. A 2D nanomagnetic grid representing the magnetic energy minimizing co-processor (MEMCoP) device [98] with spin-orbit torque (SOT) programmability is shown in Fig. 4.6. On the left, a top view of a nanomagnetic grid with a hypothetical magnetic layout is shown. Disks marked with an X represent computing elements, while the rest are non-computing elements. The expanded image on the right shows the wiring scheme corresponding to this array. Each element of the array has an n-type metal-oxide-semiconductor (NMOS) access transistor and is wired to a word line (WL), source line (SL), and bit line (BL) for reading and writing as in a traditional random-access memory (RAM) array. These wires are made of a heavy metal such as platinum or tantalum for application of SOT. A procedure to achieve the layout shown on the left is outlined here. Initially, each element in the array is in an in-plane state due to dipole coupling. Once the desired magnetic layout is determined, the computing elements (i.e. those marked with an X in Fig. 4.6) are clocked to the out of plane direction using polarized current applied through their access transistors. Then, once the clocking current is removed, SOT current is simultaneously applied to the non-computing cells through the BL.
As stated earlier, this framework targets QUBO problems that can be mapped onto a 2D grid of nanomagnets in the xy-plane, while the z-direction is used to interface with peripheral circuits. Expanding this architecture to harness dipole coupling in a 3D arrangement could allow for more complex optimization problems to be studied. However, interfacing with peripherals and accessing the magnetic ground states would be challenging. We refer the interested reader to these articles [136, 93, 137] to see a detailed description of problem mapping and extraction of results from the 2D magnetic layout.

4.7 Conclusion

In conclusion, it was found that dipole coupling energy between adjacent CoFeB nanodisks decreases as disk spacing increases and increases with disk thickness. This relationship governs the critical SOT current density required to switch disks within a dipole-coupled array, as it was observed that critical SOT current density decreases as disk spacing increases in both three disk and five disk arrangements, approaching the isolated disk’s critical current density at large disk spacings. Additionally, a phase plot showing the dimensions at which the difference between the energies of in-plane states and vortex states is minimized was produced. For disks with dimensions chosen
along this phase boundary, it was found that by varying the slope of a SOT current ramp applied to a control disk, the stable states of adjacent disks could be chosen. At very high current slopes, the adjacent disks prefer vortex states, but at low current slopes, they prefer in-plane states. This phenomenon was observed at a SOT current density an order of magnitude below the previously determined critical current density, which suggests that by utilizing SOT current dynamics, the amplitude of current necessary to switch these disks between computing and non-computing states is much lower than expected. Thus, selective programmability of disks within a dipole-coupled array was demonstrated using SOT current slope as the control method. The next chapter focuses on a novel ‘Transverse Read’ in domain wall memories.
Chapter 5: Transverse Read in Domain Wall Memories

5.1 Introduction

Spintronic memories e.g. Spin-Transfer Torque Magnetic RAM (STT-MRAM) [138], and Domain Wall Memory (DWM) [139, 66] are considered as promising replacements of charge-based memories. These magnetic memories store data by the magnetization of the structure instead of charge making it non-volatile. Moreover, domain wall memories offer high density (4× higher than STT-MRAM) and similar standby power and access time compared to other emerging memories. These promises make DWM a perfect replacement of SRAM based cache [140]. Interestingly, the evolution period of magnetic memories has also experienced a hike of research on in-memory computing and hardware-based neuromorphic computing framework, and magnetic memories are well-suited in this paradigm [98, 133]. Often these frameworks demand a highly dense, sparse non-volatile and ultra-low power storage, therefore, domain wall based memories (DWM) such as “Racetrack” memories can be a suitable alternative [141, 81, 142]. Weights are stored in many of these computing architecture as one hot encoding, and periodically requires to calculate number of ‘1’s or to check parity [143] for updating or testing. Another crucial and exciting application domain of DWM is hardware security, where instead of actual data, parity bit(s) carry more importance [144]. These algorithms often scans the number of ‘1’s (or ‘1’s) instead of the whole data. Since DW memories are one of the key players in futuristic memory applications, these area will harness the ultra-density of DWM exceedingly. In this chapter, we discuss the transverse read operation in both PMA and IMA nanowires.

5.2 Domain Wall Memory Nanowire

Fig. 5.1 depicts the schematic of a typical DWM cell consisting of a ferromagnetic nanowire with multiple access ports for reading and writing, and several access transistors. Shift-lines (SLs) control the direction of shifting; Bit-lines (BL, BLB) carry the data; word lines (RWL, WWL) enable reading and writing. A domain wall (DW) separates two oppositely magnetized domains and has a non-uniform magnetization. Typically, an access port is a magnetic tunnel junction (MTJ) where the free layer is a domain of the nanowire. Several number of pinning sites are introduced in the nanowire to forcefully form domain walls in specific locations. Popular methods of creating pinning sites include the patterning of notches periodically along the nanowire, using specific material composition, and shape and geometry of the nanowire etc. [145, 146]. We enact the first method for pining the walls across the nanowire.

![Anatomy of a typical DWM nanowire. Shift-lines (SLs) control the direction of shifting; bit-lines (BL, BLB) carry the data; word lines (RWL, WWL) enable reading and writing.](image)

5.2.1 DWM Read, Write and Shift Operation

Obviously multiple bits share same access ports in a DWM nanowire, as a result, bi-directional shifting of the domains is often required. Usually, reading the information stored in a DWM element is carried out by moving domains along the wire under an access port, namely read-head, and after reading, shifting in opposite direction is required to restore the original data making the process slower and power hungry. Extra “padding” domains (shown in gray) are required on each side of the data domains to prevent data loss while shifting to an extremity. Shift-based writing has also been proposed to accelerate and reduce energy compared to current-based writes [147]. As shown
in Fig. 5.1, the read/write port has fixed domains of opposing polarization in-plane, but orthogonal to the nanowire, allowing a specific polarization to be shifted into the aligned domain.

Fig. 5.2 shows the signals for read, write and shifting operation. A DWM nanowire is accommodated with bi-directional shifting ability which turns it as an excellent shift register. The leftmost and the rightmost transistors in Fig. 5.1 control the direction of shifting data. A shift operation initiates with turning on any of these transistors, and a current enters correspondingly, generates spin torque, and if the torque is strong enough, domain walls are de-pinned and shifted. Fig. 5.1 also shows a typical write-head of a DWM nanowire. Earlier version of a DWM nanowire includes a current carrying metal like MRAM to write into the domain associated with it. However, this poses the similar shortcomings like toggle MRAM. Emergence of STT led researchers to devise current induced writing of DWM nanowire with the aid of an MTJ [148, 149]. Nowadays, in several architecture, the MTJ based write-head is merged with the read-head, therefore, requires special attention while designing.

The traditional reading process starts with shifting the desired cell under (or on top of) the read-head. After shifting successfully, the access transistor associated with the read-head is turned on by raising the RL signal to Vdd. In the meantime, a smaller voltage ($V_{\text{read}}=0.25\text{Vdd}$) is applied into the bit-line (BL). The amount of current flowing through the read-port to Gnd determine the content of the desired cell.

![Figure 5.2: Signals associated with read, write and shift operation.](image-url)
This process is exhaustive and extremely power hungry in a sense that the domain walls are moved from their original position, and to be brought back to the initial position, thus requires shifting in reverse direction. This also limits the storage capability by half as similar number of cells are left to be empty to hold the data during shift operation. The back and forth shifting often incurs data corruption due to overwriting.

In this paper, we propose a novel transverse read (TR) mechanism to read the number of ‘1’s (or ‘1’s) in the data stored in a racetrack. We exploit the physics of current-driven domain wall motion motivated from the primal theoretical work by Berger et al. [150]. We modeled a DW nanowire with both in-plane and perpendicular magnetic anisotropy (IMA, PMA) to verify the efficacy of TR. We widen the range of nanowire dimension to inspect the scalability of proposed TR. Additionally, we devise a technique to employ TR in parallel to read a longer nanowire in fewer clock cycle to reduce the read access time. Finally, we compare the outcomes of TR method with the conventional reading (CR) while applying in a DWM nanowire with 15 bits. Surprisingly, our experiments shows an achievement of 3000x reduction of power by employing TR instead of CR.

5.3 Transverse Read

In this paper, we propose a novel transverse read (TR) mechanism [151] that provides more global information about the data stored within a nanowire. In particular, by applying a transverse read current $J_{TR} \ll J_S$, where $J_S$ is the shifting current, TR can report the number of ‘1’s (or ‘0’s) among the data stored in a racetrack. To verify the efficacy of our technique, we applied TR in domain-wall nanowires with both in-plane and perpendicular magnetic anisotropy (IMA/PMA), respectively. In both cases, our data shows that the more domains we wish to include in a TR, the sense margin between different quantities of ‘1’s decreases, making the sense circuit sensitivity a key factor of TR.

5.3.1 Motivation

There has been significant effort in designing improved data placement techniques that minimize shifting [152] during a reading process, which has significantly mitigated this concern. While shift minimization has largely addressed the dynamic energy concerns of DWM, shifting has not been
entirely eliminated. Moreover, interest in improved memory density continues to encourage the trend of ultra-miniaturization of memory structures through technology scaling. Scaling in this fashion is well known to introduce fabrication imperfections which can lead to a variety of faults in DWM memories including alignment and pinning faults during shifting. Thus, fault tolerance approaches are critical to make DWM technology mature and commercially viable [153].

Unfortunately, traditional error correcting codes cannot directly correct these faults because of the restricted access to individual domains of the device. Gaining some global insight into what data are stored in the domains, in parallel, can enable key tools to provide fault tolerance [154]. These algorithms often scan the number of ‘1’s (or ‘0’s) instead of the whole data. Since DW memories are one of the key players in futuristic memory applications, these area will harness the ultra-density of DWM exceedingly.

Additionally, interest in both in-memory computing and hardware-based neuromorphic computing has grown significantly and magnetic memories are particularly well-suited for such paradigms [98, 133]. For example, these techniques require a highly dense, non-volatile and ultra-low power storage making DWM an ideal storage choice. System operations may require parity checking during learning or testing. This requires global knowledge of data within the racetrack.

Thus, we also demonstrate how TR can be applied to partial segments of the nanowire, such as from an end to an access point or between access points.

This enables a segmented TR which allows access to all of the bits of an arbitrarily long nanowire in a maximum of three steps while maintaining isolated current paths. Finally, we applied TR to multiple segments of a 16-bit nanowire in parallel and read all the domains in 3 steps.

5.3.2 Theoretical Analysis

This strip shown in Fig. 5.1 is symmetric in a sense that the read port is in the middle of the nanowire and it can store 7 bits. The reading process starts by turning the rightmost transistor off, and turning the leftmost transistor and the transistor connected to the read port on to ensure that the current flows from the left most domain to the read port. Fig. 5.4a shows the resultant structure of the nanowire. A read current I less than the critical current (see Eq. 5.5) for domain wall motion enters into the leftmost domain, and exited through the fixed layer of the read port. The resistance values of the nanowire for different bit patterns are calculated by measuring the
The significant mismatch between the density of states (DOS) of spin-up and spin-down electrons changes the resistance of a ferromagnetic nanowire with the existences of domain walls. The structure of a wall depends on the interplay between anisotropy and exchange energies. The latter one minimizes when neighboring spins align in the same direction. On the other hand anisotropy energy is responsible for favoring the spins to lie in a preferred direction, easy axis. Interaction between these energies results in two categories of wall: transverse, when the exchange energies settle into a minimum state and in which the magnetization always rotates across the wall and, and vortex, when the magneto-static energy increases with the thickness and width.

![Diagram of domain wall structure](image.png)

Figure 5.3: Conduction electrons bend at the wall exerting a force on the wall.

The resistance of a DWM nanowire arises from two physical phenomena: (i) back scattering of electrons at the wall, and (ii) forces on the transmitting electrons through the wall. When a current flows through a ferromagnetic metal, a portion of the polarized electrons back scatters from the wall while trying to enter next domain. The impact of back scattering on the resistance of the nanowire was first investigated by Cabrera and Falicov [155] by calculating the reflection co-efficient. Rest of the electrons transmit through the wall and experience two different phenomena affecting the resistance: (i) electrons deviate from the original direction of flow at the wall, and (ii) evolution of
spin transfer torque (STT) in next domain. The 4 – s conduction band electrons interact with the 3 – d band electrons of magnetic domain generating a exchange interaction torque [156, 157]. At a transverse wall the magnetization flips over a plane parallel to the wall. As a result, the conduction electrons bend at the wall while crossing it (shown in Fig. 5.3). There is a transfer of momentum between the passing electrons and the local magnetic moment of the wall. This transfer process is adiabatic and included in Eq. 5.2. Due to the conservation of energy, the electrons exert a force on the wall to move [158, 77]. However, due to the presence of local strains, inhomogeneities and shape anisotropy, there exists a pinning force \( F_p \) on the wall and can be calculated from

\[
| F_p^x | \leq 2H_c M_s A
\]

for a static wall, where the terms are explained in Table. 5.2.

The critical current \( J_c \) necessary to move the wall can be calculated as

\[
J_c = \frac{2H_c}{a\theta}
\]  
(5.1)

such as

\[
\beta = \frac{M_s R_1}{\rho}.
\]

Since the transverse read should not move the wall, thus the maximum read current is less than \( J_c \). Assuming that the current flow is in +x direction while the magnetic domains are out of plane (+z or -z), the resistivity tensor can be expressed [158] as

\[
\rho_1 = \begin{vmatrix}
\rho & -\beta \rho \frac{M_{s1}}{M_s} & 0 \\
\beta \rho \frac{M_{s1}}{M_s} & \rho & 0 \\
0 & 0 & \rho
\end{vmatrix}
\]

\[
\rho_2 = \begin{vmatrix}
\rho & \beta \rho \frac{M_{s2}}{M_s} & 0 \\
-\beta \rho \frac{M_{s2}}{M_s} & \rho & 0 \\
0 & 0 & \rho
\end{vmatrix}
\]

The resistance of the nanowire due to the presence of domain walls is calculated as

\[
R = R_0 + \Delta R = R_0(1 + \frac{K_c N_w t \beta^2}{L})
\]
5.4 Simulation Setup

In this work, we use a finite difference solver to simulate micromagnetic behavior of a DW nanowire, and calculate resistance values for different combinations of data stored in the nanowire. During the simulation we neglect the thermal fluctuations and assume the absence of an external magnetic field. The micromagnetic model in this paper is represented by Landau-Lifshitz-Gilbert (LLG) equation with the inclusion of current-induced torques [159] (assuming the current flows into $\vec{x}$-direction) as

$$\frac{d\vec{m}}{dt} = -\gamma\vec{m} \times \vec{H}_{\text{eff}} + \alpha\vec{m} \times \frac{d\vec{m}}{dt} - v_j \frac{\partial \vec{m}}{\partial x} + \beta v_j \vec{m} \times \frac{\partial \vec{m}}{\partial x} \tag{5.2}$$

where

$$\vec{H}_{\text{eff}} = \vec{H}_{\text{ex}} + \vec{H}_{\text{d}} \tag{5.3}$$

and

$$\vec{H}_{\text{ex}} = \frac{2A_{ex}}{\mu_0 M_s} \Delta m, \quad \vec{H}_{\text{d}} = -\nabla \phi \tag{5.4}$$

Each symbol of the Eqs. 5.2,5.3,5.4 is explained in Table. 5.2. The first two terms on the right side of the Eq. 5.2 describe the torque generated by the effective field and damping. The next two terms representing the current-induced torques are described in section above and can be expressed as

$$v_j = \eta j$$

such that

$$\eta = \frac{g\mu_B P}{2eM_s}$$

![Figure 5.4: The DWM segment used for simulation.](image)
Fig. 5.4 illustrates the transverse read of a DWM segment consisting of 4 domains. Theoretically, the transverse read is feasible for larger number of domains, however, the shortcomings of the finite difference solver we used limit the number of domains to 4. The reading process starts with turning ‘ON’ the leftmost transistor and the transistor beneath the MTJ. The read current enters from the left side of the segment and exits through the fixed layer. We calculated the resistance of the nanowire for bit patterns by measuring the voltage difference between the leftmost domain and the fixed layer of the read port. In this study, width of the nanowire was varied from 10 nm to 300 nm, and thickness from 1 nm to 40 nm to find the scalability of TR mechanism. For simulation we used a finite difference cell having dimension of $3 \times 3 \times 1 \text{ nm}^3$.

Resistance Modeling: The resistance of a voxel of dimension of $dx$, $dy$ and $dz$ is $r = \rho dx/dydz$ when the current flows into $\vec{x}$ direction. The AMR is given by $AMR_c \ast \vec{n}_1 \ast \vec{n}_2$ and the GMR is given by $-1/2GMR_C \ast (\vec{n}_1 \cdot \vec{n}_2)$. The resistance is calculated as $R = r(1 + AMR + GMR)$.

We considered DW nanowires with both PMA and IMA for the first study (Fig. 5.4a), and only PMA for the longer nanowire (Figs. 5.4(b,c)) as it is independent of magneto-crystalline anisotropy. NiFe and CoFeB-MgO were used for simulating IMA and PMA nanowires. Table. 5.1 lists the material properties of them at room temperature (RT).

<table>
<thead>
<tr>
<th></th>
<th>$A_{ex}$ (J/m)</th>
<th>$M_s$ (A/m)</th>
<th>$\alpha$</th>
<th>$K_{u1}$ (J/m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NiFe</td>
<td>$1.3 \times 10^{11}$</td>
<td>$8.0 \times 10^5$</td>
<td>0.02</td>
<td>$\sim 0$</td>
</tr>
<tr>
<td>CoFeB-MgO</td>
<td>$2.0 \times 10^{11}$</td>
<td>$1.2 \times 10^6$</td>
<td>0.01</td>
<td>$10^6$</td>
</tr>
</tbody>
</table>

Our experiments also revealed all permutations containing same number of ‘1’s do not correspond to precisely the same resistance values. Permutations that place the ‘1’s closer to the fixed layer tend to be slightly higher resistance. Thus, for the configuration in Fig. 5.4, “1000” has a lower resistance than “0001.” Fig. 5.6 averages the resistances for different permutations of ‘1’s.
Table 5.2: Definition of symbols used in this chapter.

<table>
<thead>
<tr>
<th>Terms</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\vec{m}$</td>
<td>Magnetization</td>
</tr>
<tr>
<td>$H_{eff}$</td>
<td>Effective field on $\vec{m}$ including demagnetizing ($\vec{H}<em>d$), anisotropy, and exchange ($\vec{H}</em>{ex}$) fields.</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Gyromagnetic ratio</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Gilbert damping factor</td>
</tr>
<tr>
<td>$g$</td>
<td>Landé factor</td>
</tr>
<tr>
<td>$\mu_B$</td>
<td>Bohr magneton</td>
</tr>
<tr>
<td>$M_s$</td>
<td>Saturation magnetization</td>
</tr>
<tr>
<td>$P$</td>
<td>Conduction electron spin polarization</td>
</tr>
<tr>
<td>$A_{ex}$</td>
<td>Exchange constant</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>Magnetic permeability in vacuum</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Magnetic scalar potential</td>
</tr>
<tr>
<td>$H_c$</td>
<td>Coercive Field</td>
</tr>
<tr>
<td>$A$</td>
<td>Area of the wall</td>
</tr>
<tr>
<td>$a$</td>
<td>Domain wall spacing</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Tangent of the Hall angle of the material</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Resistivity of the material</td>
</tr>
<tr>
<td>$\vec{M}<em>{z1}$, $\vec{M}</em>{z2}$</td>
<td>Magnetization of domains next to the wall</td>
</tr>
<tr>
<td>$L, t$</td>
<td>Length and thickness of the nanowire</td>
</tr>
<tr>
<td>$N_w$</td>
<td>Number of walls along the nanowire</td>
</tr>
<tr>
<td>$K_c$</td>
<td>Material dependent parameter</td>
</tr>
</tbody>
</table>

5.4.1 Transverse Read Current

During a transverse read there is a transfer of momentum between the passing electrons and the local magnetic moment of the wall. This transfer process is adiabatic. Due to the conservation of energy, the electrons exert a force on the wall to move [158]. However, presence of local strains,
inhomogeneities, and shape anisotropy there apply a pining force \( F_p \) on the wall. The critical current \( J_S \) necessary to move the wall can be calculated as shown in Eq. 5.5.

\[
|F_p| \leq 2H_c M_s A, \quad J_S = \frac{\mu H_c}{a R_H}
\]  

(5.5)

The TR current \( J_{TR} \) must be smaller than the critical current \( J_S \) to avoid domain wall motion. From the Eq. 5.5 describes that nanowires with higher coercivity and lower Hall resistance can have a higher \( J_{TR} \), resulting in a better sense margin.

Figure 5.5: A transverse read (a) from the right to the access port and (b) from the left to the access port.

5.4.2 Device Operation During Transverse Read

The nanowire represented in Fig. 5.5 is symmetric in a sense that the read port is in the middle and it can store nine bits. The TR process starts by applying a read current \( J_{TR} \) between the right extremity and the access point [Fig. 5.5(a)]. This is accomplished by activating SL and RWL to create a current through BL to GND with BLB disconnected through the peripheral circuitry (see Fig. 5.1). Similarly a current from the left is possible by activating BLB and disconnecting BL [Fig. 5.5(b)]. Fig. 5.4 shows the resultant structure of the left portion of the TR. The resistance values for different bit patterns are calculated by measuring the voltage difference between the leftmost domain and the fixed layer of the read port. Due to symmetry, similar results are achieve when reading from the right.
5.5 Results and Discussion

Fig. 5.6 shows the average resistance values for all possible cases (5 in this scenario) calculated from the simulation setup shown in Fig. 5.4a. A ‘1’ in the data bit is considered when the domain storing that is directed anti-parallel to the fixed layer magnetization. As a result, increasing number of ‘1’s elevates the resistance level. One salient observation is the resistance values of CoFeB-MgO nanowire for any combination is higher that NiFe counterpart.

The key reasons are twofold: (i) higher resistivity of CoFeB, and (ii) the perpendicular injection of electrons into a domain with respect to the magnetization discussed briefly in section 3. Fig. 5.6 averages the resistances for different combinations e.g. 6 combinations with two ‘1’s, 4 combination with one and three ‘1’s etc. However, all the combinations containing same number of ‘1’s does not correspond to the same resistance values.

Additionally, the sense margin between two consecutive levels of ‘1’s decreases as the number of ‘1’s increases, e.g., it is at a minimum between “0111” and “1111” states. As the TR shares the same current path as the shift operation, care must be taken to ensure inadvertent shifting does not occur. Thus, $J_{TR}$ must remain sufficiently below shifting currents ($J_S$), and from Table 5.3, we see that this is the case for both IMA (42% less) and PMA (48% less).
Figure 5.7: Analysis of variation of CoFeB DW nanowire.

5.5.1 Resistance Deviation Due to Process Variation

To better understand how the resistance for different combination varies, we fabricated CoFeB nanowires having similar dimensions used in simulation. Fabrication was conducted using e-beam lithography and CoFeB was deposited using sputtering and samples were then characterized with a scanning electron microscope (SEM) and an atomic force microscope (AFM) to measure height, length, and width. The SEM and AFM images in Fig. 5.7(a,b) capture the shape and height variations of the nanowires. Table 5.3 reveals the maximum resistance deviation for both nanowires along with the comparative facts between them. Interestingly the deviation in both cases is significant which infers a healthy sense margin is possible among different combinations.
Surely the sense margin between two consecutive number of ‘1’s is lower and our experiment shows it is minimum between ‘0111’ and ‘1111’ states. However, with the read current stated in Table 5.3 the minimum sensing voltages, 93 mV and 115 mV, are well above the challenging paradigm. Since shift operation shares the similar current path like TR, there is a upper limit of read current which is lower than depinning currents ($J_{depin} = 4.3 \times 10^7 A/cm^2$ for NiFe and $6.7 \times 10^7 A/cm^2$ for CoFeB-MgO)

### 5.6 Parallel TRs for Longer Nanowires

DWM nanowires proposed for building memories contain larger numbers of data bits, e.g., 16, 32, or higher. It may be impractical to scale a TR to these numbers of domains. A solution is dividing the nanowire into multiple segments of a length the TR is capable of discerning using multiple access points. We propose a parallel TR technique to access these segments such that the entire nanowire is queried in three or fewer sequential steps. To study the feasibility of applying TR in parallel, we consider a DWM nanowire with 16 domains and 4 read-heads dividing it into 5 segments, shown in Fig. 5.8a. The first and fourth segments from left to right (red arrows) can be read in parallel by raising $\Phi_1$ and $\Phi_4$ to $V_{DD}$ to turn on $M_1$ and $M_4$. Meanwhile $\Phi_2$ and $\Phi_3$ are set to $V_{SS}$ (off) to prevent current flow through the other segments. The read current is provided from SL and BL2 and collected at BLB0 and BLB3. Similarly, the second and fifth segments can be read simultaneously, followed by the third segment.

<table>
<thead>
<tr>
<th></th>
<th>Read current</th>
<th>Max resistance deviation</th>
<th>Min sense margin</th>
<th>Read time</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMA</td>
<td>$2.5 \times 10^7$ A/cm$^2$</td>
<td>3.66%</td>
<td>93 mV</td>
<td>5 ns</td>
</tr>
<tr>
<td>PMA</td>
<td>$3.5 \times 10^7$ A/cm$^2$</td>
<td>3.93%</td>
<td>115 mV</td>
<td>5 ns</td>
</tr>
</tbody>
</table>
Figure 5.8: Parallel TR study. (a) 16-bit DWM segment with four access points for parallel TR; (b) Current evaluation for parallel TR of (a).

We simulated the system from Fig. 5.8a using 22nm CMOS technology. Fig. 5.8b reports of the currents flowing through the first and second segment in first two steps described above. During the first cycle, when Φ₁ is high, a read current \( I_{TR} \) of 36 \( \mu \text{A} \) flows through the first segment. The leakage current through the second segment is in the nA range, which confirms the read current only flows through the desired segment. In the next step, the active signal is pulled down to \( V_{SS} \) while \( Φ₂ \) is raised to \( V_{DD} \) to open the alternative paths. An equal current flows through the second segment demonstrating successful isolation of these paths allowing parallel TRs. If the segments are read serially, all the access ports would share a sense amplifier (SA) [160]. However, the parallel
transverse read process requires additional SAs. In the example, \( \Phi_1, \Phi_2, \Phi_3 \) would share a sense amplifier SA1 and \( \Phi_4 \) would require a second sense amplifier SA2 [161, 162]. Thus, in cycle 1, SA1 would be accessed from \( \Phi_1 \) and SA2 would be accessed from \( \Phi_4 \). In cycle 2, SA1 would be accessed from \( \Phi_2 \), and SA2 would be accessed from \( \Phi_4 \). Finally, in cycle 3, SA1 would be accessed from \( \Phi_3 \) and SA2 would be idle.

5.6.1 Comparison Between Transverse Read and Conventional Read

Table 5.4: Performance comparison between conventional and transverse read for a 16 bit DWM.

<table>
<thead>
<tr>
<th></th>
<th>Read energy</th>
<th>Read time</th>
<th>Area overhead</th>
<th>Shifting error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional Read [163, 164, 165]</td>
<td>1.4 nJ/bit</td>
<td>47 ns</td>
<td>2x</td>
<td>Prone</td>
</tr>
<tr>
<td>Transverse Read (this work)</td>
<td>0.47 pJ/bit</td>
<td>34 ns</td>
<td>1x</td>
<td>Immune</td>
</tr>
</tbody>
</table>

Table 5.4 compares the conventional and the proposed transverse reading of a 15 bit DWM nanowire with PMA. We mimic the conventional reading mechanism reported in [163, 164, 165] to a NW with similar geometry used in previous study. The worth noting fact is since the traditional way of reading incorporates shifting, the energy requirement is 3000x higher and 1.38x slower than TR. Additionally, in conventional reading each bit is shifted right or left to bring under the read-head, therefore, it requires a two times longer nanowire to hold the value. Otherwise due the shifting, the data will be lost or corrupted. The following section\textsuperscript{5.2} describes the use of TR in developing error correction coding to solve position errors.

5.7 Transverse Error Correction Coding

The goal of using the transverse read is to enable efficient correction of domain-wall memory alignment faults [154, 166]. To accomplish this, we propose to utilize values stored in the padding bits to determine the position of the nanowire during the transverse read. We can build the nanowire with a fixed domain representing a ‘1’ on the right end of the tape and a fixed domain representing a ‘0’ on the left end of the tape. Thus, during left and right shifts we will shift ‘1’s and ‘0’s into the padding bits on the right and left sides of the tape, respectively. The number of ‘1’s actually

\textsuperscript{5.2}Section 5.7 is a work of Sebastián Ollivier. Permission attached in Appendix A.
indicates the position of the data within the nanowire. As a result, if an under- or over-shift fault occurs, the calculated number of ones will differ from the expected value. Using the difference from the expected value, we can then correct the error, and ultimately the position fault.

![Number of ones on the sides](image)

Figure 5.9: Number of ones on the sides for different positions.

In Fig. 5.9 where the data bits \(d_i\) are shown in blue and the data bit aligned with the access port is shown in navy (dark blue). The padding bits on the left side (purple domains) contain ‘0’\textquotesingle s and the right side (beige domains) contain ‘1’\textquotesingle s.

Consider the case where the racetrack begins in position 1 [Fig. 5.9a] and attempts to shift to the right by one position to match position 2 [Fig. 5.9b]. Prior to the read, we can calculate the total number of ones in the nanowire using two transverse reads and a standard read. After performing the transverse reads \(L = d_0 + d_1 + d_2 + d_3\) and \(R = d_3 + d_4 + 1 + 1 + 1\). The total number of ‘1’s \(TOT = L + R - d_3\), where \(d_3\) is accessed using the standard read method. After the shift, because the data is invariant, the total number of ‘1’s should decrease by one. In particular, \(L' = d_0 + d_1 + d_2\), \(R' = d_2 + d_3 + d_4 + 1 + 1\), and \(TOT' = L' + R' - d_2\). The classical read is performed first in order to, wherever possible, hide the latency of the transverse read verification. If an under-shift occurs, the tape will actually remain in position 1 such that \(L' = L, R' = R,\) and \(TOT' = TOT\). Because \(TOT'\) remains unchanged, we know that no shift actually occurred, which can be corrected by right shifting again. Alternatively, if an over-shift occurs, the tape will move to position 3 [Fig. 5.9c]. In this case \(L' = d_0 + d_1\), \(R' = d_1 + d_2 + d_3 + d_4 + 1\), and \(TOT' = L' + R' - d_1\) and the total number of ‘1’s decreases by two rather than one, and a left shift should correct the misalignment.
5.8 Conclusion

We propose a novel transverse read mechanism for domain wall memories, which provides more global information about the stored data in place of querying each bit independently. TRs can be leveraged in fault tolerance, processing in memory, and machine learning applications, among others. Transverse read offers no cost-extension necessary for holding the data during read process. Another major advantage over state-of-the-art reading is the immunity from shifting error. DWM “racetrack” memories suffer greatly from alignment faults during shift operation. We envision that error correction due to shifting failure can harness the TR read. The next chapter, we discuss a novel fault, known as ‘Pinning Fault’ during shifting occurred in domain wall memories due to fabrication imperfections.
Chapter 6: Pinning Fault in Domain Wall Memories

6.1 Introduction

Domain wall memories are different than other emerging memories mainly for two reasons: 1) it can store multiple data-bits e.g. 64-512 bits in one cell, and 2) inherent ability of bi-directional shifting. As stated in chapter 2 and 5, multiple bits in a domain wall nanowire share the same read and write ports. Hence, these two operations also enact shifting after writing and reading the desired bit. This makes the shift operation most critical, and any error while shifting has consequences in overall functionality of the nanowire. Moreover, shift-based reading and writing techniques require additional power and time. Though there have been significant efforts [167, 147] on reducing the power consumption and delay while writing and reading the data, a handful amount of works [153, 154, 168, 169, 170] focus on the reliability of the shift operation in domain wall memories. One of the reasons behind this less effort is the lack of error-modeling.

Due to the extreme scaling in order to achieve higher density, the probability of occurring a fault has been increasing. Apart from the existing faults [171], a novel type of fault, namely “Pinning Fault”, emerges while shifting. The underlying reason behind this emerging fault is the non-uniform pinning potential distribution created by notches with fabrication imperfection. This non-uniformity affects the critical current requirement for shifting successfully. Therefore, it is imperative to analyze the shift current variation due to the deformities of a notch. However, the fab data for studying the the process variation is not available in a great detail, thus, we adapted a model for deformities in a notch. This chapter discusses the modelling of geometric variation of a notch as well as the impacts on the critical shift current.
6.1.1 Faults During Shifting

One major fault in shift operation is the ‘position error’ modeled by Zhang et al. [171]. Two types of scenario can be happened in this type of fault, ‘stop-in-middle’ and ‘out-of-step’. In the first case, a domain wall is not pinned at a notch position, rather, stays at any part of the domain enclosed by two notches. Therefore, the intended domain to be read or written is not aligned with access ports. This fault is often know as ‘alignment fault’. On the other hand, an ‘out-of-step’ fault occurs when a domain wall shifts more than 1 bit. In that case, a wrong domain is under the access ports and a faulty read or write operation is performed.

Fig. 6.1 illustrates the ‘position errors’ while shifting a domain wall nanowire. For an example shown in Fig. 6.1a, the data stored in the nanowire are to be shifted one bit. In a case of successful shifting, the data look like that in Fig. 6.1b. However, if a ‘stop-in-middle’ fault occurs, domain walls will not be aligned with the notch locations as shown in Fig. 6.1c. In this case, an uncertain data will be read. On the other hand, Fig. 6.1d depicts an ‘out-of-step’ error where, the data shifted two bits instead of one bit. Therefore, a wrong data will be accessed.

Figure 6.1: Stored data prior shifting,(b) desired data bits after perfect shift operation,(c) "stop-in-middle" error, (d) an "out-of-step" error.

In this chapter, we have modeled a novel error that may occur during shifting: "Pinning Fault". A domain wall is pinned in pinning sites which are created by intentionally engineered notes at regular intervals. Any deformation of a notch creates non-uniform pinning strength, and at some
degree of deformity the pinning strength is high enough to depin the wall. In that case, that wall comes stuck in that position and does not move while apply a critical shift current. To understand the fault, we briefly discuss the energies of a domain wall and pinning potential in a notch.

![Figure 6.2: Phase boundary between transverse and vortex wall.](image)

### 6.2 Domain Wall and Pinning Sites

When two adjacent domain contain two different bits (one ‘0’ and one ‘1’), a reorientation of spins happen at the boundary. The region where the spins get reoriented is known as a domain wall. Based on the dimensions of the nanowire, two types of wall can be formed: 1) transverse wall (TW), 2) vortex wall (VW) [172]. Fig. 6.2 illustrates the phase boundary between transverse wall (TW) and vortex wall (VW) between two domains. Clearly a wider nanowire has the tendency to form a vortex domain-wall because the interplay between the quantum exchange energy exchange and magneto-static energy results in a curling magnetization with no remanence near zero field.

In this work, we have only considered the dimensions of nanowire that create transverse walls. However, a transverse wall can be of two types: 1) Néel wall, and 2) Bloch wall. The dimensions we chose in this study favors the Bloch walls since the nanowires are thinner. The width of a Bloch TW wall is determined by
\[ \Delta = \pi S \sqrt{\frac{2J}{K_u a}} \]  

(6.1)

where \( S \) is the dot product of two anti-aligned spins, \( K_U \) is the magneto-crystalline anisotropy, \( a \) is the lattice constant, and \( J \) is the exchange co-efficient.

6.2.1 Pinning of DW

As stated earlier, pinning sites are created along the nanowire by intentionally fabricated notches. Typically these notches are patterned after a certain distance. The purpose of patterning notches is twofold: 1) pinning the wall at certain locations so that DWs can shift in lockstep fashion, 2) in case of subsequent shifting, no domain wall is annihilated. The pinning potential depends on the dimensions of a notch, and be expressed by the exchange and anistropic energies per unit area as

\[ E_{DW} = E_{DW}^{ex} + E_{DW}^{anis} = JS^2 \frac{\pi^2}{a^2 N} + \frac{NK_u a}{2} \]  

(6.2)

A current pulse with adequate amplitude can depin the wall from the notch positions, and if depinned, a wall can travel along the nanowire till the next pinning site. The current requirement to shift the DWs depends on the dimensions, material properties of the nanowire as well as the notch dimensions. However, a process variation can generate non-uniform pinning potential at the notches, which can alter the current requirement for a successful shifting.

Figure 6.3: A domain wall is pinned at a triangular notch.

several literature \[173, 174, 175\] modeled the pinning potential as
The purpose of this study is to observe if there is any change of critical and upper-bound of shift current from the benchmark. We measured the critical current of $6.7 \times 10^7 A/m^2$ and the upper-bound current of $8.2 \times 10^7 A/m^2$. Thereafter, we wanted to create a distribution of these two currents based on the variation we considered.

Fig. 6.5 captures the normal distribution of notch width and depth of the deformed notch. It is worth mentioning that, these distributions are only applicable to the deformed notch. Rest of the notches had standard values.
6.3.1 Modeling the Deformed Side

As stated earlier, we considered one side of the triangular notch is not a straight line rather an arc. We modeled the curvature of the arc as a line segment of a circle where the length of the segment (S) matches with the length of the deformed side. The degree of curvature was modeled by the radius of the circle. Fig. 6.6 shows the model we used to model the curvature. We considered both the convex and concave arcs. The length of the arc is given by the Eq. 6.4.

\[ S = r\theta \]  \hspace{1cm} (6.4)

In case of a perfect notch, the radius will be close to infinity as

\[ r = \lim_{\theta \to 0} \frac{S}{2\theta} \]  \hspace{1cm} (6.5)

where S matches with the length of the side of a notch, r is the radius of the imaginary circles and \( \theta \) is the angle at the center created by the arc. Assuming the length of arc (S) is equal to the side of the triangle, a distribution of r was generated to capture the variation of curvatures.
6.4 Variation of Pinning Potential with Process Variation

Pinning potential in a notch location varies with the variation of dimensions and the shape of the notch as suggested by the Eq. 6.3.

Fig. 6.7 shows the radius if the virtual circle and corresponding angle to mimic the deformity different position for which the critical current is $6.4 \times 10^7 A/m^2$. Below this value, any wall does not move beyond the deformed notch creating a pinning fault. Considering all the variations, the pinning potential varies as shown in Fig. 6.8. Blue line and red line represent a triangular notch and an oval shaped notch.

Figure 6.7: Angle and radius of notch to achieve shift current for different positions.
6.5 Results and Discussion

In our experiment for a perfect domain wall nanowire, we considered triangular notches with a width of 50 nm and a depth (or height) of 30 nm. These notches are located at top edge of the nanowire along the length. Fig. 6.9 shows a typical domain wall nanowire with 4 domains separated by carefully engineered triangular notches.

![Figure 6.9: A domain wall nanowire with 4 domains.](image)

We simulated a similar nanowire with 16 domains. The nanowire is 3200 nm long with each domain of 200 nm length. The width and thickness of the nanowire are 100 nm and 4 nm. The material properties of the nanowire are listed in Table 6.1.

<table>
<thead>
<tr>
<th>$A_{ex}$ (J/m)</th>
<th>$M_s$ (A/m)</th>
<th>$\alpha$</th>
<th>$K_u 1$ (J/m$^3$)</th>
<th>current pulse width</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2.0 \times 10^{11}$</td>
<td>$6.5 \times 10^5$</td>
<td>0.02</td>
<td>$10^6$</td>
<td>0.5 ns</td>
</tr>
</tbody>
</table>
At the beginning we initialized the nanowire with random magnetization, and allowed the nanowire to relax. Thereafter we applied two consecutive current pulses of 0.5 ns with 3 ns interval, and subsequently tracked the position of a domain wall along the nanowire length. The 3 ns interval between two current pulses is the relaxation time after 1-bit shifting. Fig. 6.10b shows the current pulse applied to the nanowire in x direction. The displacement of the domain wall is depicted in Fig. 6.10a. The domain wall travelled \( \approx 400 \text{ nm} \) in two step of 1-bit shifting.

![Figure 6.10: Domain wall displacement with applied current pulses for shifting.](image)

We tuned the shift to find the critical current for shifting in this perfect scenario. For a nanowire containing 16 domains with no deformed triangular notches require a minimum current of \( 5.1 \times 10^7 \text{A/m}^2 \) to successfully shift one bit. We also calculated the upper-bound of shift current for which a domain wall shifts two bits at a time. The upper-bound current for this arrangement is \( 7.9 \times 10^7 \text{A/m}^2 \). We took these values as the benchmark for a nanowire with perfect notches.

During the experiment, we varied only a variable from depth, width and curvature while keeping the others fixed. Initially, we put the deformed notch at the center of the nanowire, and measure the benchmark currents. We observed a significant variation of currents because of the deformity. Later, we examined if there is any effect of notch position along the nanowire. The interesting observation is that deformity near the shift port from which the current enters incurs most variations.

Fig. 6.11 depicts the current variation with notch depth, width and curvature. current requirement for shifting is almost invariant width depth and width for the deformed notches that are not too close to the shift port. On the other hand, variation of critical current due to the cois significant in all notch position, and increase drastically.
Figure 6.11: Current variation with (a) notch width and depth, (b) curvature.

6.6 Conclusion

In this chapter, we measured the pinning fault by modeling the deformation of a notch. We distinguish the cases whether a pinning fault occurred or not by measuring the critical current for shifting. This model can be verified by fabricating the nanowires in future.
Chapter 7: Conclusion and Future Work

7.1 Synopsis

The unique properties of spintronic devices together with its growing success and the global investments have increased its potential to become a universal memory. In this dissertation we have studied how these devices are accessed (reading) and reconfigured (programmed) and the impact of process variation on these mechanisms. First, we have developed a reading mechanism to detect the output of a nanomagnet-based computing hardware. We have studied the constraints and challenges and devised three read techniques that can handle process variation. A variability tolerant differential read circuit for the architecture has been proposed. The read circuit works in two phases, the pre-charge phase followed by the sense phase. It compares the output state against its complement. To reduce variation effects in the read circuit we have introduced redundancy by reading a pair of outputs against a pre-fabricated nanomagnets. The read circuit is low power and is non-destructive with the cell states retained after read. We calculated the resistance variation due to the fabrication imperfection and designed a pre-amplifier to boost the sense margin in order to increase the accuracy of detection.

Second, we have presented how spin-orbital torque can counterbalance the dipolar coupling between two adjacent nanomagnets to program a magnetic grid. The programmed cells act as isolated cells and do not participate in the computation. We discovered the effect of the speed of the ramp of current on the programming. In the next chapter we developed a novel transverse read to gather the global information such as the number of ‘1’s stored in a domain wall memory without shifting the data. This method consumes very low power and immune to the shifting faults. Using the output from the transverse read we have demonstrated an error correction code to solve alignment faults.
Finally, we modeled the pinning fault in domain wall memories due to the deformity of the notches. We considered the variation in notch width, depth and curvatures of the side. Our detecting aspect is the critical current for shifting. If the current required for shifting the data is more than the critical current in case of a perfect nanowire, there is a chance of occurring a pinning fault.

7.2 Future Work

In future we would like to explore the read techniques that can have adaptable reference to read muti-state nanomagnets useful for the analog computation. In the programmability work, the additional terminal to provide SOT imposes additional challenges while integrating with CMOS technology. We would like to extend this work to find a more system-level solution. The transverse read method has the potential to be used in computational paradigm. So far, we have harnessed this for error correction scheme. We would like to explore various computing algorithms which can be benefited from this. We also would like to gain more insight of the pinning fault from fabrication perspective. In future we will study the scaling impact on the pinning fault.
References


S. Bhanja, D. Karunaratne, R. Panchumarthy, S. Rajaram, and S. Sarkar, “Experimental demonstration of non-boolean computing in computer vision with nanomagnets,”


Appendix A: Copyright Permissions

The following are Copyright permissions for use of materials in Chapters 2, 3, 4, 5, 6 and 7.

---

**Re: Copyright Info**

MC Mike Casey <mike.casey@nowpublishers.com>
To: Roxy, Kawisher

Yesterday at 4:31 AM

Hi Kawisher,

We allow all our authors to retain the copyright so there is no issue. I'd appreciate you citing the FnT issue obviously.

Regards, Mike

---

From: Roxy, Kawisher <kawisher@usf.edu>
Sent: Friday, November 20, 2020 7:34:21 PM
To: mike.casey@nowpublishers.com <mike.casey@nowpublishers.com>
Subject: Copyright Info

Hello,


I am wondering is there any copyright issue?

Best,
Kawisher Ahmed Roxy

---

**[EXTERNAL EMAIL]** DO NOT CLICK links or attachments unless you recognize the sender and know the content is safe.
Appendix A (Continued)

Reading Nanomagnetic Energy Minimizing Coprocessor

Author: Kawsher A. Roxy
Publication: Nanotechnology, IEEE Transactions on
Publisher: IEEE
Date: March 2018

Copyright © 2018, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author’s approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis online.
3) In placing the thesis on the author’s university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity’s name goes here]’s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.
Appendix A (Continued)

A Novel Transverse Read Technique for Domain-Wall “Racetrack” Memories

Author: Kawshir Roxy
Publication: Nanotechnology, IEEE Transactions on
Publisher: IEEE
Date: 2020

Copyright © 2020, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © [Year of original publication] IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author’s approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author’s university website, please display the following message in a prominent place on the website: in reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity’s name goes here]’s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.
Appendix A (Continued)

Variability tolerant reading of nanomagnetic energy minimizing co-processor

Conference Proceedings:
2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)

Author: Kawsheer A Roxy

Publisher: IEEE

Date: Aug. 2017

Copyright © 2017, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © [Year of original publication]

2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.

3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]

2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis online.

3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.
Appendix A (Continued)

Exploring the readability of nano-magnetic energy minimizing co-processor
Conference Proceedings:
2017 IEEE 17th International Conference on Nanotechnology (IEEE-NANO)
Author: Kawser A Roxy
Publisher: IEEE
Date: July 2017

Copyright © 2017, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © [Year of original publication]
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author’s approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis online.
3) In placing the thesis on the author’s university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity’s name goes here]’s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.
Leveraging Transverse Reads to Correct Alignment Faults in Domain Wall Memories

IEEE Requesting permission to reuse content from an IEEE publication

Conference Proceedings:
2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)
Author: Sébastien Ollivier
Publisher: IEEE
Date: Jun 2019

Copyright © 2019, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author’s approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.
About the Author

Kawsher Ahmed Roxy obtained his bachelor degree in electrical and electronics engineering from the Bangladesh University of Engineering and Technology (BUET), Bangladesh in 2014. He received Masters degree in electrical engineering from the University of South Florida, Tampa, Florida, USA in 2019. He joined the University of South Florida for his Ph.D. in 2015. His primary research interests were non-volatile memory, device modeling, and memory circuits.